ΑΡΙΣΤΟΤΕΛΕΙΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΘΕΣΣΑΛΟΝΙΚΗΣ

ΑΝΟΙΧΤΑ
ΑΚΑΔΗΜΑΙΚΑ
ΜΑΘΗΜΑΤΑ

Στατιστική

9ο Μάθημα: Εφαρμογές Στατιστικής ΙΙ: Στατιστικοί Έλεγχοι

Γεώργιος Μενεξές
Τμήμα Γεωπονίας

Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο
ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ
ΕΣΠΑ 2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ
Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

Άδειες Χρήσης

- Το παρόν εκπαιδευτικό υλικό υπόκειται σε άδειες χρήσης Creative Commons.
- Για εκπαιδευτικό υλικό, όπως εικόνες, που υπόκειται σε άλλου τύπου άδειας χρήσης, η άδεια χρήσης αναφέρεται ρητώς.

Χρηματοδότηση

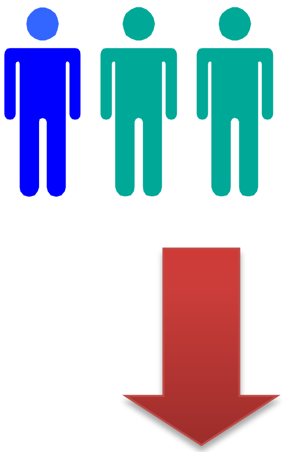
- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «Ανοικτά Ακαδημαϊκά Μαθήματα στο Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης» έχει χρηματοδοτήσει μόνο τη αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

9ο Μάθημα

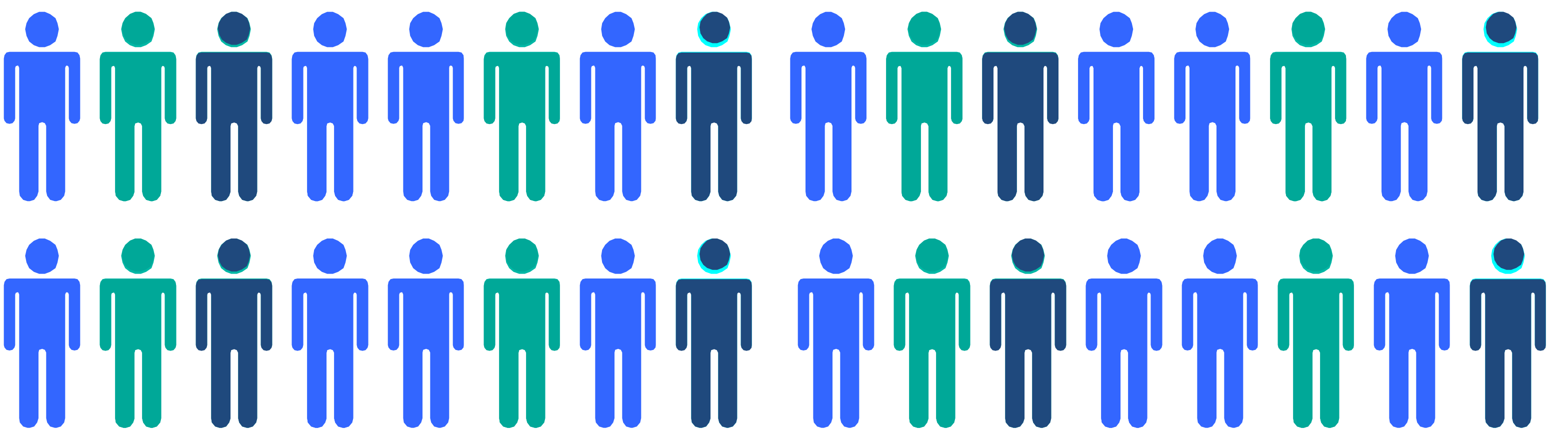
Εφαρμογές Στατιστικής ΙΙ: Στατιστικοί Έλεγχοι

Υπενθύμιση…



Από το δείγμα…

…γενικεύουμε για τον αντίστοιχο Πληθυσμό



Εφαρμογή 1

- Σε δείγμα 27 φυτών μηδικής ο μέσος όρος ύψους βρέθηκε ίσος με 84,8 cm με παραλλακτικότητα 93,90.

Α) Δεν έχουμε καμιά πληροφορία για το μέσο όρο μ του αντίστοιχου πληθυσμού μηδικής και θέλουμε να ελέγξουμε την υπόθεση ότι ο πραγματικός μέσος όρος μ είναι ίσος με 82 cm (σε ε.σ. α=0,05).

Β) Έχουμε λόγους να υποθέσουμε ότι ο πραγματικός μέσος όρος μ είναι μικρότερος από 84,8 και μάλιστα θέλουμε να ελέγξουμε την υπόθεση ότι είναι μικρότερος ή ίσος από 80 cm (σε ε.σ. α=0,05).

Γ) Έχουμε λόγους να υποθέσουμε ότι ο πραγματικός μέσος όρος μ είναι μεγαλύτερος από 84,8 και μάλιστα θέλουμε να ελέγξουμε την υπόθεση ότι είναι μεγαλύτερος ή ίσος από 90 cm (σε ε.σ. α=0,05).

Εφαρμογή 1 (*συνέχεια*)

- Το πρόβλημα ανάγεται στον Έλεγχο Υπόθεσης ότι ο μέσος όρος ενός πληθυσμού έχει μια συγκεκριμένη τιμή.
- Στην Α) περίπτωση ο έλεγχος είναι δίπλευρος ή αμφίπλευρος (*two-tailed*).
- Στις περιπτώσεις Β) και Γ) οι έλεγχοι είναι μονόπλευροι (*one-tailed*).

Εφαρμογή 1 (*συνέχεια*)

- Στην A) περίπτωση:

  H_0: μ=82

  H_1: μ≠82

  Σε ε.σ. *α*=0,05.

- Στη B) περίπτωση:

  H_0: μ=80 (ή μ≤80)

  H_1: μ>80

  Σε ε.σ. *α*=0,05.

- Στη Γ) περίπτωση:

  H_0: μ=90 (ή μ≥90)

  H_1: μ<90

  Σε ε.σ. *α*=0,05.

Δίπλευρος Έλεγχος

Μονόπλευροι Έλεγχοι

Εφαρμογή 1 (*συνέχεια*)

H_0	H_1	*Προϋποθέσεις*	*Απορρ. περιοχή R*	*Επεξηγήσεις*
$\mu=\mu_0$	$\mu>\mu_0$	σ^2 γνωστό	$R=\{z>z_a\}$	
		σ^2 άγνωστο $n\geq 30$	$R=\{t>z_a\}$	
B) περίπτωση		σ^2 άγνωστο $n<30$	$R=\{t>t_{n-1;\,a}\}$	όπου $z=\dfrac{(\bar{x}-\mu_0)\sqrt{n}}{\sigma}$
	$\mu<\mu_0$	σ^2 γνωστό	$R=\{z<-z_a\}$	
		σ^2 άγνωστο $n\geq 30$	$R=\{t<-z_a\}$	και $t=\dfrac{(\bar{x}-\mu_0)\sqrt{n}}{s}$
Γ) περίπτωση		σ^2 άγνωστο $n<30$	$R=\{t<-t_{n-1;\,a}\}$	
	$\mu\neq\mu_0$	σ^2 γνωστό	$R=\{\lvert z\rvert>z_{a/2}\}$	
		σ^2 άγνωστο $n\geq 30$	$R=\{\lvert t\rvert>z_{a/2}\}$	
A) περίπτωση		σ^2 άγνωστο $n<30$	$R=\{\lvert t\rvert>t_{n-1;\,a/2}\}$	

Πηγή:
Κολυβά & Μπόρα-Σέντα (1995)

Απάντηση

- Υπολογίζουμε το στατιστικό $t = \dfrac{\bar{Y} - \mu_0}{s/\sqrt{n}}$

 και για τις 3 περιπτώσεις (Α, Β και Γ)

$$t = \frac{84,8 - 82}{9,7/\sqrt{27}} = 1,501 \ (\text{Α } \textit{περίπτωση})$$

$$t = \frac{84,8 - 80}{9,7/\sqrt{27}} = 2,574 \ (\text{Β } \textit{περίπτωση})$$

$$t = \frac{84,8 - 90}{9,7/\sqrt{27}} = -2,778 \ (\text{Γ } \textit{περίπτωση})$$

Απάντηση (*συνέχεια*)

- Περίπτωση Α): Η απορριπτική περιοχή είναι:

$$R = \left\{ |t| > t_{n-1;a/2} \right\} \Rightarrow$$

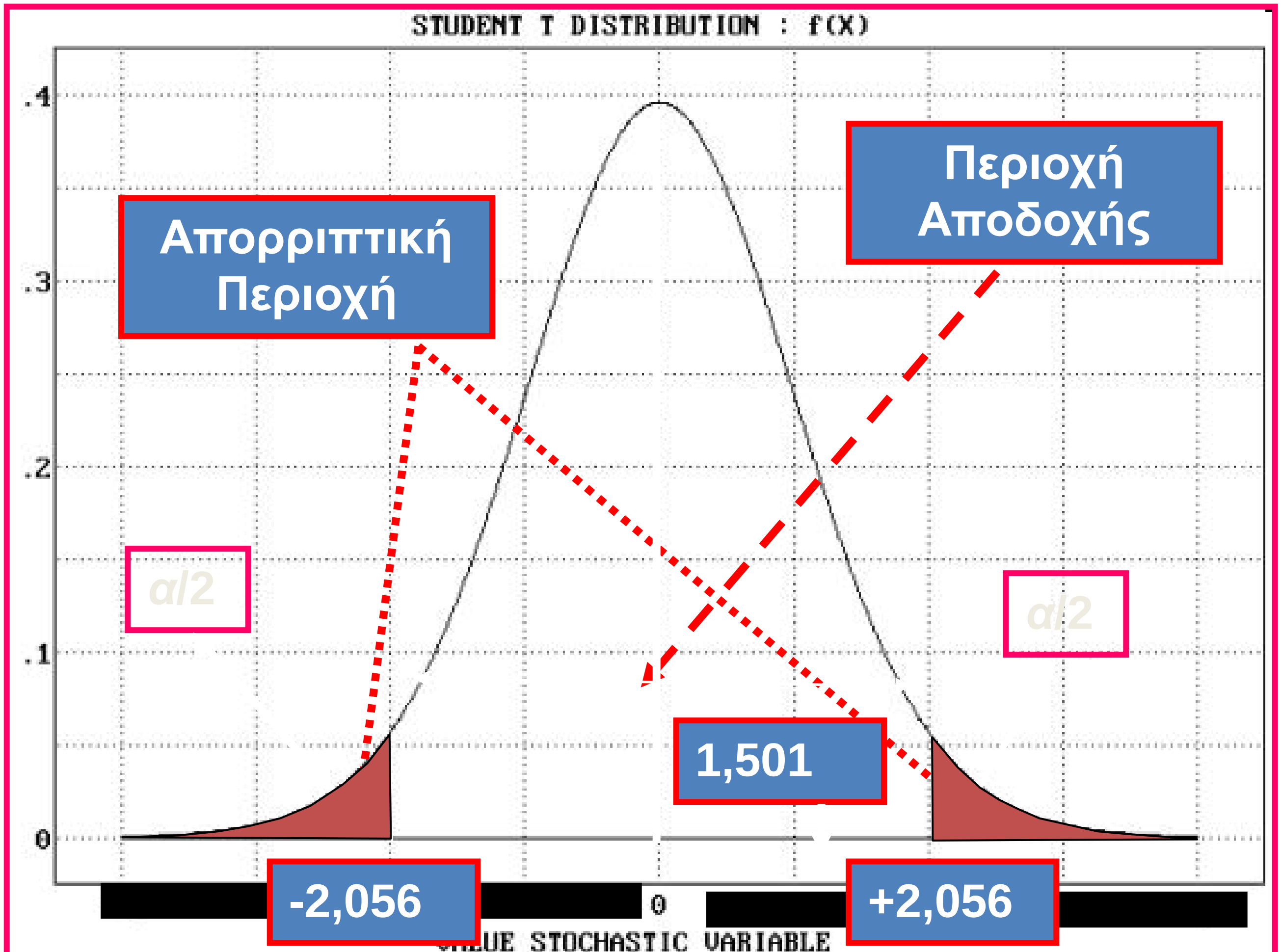
$$R = \left\{ |t| > t_{26;0,025} \right\} = \left\{ |t| > 2,056 \right\}$$

- Δηλ. περιλαμβάνει τις τιμές που είναι μικρότερες από -2,056 ή μεγαλύτερες από +2,056. Η τιμή t=1,501 που υπολογίσαμε από το δείγμα βρίσκεται μέσα στην περιοχή αποδοχής. Άρα η μηδενική υπόθεση δεν απορρίπτεται σε ε.σ. α=0,05.

Πίνακας Π.2 Κατανομή t

Βαθμοί ελευθε-ρίας	Πιθανότητα μιας απολύτως μεγαλύτερης τιμής (βλ. σχ. 4.8)								
	0,5	0,4	0,3	0,2	0,1	0,05	0,02	0,01	0,001
1	1,000	1,376	1,963	3,078	6,314	12,706	31,821	63,657	636.619
2	0,816	1,061	1,386	1886	2,920	4,303	6,965	9,925	31,598
3	0,765	0,978	1,250	1,638	2,353	3,182	4,541	5,841	12,941
4	0,741	0,941	1,190	1,533	2,132	2,776	3,747	4,604	8,610
5	0,727	0,920	1,156	1,476	2,015	2,571	3,365	4,032	6,859
6	0,718	0,906	1,134	1,440	1,943	2,447	3,143	3,707	5,959
7	0,711	0,896	1,119	1,415	1,895	2,365	2,998	3,499	5,405
8	0,706	0,889	1,108	1,397	1,860	2,306	2,896	3,355	5,041
9	0,703	0,883	1,100	1,383	1,833	2,262	2,821	3,250	4,781
10	0,700	0,879	1,093	1,372	1,812	2,228	2,764	3,169	4,587
11	0,697	0,876	1,088	1,363	1,796	2,201	2,718	3,106	4,437
12	0,695	0,873	1,083	1,356	1,782	2,179	2,681	3,055	4,318
13	0,694	0,870	1,079	1,350	1,771	2,160	2,650	3,012	4,221
14	0,692	0,868	1,076	1,345	1,761	2,145	2,624	2,977	4,140
15	0,691	0,866	1,074	1,341	1,753	2,131	2,602	2,947	4,073
16	0,690	0,865	1,071	1,337	1,746	2,120	2,583	2,921	4,015
17	0,689	0,863	1,069	1,333	1,740	2,110	2,567	2,898	3,965
18	0,688	0,862	1,067	1,330	1,734	2,101	2,552	2,878	3,922
19	0,688	0,861	1,066	1,328	1,729	2,093	2,539	2,861	3,883
20	0,687	0,860	1,064	1,325	1,725	2,086	2,528	2,845	3,850
21	0,686	0,859	1,063	1,323	1,721	2,080	2,518	2,831	3,819
22	0,686	0,858	1,061	1,321	1,717	2,074	2,508	2,819	3,792
23	0,685	0,858	1,060	1,319	1,714	2,069	2,500	2,807	3,767
24	0,685	0,857	1,059	1,318	1,711	2,064	2,492	2,797	3,745
25	0,684	0,856	1,058	1,316	1,708	2,060	2,485	2,787	3,725
26	0,684	0,856	1,058	1,315	1,706	2,056	2,479	2,779	3,707
27	0,684	0,855	1,057	1,314	1,703	2,052	2,473	2,771	3,690
28	0,683	0,855	1,056	1,313	1,701	2,048	2,467	2,763	3,674
29	0,683	0,854	1,055	1,311	1,699	2,045	2,462	2,756	3,659
30	0,683	0,854	1,055	1,310	1,697	2,042	2,457	2,750	3,646
40	0,681	0,851	1,050	1,303	1,684	2,021	2,423	2,704	3,551
60	0,679	0,848	1,046	1,296	1,671	2,000	2,390	2,660	3,460
120	0,677	0,845	1,041	1,289	1,658	1,980	2,358	2,617	3,373
∞	0,674	0,842	1,036	1,282	1,645	1,960	2,326	2,576	3,291

Περιοχή Αποδοχής και Απόρριψης



Απάντηση (*συνέχεια*)

- Περίπτωση Β): Η απορριπτική περιοχή είναι:

$$R = \left\{ t > t_{n-1;a} \right\} \Rightarrow$$

$$R = \left\{ t > t_{26;0,05} \right\} = \left\{ t > 1,706 \right\}$$

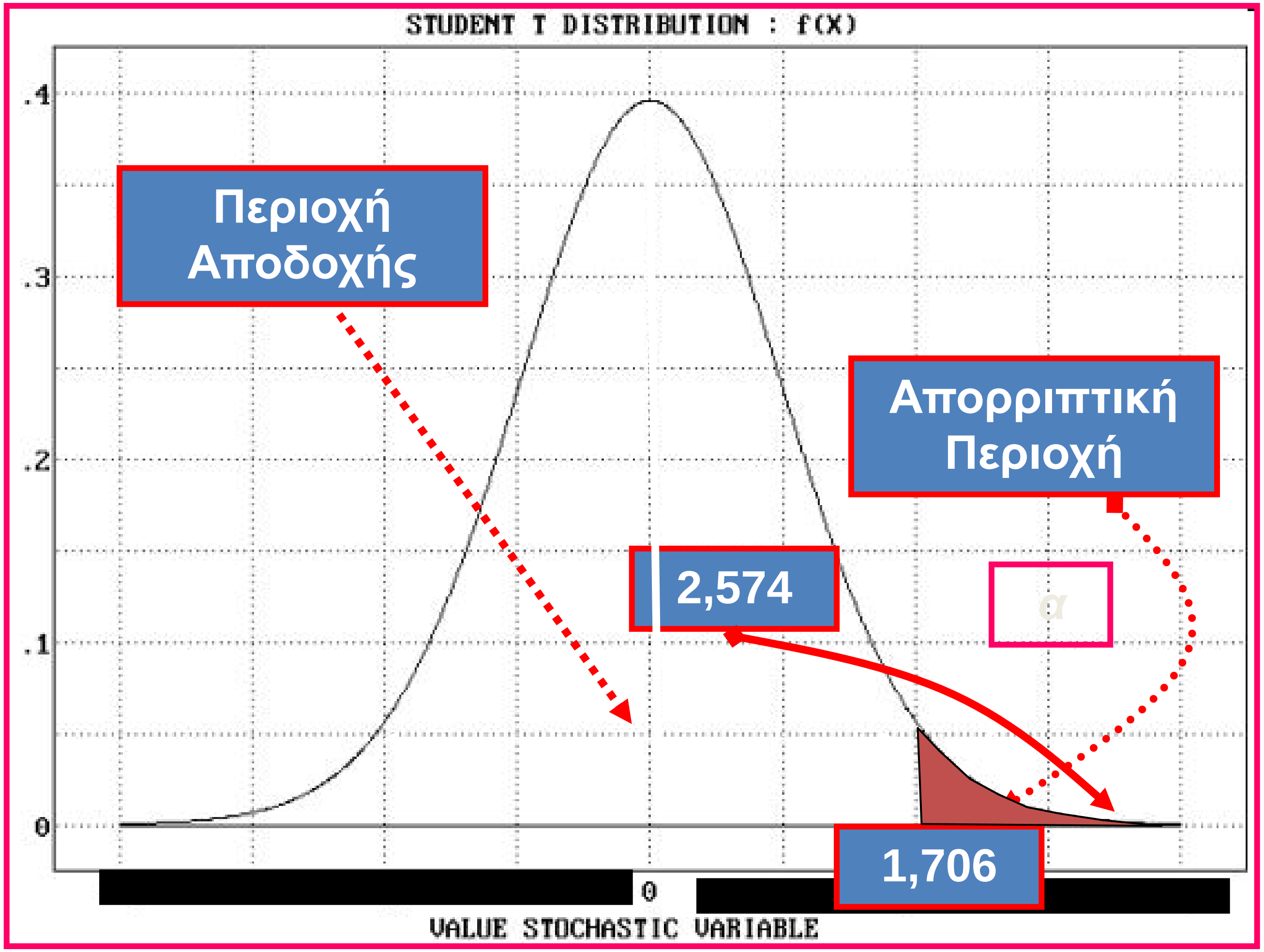
- Η περιοχή απόρριψης περιλαμβάνει την επιφάνεια μόνο στο δεξί άκρο (ουρά) της καμπύλης και πιο συγκεκριμένα τις τιμές που είναι μεγαλύτερες από 1,706. Η τιμή t=2,574 που υπολογίσαμε βρίσκεται μέσα στην απορριπτική περιοχή και συνεπώς η μηδενική υπόθεση απορρίπτεται σε ε.σ. α=0,05.

Πίνακας Π.2 Κατανομή t

Βαθμοί ελευθερίας	Πιθανότητα μιας απολύτως μεγαλύτερης τιμής (βλ. σχ. 4.8)								
	0,5	0,4	0,3	0,2	0,1	0,05	0,02	0,01	0,001
1	1,000	1,376	1,963	3,078	6,[illegible]14	12,706	31,821	63,657	636.619
2	0,816	1,061	1,386	1886	2,[illegible]20	4,303	6,965	9,925	31,598
3	0,765	0,978	1,250	1,638	2,[illegible]53	3,182	4,541	5,841	12,941
4	0,741	0,941	1,190	1,533	2,[illegible]32	2,776	3,747	4,604	8,610
5	0,727	0,920	1,156	1,476	2,[illegible]15	2,571	3,365	4,032	6,859
6	0,718	0,906	1,134	1,440	1,[illegible]43	2,447	3,143	3,707	5,959
7	0,711	0,896	1,119	1,415	1,[illegible]95	2,365	2,998	3,499	5,405
8	0,706	0,889	1,108	1,397	1,[illegible]60	2,306	2,896	3,355	5,041
9	0,703	0,883	1,100	1,383	1,[illegible]33	2,262	2,821	3,250	4,781
10	0,700	0,879	1,093	1,372	1,[illegible]12	2,228	2,764	3,169	4,587
11	0,697	0,876	1,088	1,363	1,796	2,201	2,718	3,106	4,437
12	0,695	0,873	1,083	1,356	1,782	2,179	2,681	3,055	4,318
13	0,694	0,870	1,079	1,350	1,771	2,160	2,650	3,012	4,221
14	0,692	0,868	1,076	1,345	1,761	2,145	2,624	2,977	4,140
15	0,691	0,866	1,074	1,341	1,753	2,131	2,602	2,947	4,073
16	0,690	0,865	1,071	1,337	1,746	2,120	2,583	2,921	4,015
17	0,689	0,863	1,069	1,333	1,740	2,110	2,567	2,898	3,965
18	0,688	0,862	1,067	1,330	1,734	2,101	2,552	2,878	3,922
19	0,688	0,861	1,066	1,328	1,729	2,093	2,539	2,861	3,883
20	0,687	0,860	1,064	1,325	1,725	2,086	2,528	2,845	3,850
21	0,686	0,859	1,063	1,323	1,721	2,080	2,518	2,831	3,819
22	0,686	0,858	1,061	1,321	1,717	2,074	2,508	2,819	3,792
23	0,685	0,858	1,060	1,319	1,714	2,069	2,500	2,807	3,767
24	0,685	0,857	1,059	1,318	1,711	2,064	2,492	2,797	3,745
25	0,684	0,856	1,058	1,316	[illegible]	2,060	2,485	2,787	3,725
26	0,684	0,856	1,058	1,315	1,706	2,056	2,479	2,779	3,707
27	0,684	0,855	1,057	1,314	1,703	2,052	2,473	2,771	3,690
28	0,683	0,855	1,056	1,313	1,701	2,048	2,467	2,763	3,674
29	0,683	0,854	1,055	1,311	1,699	2,045	2,462	2,756	3,659
30	0,683	0,854	1,055	1,310	1,697	2,042	2,457	2,750	3,646
40	0,681	0,851	1,050	1,303	1,684	2,021	2,423	2,704	3,551
60	0,679	0,848	1,046	1,296	1,671	2,000	2,390	2,660	3,460
120	0,677	0,845	1,041	1,289	1,658	1,980	2,358	2,617	3,373
∞	0,674	0,842	1,036	1,282	1,645	1,960	2,326	2,576	3,291

Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης

Περιοχή Αποδοχής και Απόρριψης



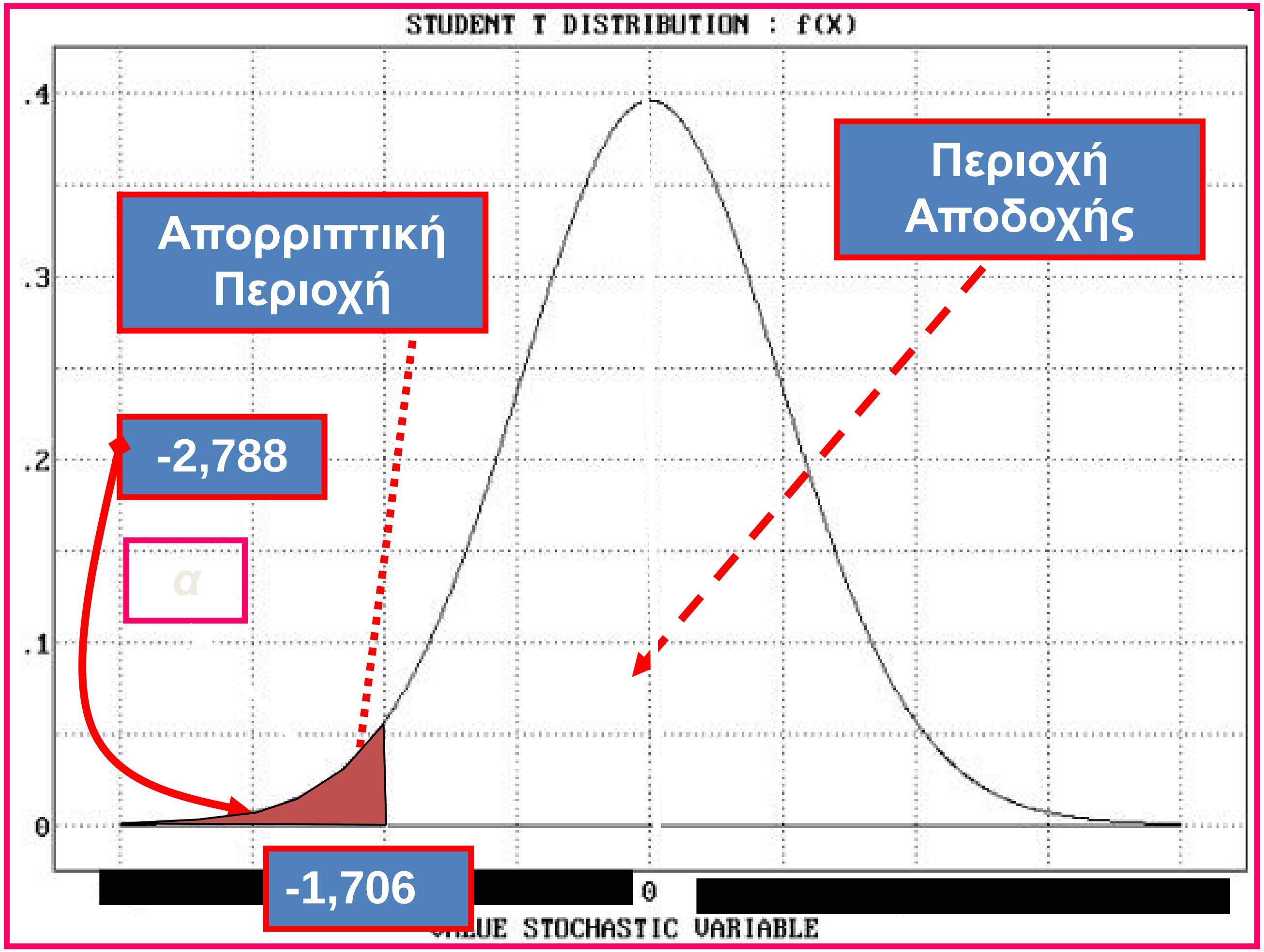
Απάντηση (*συνέχεια*)

- Περίπτωση Γ): Η απορριπτική περιοχή είναι:

$$R = \left\{ t < -t_{n-1;a} \right\} \Rightarrow$$
$$R = \left\{ t < -t_{26;0,05} \right\} = \left\{ t < -1,706 \right\}$$

- Η περιοχή απόρριψης περιλαμβάνει την επιφάνεια μόνο στο αριστερό άκρο (ουρά) της καμπύλης και πιο συγκεκριμένα τις τιμές που είναι μικρότερες από -1,706. Η τιμή *t*=-2,788 που υπολογίσαμε βρίσκεται μέσα στην απορριπτική περιοχή και συνεπώς η μηδενική υπόθεση απορρίπτεται σε ε.σ. *α*=0,05.

Περιοχή Αποδοχής και Απόρριψης



Εφαρμογή 2

- Ένας ερευνητής θέλει να ελέγξει αν μια νέα σειρά πειραματόζωων είναι πιο **ομοιόμορφη** από άποψη σωματικού βάρους σε σύγκριση με την εν χρήσει.
- Από ένα δείγμα 25 ζώων της νέας σειράς υπολογίζει την παραλλακτικότητα σε 750 (s^2) ενώ η εν χρήσει σειρά έχει παραλλακτικότητα 625 (σ^2). Ποιον στατιστικό έλεγχο θα πραγματοποιήσει ο ερευνητής;

Εφαρμογή 2 (*συνέχεια*)

- Το πρόβλημα ανάγεται στον Έλεγχο Υπόθεσης ότι η παραλλακτικότητα ενός πληθυσμού έχει μια συγκεκριμένη τιμή.
- Στη συγκεκριμένη περίπτωση είναι πιο λογικό να υποθέσουμε ότι ο έλεγχος θα αφορά στην ερευνητική υπόθεση ότι η νέα σειρά πειραματόζωων έχει παραλλακτικότητα μεγαλύτερη από 625.

Απάντηση

Ο στατιστικός έλεγχος που πρέπει να πραγματοποιηθεί είναι:

H_0: σ^2=625 (ή $\sigma^2 \leq 625$)

H_1: σ^2>625

Σε ε.σ. α=0,05.

Απάντηση (*συνέχεια*)

H_0	H_1			
$\sigma^2=\sigma_0^2$	$\sigma^2>\sigma_0^2$		$R=\{X^2>X^2_{n-1;\,a}\}$	όπου $X^2=\frac{(n-1)s^2}{\sigma_0^2}$
	$\sigma^2<\sigma_0^2$		$R=\{X^2<\mathscr{X}^2_{n-1;\,1-a}\}$	
	$\sigma^2\neq\sigma_0^2$		$R=\{X^2>X^2_{n-1;\,a/2}$ ή $X^2<\mathscr{X}^2_{n-1;\,1-a/2}\}$	

Υπολογίζουμε το στατιστικό:

$$X^2=\frac{(n-1)s^2}{\sigma_0^2}\Rightarrow$$

$$X^2=\frac{(25-1)750}{625}=28{,}8$$

Πηγή: Κολυβά & Μπόρα-Σέντα (1995)

Απάντηση (*συνέχεια*)

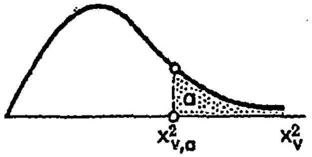
- Η απορριπτική περιοχή είναι:

$$R = \left\{ X^2 > X^2_{n-1;a} \right\} \Rightarrow$$

$$R = \left\{ X^2 > X^2_{24;0,05} \right\} = \left\{ X^2 > 36,42 \right\}$$

- Περιλαμβάνει το δεξί άκρο της X^2 Κατανομής, δηλ. τις τιμές που είναι μεγαλύτερες από 36,42. Η τιμή που υπολογίσαμε από το δείγμα 28,8 βρίσκεται μέσα στην περιοχή αποδοχής και επομένως η μηδενική υπόθεση δεν μπορεί να απορριφθεί σε ε.σ. *α*=0,05.

ΠΙΝΑΚΑΣ 7: Κριτικές τιμές της κατανομής χι-τετράγωνο



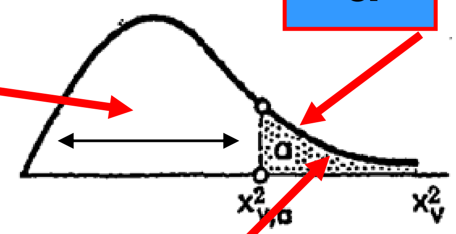
ν	α									
	.995	.990	.975	.950	.900	.100	.050	.025	.010	.005
1	0.0^4393	0.0^3157	0.0^3982	0.0^2393	0.0158	2.71	3.84	5.02	6.63	7.88
2	0.0100	0.0201	0.0506	0.103	0.211	4.61	5.99	7.38	9.21	10.60
3	0.072	0.115	0.216	0.352	0.584	6.25	7.81	9.35	11.34	12.84
4	0.207	0.297	0.484	0.711	1.064	7.78	9.49	11.14	13.28	14.86
5	0.412	0.554	0.831	1.145	1.61	9.24	11.07	12.83	15.09	16.75
6	0.676	0.872	1.24	1.64	2.20	10.64	12.59	14.45	16.81	18.55
7	0.989	1.24	1.69	2.17	2.83	12.02	14.07	16.01	18.48	20.28
8	1.34	1.65	2.18	2.73	3.49	13.36	15.51	17.53	20.09	21.96
9	1.73	2.09	2.70	3.33	4.17	14.68	16.92	19.02	21.67	23.59
10	2.16	2.56	3.25	3.94	4.87	15.99	18.31	20.48	23.21	25.19
11	2.60	3.05	3.82	4.57	5.58	17.28	19.68	21.92	24.73	26.76
12	3.07	3.57	4.40	5.23	6.30	18.55	21.03	23.34	26.22	28.30
13	3.57	4.11	5.01	5.89	7.04	19.81	22.36	24.74	27.69	29.82
14	4.07	4.66	5.63	6.57	7.79	21.06	23.68	26.12	29.14	31.32
15	4.60	5.23	6.26	7.26	8.55	22.31	25.00	27.49	30.58	32.80
16	5.14	5.81	6.91	7.96	9.31	23.54	26.30	28.85	32.00	34.27
17	5.70	6.41	7.56	8.67	10.09	24.77	27.59	30.19	33.41	35.72
18	6.26	7.01	8.23	9.39	10.86	25.99	28.87	31.53	34.81	37.16
19	6.84	7.63	8.91	10.12	11.65	27.20	30.14	32.85	36.19	38.58
20	7.43	8.26	9.59	10.85	12.44	28.41	31.41	34.17	37.57	40.00
21	8.03	8.90	10.28	11.59	13.24	29.62	32.67	35.48	38.93	41.40
22	8.64	9.54	10.98	12.34	14.04	30.81	33.92	36.78	40.29	42.80
23	9.26	10.20	11.69	13.09	14.85	32.01	35.17	38.08	41.64	44.18
24	9.89	10.86	12.40	13.85	15.66	33.20	36.42	39.36	42.98	45.56
25	10.52	11.52	13.12	14.61	16.47	34.38	37.65	40.65	44.31	46.93
26	11.16	12.20	13.84	15.38	17.29	35.56	38.89	41.92	45.64	48.29
27	11.81	12.88	14.57	16.15	18.11	36.74	40.11	43.19	46.96	49.64
28	12.46	13.56	15.31	16.93	18.94	37.92	41.34	44.46	48.28	50.99
29	13.12	14.26	16.05	17.71	19.77	39.09	42.56	45.72	49.59	52.34
30	13.79	14.95	16.79	18.49	20.60	40.26	43.77	46.98	50.89	53.67
40	20.71	22.16	24.43	26.51	29.05	51.81	55.76	59.34	63.69	66.77
50	27.99	29.71	32.36	34.76	37.69	63.17	67.50	71.42	76.15	79.49
60	35.53	37.48	40.48	43.19	46.46	74.40	79.08	83.30	88.38	91.95
70	43.28	45.44	48.76	51.74	55.33	85.53	90.53	95.02	100.4	104.2
80	51.17	53.54	57.15	60.39	64.28	96.58	101.9	106.6	112.3	116.3
90	59.20	61.75	65.65	69.13	73.29	107.6	113.1	118.1	124.1	128.3
100	67.33	70.06	74.22	77.93	82.36	118.5	124.3	129.6	135.8	140.2

Περιοχή Αποδοχής και Απόρριψης

Περιοχή Αποδοχής

α

747



Περιοχή Απόρριψης

Παρατηρήσεις

- Στην περίπτωση

  H_0: $\sigma^2=625$ (ή $\sigma^2 \geq 625$)

  H_1: $\sigma^2<625$

  Σε ε.σ. $\alpha=0{,}05$.

- Τότε η Απορριπτική Περιοχή είναι στο αριστερό άκρο της κατανομής X^2:

$$\boxed{R=\left\{ X^2 < X^2_{n-1;1-a} \right\}}$$

Παρατηρήσεις (*συνέχεια*)

- Στην περίπτωση

  H_0: $\sigma^2=625$

  H_1: $\sigma^2\neq 625$

  Σε ε.σ. $\alpha=0{,}05$.

- Τότε η Απορριπτική Περιοχή είναι:

$$\boxed{R=\left\{X^2 > X^2_{n-1;a/2} \ \acute{\eta}\ X^2 < X^2_{n-1;1-a/2}\right\}}$$

Περιλαμβάνει και τα δύο άκρα της κατανομής X^2.

Εφαρμογή 3

- Σύγκριση δύο παραλλακτικοτήτων σε ε.σ. *α* (έχουμε 3 περιπτώσεις):

$$H_0 : \sigma_1^2 = \sigma_2^2$$

$$H_1 : \sigma_1^2 \neq \sigma_2^2$$

$$H_1 : \sigma_1^2 > \sigma_2^2$$

$$H_1 : \sigma_1^2 < \sigma_2^2$$

Εφαρμογή 3 (*συνέχεια*)

$$s_1^2 = 77,91 \quad \mu\varepsilon\ n = 10$$

$$s_2^2 = 47,42 \quad \mu\varepsilon\ n = 28$$

$$\mathrm{H}_0 : \sigma_1^2 = \sigma_2^2$$

$$\mathrm{H}_1 : \sigma_1^2 > \sigma_2^2$$

$$\sigma\varepsilon\ \alpha = 0,05$$

Ο Στατιστικός Έλεγχος που μας ενδιαφέρει

Εφαρμογή 3 (*συνέχεια*)

H0	H1			
$\sigma_1^2=\sigma_2^2$	$\sigma_1^2>\sigma_2^2$	$s_1>s_2$	$R=\left\{\dfrac{s_1^2}{s_2^2}>F_{n-1,\,m-1;\,a}\right\}$	
ή $\dfrac{\sigma_1^2}{\sigma_2^2}=1$	$\sigma_1^2<\sigma_2^2$	$s_2>s_1$	$R=\left\{\dfrac{s_2^2}{s_1^2}>F_{m-1,\,n-1;\,a}\right\}$	
	$\sigma_1^2\neq\sigma_2^2$	$s_1^2\geq s_2^2$	$R=\left\{\dfrac{s_1^2}{s_2^2}\geq F_{n-1,\,m-1;\,a/2}\right\}$	
		$s_2^2\geq s_1^2$	$R=\left\{\dfrac{s_2^2}{s_1^2}\geq F_{m-1,\,n-1;\,a/2}\right\}$	

Πηγή: Κολυβά & Μπόρα-Σέντα (1995)

Απάντηση

- Υπολογίζουμε το λόγο:

$$F = \frac{s_1^2}{s_2^2} \Rightarrow \frac{77,91}{47,42} = 1,64$$

- Η Απορριπτική Περιοχή είναι:

$$R = \left\{ F = \frac{s_1^2}{s_2^2} > F_{n-1,m-1;a} \right\} \Rightarrow$$

$$R = \left\{ F > F_{9,27;0,05} \right\} = \left\{ F > 2,25 \right\}$$

Περιλαμβάνει το δεξί άκρο της κατανομής F.

Πίνακας 9γ

Τιμές της F κατανομής (α=0,05)

n_2	n_1: 1	2	3	4	5	6	7	8	9	10	11	12	14	16	20	24	30	40	50	75	100	200	500	∞
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,30	2,25	2,20	2,16	2,13	2,08	2,03	1,97	1,93	1,88	1,84	1,80	1,76	1,74	1,71	1,68	1,67
28	4,20	3,34	2,95	2,71	2,56	2,44	2,36	2,29	2,24	2,19	2,15	2,12	2,06	2,02	1,96	1,91	1,87	1,81	1,78	1,75	1,72	1,69	1,67	1,65
29	4,18	3,33	2,93	2,70	2,54	2,43	2,35	2,28	2,22	2,18	2,14	2,10	2,05	2,00	1,94	1,90	1,85	1,80	1,77	1,73	1,71	1,68	1,65	1,64
30	4,17	3,32	2,92	2,69	2,53	2,42	2,34	2,27	2,21	2,16	2,12	2,09	2,04	1,99	1,93	1,89	1,84	1,79	1,76	1,72	1,69	1,66	1,64	1,62
32	4,15	3,30	2,90	2,67	2,51	2,40	2,32	2,25	2,19	2,14	2,10	2,07	2,02	1,97	1,91	1,86	1,82	1,76	1,74	1,69	1,67	1,64	1,61	1,59
34	4,13	3,28	2,88	2,65	2,49	2,38	2,30	2,23	2,17	2,12	2,08	2,05	2,00	1,95	1,89	1,84	1,80	1,74	1,71	1,67	1,64	1,61	1,59	1,57
36	4,11	3,26	2,86	2,63	2,48	2,36	2,28	2,21	2,15	2,10	2,06	2,03	1,98	1,93	1,87	1,82	1,78	1,72	1,69	1,65	1,62	1,59	1,56	1,55
38	4,10	3,25	2,85	2,62	2,46	2,35	2,26	2,19	2,14	2,09	2,05	2,02	1,96	1,92	1,85	1,80	1,76	1,71	1,67	1,63	1,60	1,57	1,54	1,53
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,07	2,04	2,00	1,95	1,90	1,84	1,79	1,74	1,69	1,66	1,61	1,59	1,55	1,53	1,51
42	4,07	3,22	2,83	2,59	2,44	2,32	2,24	2,17	2,11	2,06	2,02	1,99	1,94	1,89	1,82	1,78	1,73	1,68	1,64	1,60	1,57	1,54	1,51	1,49
44	4,06	3,21	2,82	2,58	2,43	2,31	2,23	2,16	2,10	2,05	2,01	1,98	1,92	1,88	1,81	1,76	1,72	1,66	1,63	1,58	1,56	1,52	1,50	1,48
46	4,05	3,20	2,81	2,57	2,42	2,30	2,22	2,14	2,09	2,04	2,00	1,97	1,91	1.87	1,80	1,75	1,71	1,65	1,62	1,57	1,54	1,51	1,48	1,46
48	4,04	3,19	2,80	2,56	2,41	2,30	2,21	2,14	2,08	2,03	1,99	1,96	1,90	1,86	1,79	1,74	1,70	1,64	1,61	1,56	1,53	1,50	1,47	1,45
50	4,03	3,18	2,79	2,56	2,40	2,29	2,20	2,13	2,07	2,02	1,98	1,95	1,90	1,85	1,78	1,74	1,69	1,63	1,60	1,55	1,52	1,48	1,46	1,44
55	4,02	3,17	2,78	2,54	2,38	2,27	2,18	2,11	2,05	2,00	1,97	1,93	1,88	1,83	1,76	1,72	1,67	1,61	1,58	1,52	1,50	1,46	1,43	1,41
60	4,00	3,15	2,76	2,52	2,37	2,25	2,17	2,10	2,04	1,99	1,95	1,92	1,86	1,81	1,75	1,70	1,65	1,59	1,56	1,50	1,48	1,44	1,41	1,39
65	3,99	3,14	2,75	2,51	2,36	2,24	2,15	2,08	2,02	1,98	1,94	1,90	1,85	1,80	1,73	1,68	1,63	1,57	1,54	1,49	1,46	1,42	1,39	1,37
70	3,98	3,13	2,74	2,50	2,35	2,23	2,14	2,07	2,01	1,97	1,93	1,89	1,84	1,79	1,72	1,67	1,62	1,56	1,53	1,47	1,45	1,40	1,37	1,35
80	3,96	3,11	2,72	2,48	2,33	2,21	2,12	2,05	1,99	1,95	1,91	1,88	1,82	1,77	1,70	1,65	1,60	1,54	1,51	1,45	1,42	1,38	1,35	1,32
100	3,94	3,09	2,70	2,46	2,30	2,19	2,10	2,03	1,97	1,92	1,88		1,79	1,75	1,68	1,63	1,57	1,51	1,48	1,42	1,39	1,34	1,30	1,28
125	3,92	3,07	2,68	2,44	2,29	2,17	2,06	2,01	1,95	1,90	1,86	1,83	1,77	1,72	1,65	1,60	1,55	1,49	1,45	1,39	1,36	1,31	1,27	1,25
150	3,91	3,06	2,67	2,43	2,27	2,16	2,07	2,00	1,94	1,89	1,85	1,82	1,76	1,71	1,64	1,59	1,54	1,47	1,44	1,37	1,34	1,29	1,25	1,22
200	3,89	3,04	2,65	2,41	2,26	2,14	2,05	1,98	1,92	1,87	1,83	1,80	1,74	1,69	1,62	1,57	1,52	1,45	1,42	1,35	1,32	1,26	1,22	1,19
400	3,86	3,02	2,62	2,39	2,23	2,12	2,03	1,96	1,90	1,85	1,81	1,78	1,72	1,67	1,60	1,54	1,49	1,42	1,38	1,32	1,28	1,22	1,16	1,13
1000	3,85	3,00	2,61	2,38	2,22	2,10	2,02	1,95	1,89	1,84	1,80	1,76	1,70	1,65	1,58	1,53	1,47	1,41	1,36	1,30	1,26	1,19	1,13	1,08
∞	3,84	2,99	2,60	2,37	2,21	2,09	2,01	1,94	1,88	1,83	1,79	1,75	1,69	1,64	1,57	1,52	1,46	1,40	1,35	1,28	1,24	1,17	1,11	1,00

Απάντηση (*συνέχεια*)

- Το στατιστικό F που υπολογίσαμε από το δείγμα 1,64 είναι μικρότερο από το θεωρητικό 2,25 και συνεπώς βρίσκεται μέσα στην περιοχή αποδοχής της μηδενικής υπόθεσης.
- Άρα με βάση τα διαθέσιμα δεδομένα η μηδενική υπόθεση δεν μπορεί να απορριφθεί σε ε.σ. α=0,05.
- Οι δύο παραλλακτικότητες δεν διαφέρουν στατιστικά σημαντικά σε ε.σ. α=0,05.

Εφαρμογή 4

$$s_1^2 = 1,34 \quad \mu\varepsilon\ n = 17$$

$$s_2^2 = 29,60 \quad \mu\varepsilon\ n = 4$$

$$\mathrm{H}_0 : \sigma_1^2 = \sigma_2^2$$

$$\mathrm{H}_1 : \sigma_1^2 < \sigma_2^2$$

$$\sigma\varepsilon\ \alpha = 0,05$$

Ο Στατιστικός Έλεγχος που μας ενδιαφέρει

Εφαρμογή 4 (*συνέχεια*)

H_0	H_1			
$\sigma_1^2=\sigma_2^2$	$\sigma_1^2>\sigma_2^2$	$s_1>s_2$	$R=\left\{\frac{s_1^2}{s_2^2}>F_{n-1,\,m-1;\,a}\right\}$	
ή $\frac{\sigma_1^2}{\sigma_2^2}=1$	$\sigma_1^2<\sigma_2^2$	$s_2>s_1$	$R=\left\{\frac{s_2^2}{s_1^2}>F_{m-1,\,n-1;\,a}\right\}$	
	$\sigma_1^2\neq\sigma_2^2$	$s_1^2\geq s_2^2$	$R=\left\{\frac{s_1^2}{s_2^2}\geq F_{n-1,\,m-1;\,a/2}\right\}$	
		$s_2^2\geq s_1^2$	$R=\left\{\frac{s_2^2}{s_1^2}\geq F_{m-1,\,n-1;\,a/2}\right\}$	

Πηγή: Κολυβά & Μπόρα-Σέντα (1995)

Απάντηση

- Υπολογίζουμε το λόγο:

$$\boxed{F=\frac{s_2^2}{s_1^2}\Rightarrow\frac{29,60}{1,34}=22,09}$$

- Η Απορριπτική Περιοχή είναι:

$$R=\left\{F=\frac{s_2^2}{s_1^2}>F_{m-1,n-1;a}\right\}\Rightarrow$$

$$R=\left\{F>F_{3,16;0,05}\right\}=\left\{F>3,24\right\}$$

Περιλαμβάνει το δεξί άκρο της κατανομής F.

Απάντηση (*συνέχεια*)

- Το στατιστικό F που υπολογίσαμε από το δείγμα 22,09 είναι μεγαλύτερο από το θεωρητικό 3,24 και συνεπώς βρίσκεται μέσα στην περιοχή απόρριψης της μηδενικής υπόθεσης.
- Άρα με βάση τα διαθέσιμα δεδομένα η μηδενική υπόθεση απορρίπτεται σε ε.σ. α=0,05.
- Οι δύο παραλλακτικότητες διαφέρουν στατιστικά σημαντικά σε ε.σ. α=0,05. Στο δεύτερο πληθυσμό η παραλλακτικότητα είναι στατιστικά σημαντικά μεγαλύτερη.

Εφαρμογή 5

$$s_1^2 = 3,89 \quad \mu\varepsilon\ n = 9$$

$$s_2^2 = 8,19 \quad \mu\varepsilon\ n = 10$$

$$\mathrm{H}_0 : \sigma_1^2 = \sigma_2^2$$

$$\mathrm{H}_1 : \sigma_1^2 \neq \sigma_2^2$$

$$\sigma\varepsilon\ \alpha = 0,05$$

Ο Στατιστικός Έλεγχος που μας ενδιαφέρει

Εφαρμογή 5 (*συνέχεια*)

H_0	H_1			
$\sigma_1^2 = \sigma_2^2$	$\sigma_1^2 > \sigma_2^2$	$s_1 > s_2$	$R = \left\{ \frac{s_1^2}{s_2^2} > F_{n-1,\, m-1;\, a} \right\}$	
ή $\frac{\sigma_1^2}{\sigma_2^2} = 1$	$\sigma_1^2 < \sigma_2^2$	$s_2 > s_1$	$R = \left\{ \frac{s_2^2}{s_1^2} > F_{m-1,\, n-1;\, a} \right\}$	
	$\sigma_1^2 \neq \sigma_2^2$	$s_1^2 \geq s_2^2$	$R = \left\{ \frac{s_1^2}{s_2^2} \geq F_{n-1,\, m-1;\, a/2} \right\}$	
		$s_2^2 \geq s_1^2$	$R = \left\{ \frac{s_2^2}{s_1^2} \geq F_{m-1,\, n-1;\, a/2} \right\}$	

Πηγή: Κολυβά & Μπόρα-Σέντα (1995)

Απάντηση

- Υπολογίζουμε το λόγο:

$$\boxed{F = \frac{s_2^2}{s_1^2} \Rightarrow \frac{8,19}{3,89} = 2,10}$$

- Η Απορριπτική Περιοχή είναι:

$$R = \left\{ F = \frac{s_2^2}{s_1^2} > F_{m-1,n-1;a/2} \right\} \Rightarrow$$

$$R = \left\{ F > F_{9,8;0,025} \right\} = \left\{ F > 4,36 \right\}$$

Απάντηση (*συνέχεια*)

- Το στατιστικό F που υπολογίσαμε από το δείγμα 2,10 είναι μικρότερο από το θεωρητικό 4,36 και συνεπώς βρίσκεται μέσα στην περιοχή αποδοχής της μηδενικής υπόθεσης.
- Άρα με βάση τα διαθέσιμα δεδομένα η μηδενική υπόθεση δεν απορρίπτεται σε ε.σ. α=0,05.
- Οι δύο παραλλακτικότητες δεν διαφέρουν στατιστικά σημαντικά σε ε.σ. α=0,05.

Εφαρμογή 6

- Στον παρακάτω πίνακα δίνεται η περιεκτικότητα σε Ca (mg/100 cm^3) του γάλακτος που προσκομιζόταν σε ένα εργοστάσιο κατά τη θερινή ή χειμερινή περίοδο (οι τιμές προέρχονται από απλή τυχαία δειγματοληψία).

Θερινή Περίοδος	Χειμερινή Περίοδος
121,5	126,9
123,9	131,0
122,9	133,0
119,7	133,1
122,8	125,8
120,7	130,1
124,8	125,0
118,7	126,8
122,4	128,5
	129,0

Εφαρμογή 6 (*συνέχεια*)

- Υπάρχει διαφορά στην περιεκτικότητα Ca ανάμεσα στις δύο περιόδους;
- Σε ποια περίοδο η περιεκτικότητα σε Ca είναι μεγαλύτερη;
- Που μπορεί να οφείλεται η διαφορά;
- Η παρατηρούμενη διαφορά εκτός από στατιστική σημαντικότητα έχει και βιολογική σημαντικότητα;

Υπενθύμιση

Πηγή: Κολυβά & Μπόρα-Σέντα (1995)

Συνήθως δ=0

H_0	H_1	Προϋποθέσεις	Απορρ. περιοχή R	Επεξηγήσεις
$\mu_1-\mu_2=\delta$	$\mu_1-\mu_2>\delta$	σ_1^2, σ_2^2 γνωστά	$R=\{z>z_a\}$	$z=\dfrac{\bar{x}-\bar{y}-\delta}{\sqrt{\dfrac{\sigma_1^2}{n}+\dfrac{\sigma_2^2}{m}}}$ συνήθως δ=0
		σ_1^2, σ_2^2 άγνωστα n, m>30	$R=\{z>z_a\}$	$z=\dfrac{\bar{x}-\bar{y}-\delta}{\sqrt{\dfrac{s_1^2}{n}+\dfrac{s_2^2}{m}}}$
		$\sigma_1=\sigma_2=\sigma=$ άγνωστο n, m<30	$R=\{t>t_{n+m-2;\,a}\}$	$t=\dfrac{\bar{x}-\bar{y}-\delta}{s\sqrt{\dfrac{1}{n}+\dfrac{1}{m}}}$, $s=\sqrt{\dfrac{(n-1)s_1^2+(m-1)s_2^2}{n+m-2}}$
		$\sigma_1^2\neq\sigma_2^2$ άγνωστο n, m<30	$R=\{t>t_{v;\,a}\}$	$t=\dfrac{\bar{x}-\bar{y}-\delta}{\sqrt{\dfrac{s_1^2}{n}+\dfrac{s_2^2}{m}}}$, όταν n=m, $v=2(n-1)$ όταν $n\neq m$, $v=\dfrac{\left(\dfrac{s_1^2}{n}+\dfrac{s_2^2}{m}\right)^2}{\dfrac{(s_1^2/n)^2}{n-1}+\dfrac{(s_2^2/m)^2}{m-1}}$
		ζευγαρωτές παρατηρήσεις	$R=\left\{\dfrac{\bar{z}\sqrt{n}}{s_z}>t_{n-1;\,a}\right\}$	όπου $\bar{z}$ και s_z η μέση τιμή και η τυπική απόκλιση των x_i-y_i
	$\mu_1-\mu_2<\delta$	σ_1^2, σ_2^2 γνωστά	$R=\{z<-z_a\}$	$z=\dfrac{\bar{x}-\bar{y}-\delta}{\sqrt{\dfrac{\sigma_1^2}{n}+\dfrac{\sigma_2^2}{m}}}$ συνήθως δ=0

Υπενθύμιση (*συνέχεια*)

Πηγή: Κολυβά & Μπόρα-Σέντα (1995)

H_0	H_1	*Προϋποθέσεις*	*Απορρ. περιοχή R*	*Επεξηγήσεις*
		σ_1^2, σ_2^2 άγνωστα n, m>30	$R = \{z < -z_a\}$	$z = \dfrac{\bar{x}-\bar{y}-\delta}{\sqrt{\dfrac{s_1^2}{n}+\dfrac{s_2^2}{m}}}$
		$\sigma_1 = \sigma_2 = \sigma$ άγνωστο n, m<30	$R = \{t < -t_{n+m-2;\, a}\}$	$t = \dfrac{\bar{x}-\bar{y}-\delta}{s\sqrt{\dfrac{1}{n}+\dfrac{1}{m}}}$ και $s = \sqrt{\dfrac{(n-1)s_1^2+(m-1)s_2^2}{n+m-2}}$
		$\sigma_1^2 \neq \sigma_2^2$ άγνωστα n, m<30	$R = \{t < -t_{\nu;\, a}\}$	$t = \dfrac{\bar{x}-\bar{y}-\delta}{\sqrt{\dfrac{s_1^2}{n}+\dfrac{s_2^2}{m}}}$ όταν n=m, $\nu=2(n-1)$ όταν $n \neq m$, $\nu = \dfrac{\left(\dfrac{s_1^2}{n}+\dfrac{s_2^2}{m}\right)^2}{\dfrac{(s_1^2/n)^2}{n-1}+\dfrac{(s_2^2/m)^2}{m-1}}$
		ζευγαρωτές παρατη-ρήσεις	$R = \left\{\dfrac{\bar{z}\sqrt{n}}{s_z} < t_{n-1;\, a}\right\}$	όπου $\bar{z}$ και s_z η μέση τιμή και η τυπική απόκλιση των $x_i - y_i$
	$\mu_1-\mu_2 \neq \delta$	σ_1^2, σ_2^2 γνωστά	$R = \{\lvert z\rvert > z_{a/2}\}$	$z = \dfrac{\bar{x}-\bar{y}-\delta}{\sqrt{\dfrac{\sigma_1^2}{n}+\dfrac{\sigma_2^2}{m}}}$ συνήθως δ=0
		σ_1^2, σ_2^2 άγνωστα n, m>30	$R = \{\lvert z\rvert > z_{a/2}\}$	$z = \dfrac{\bar{x}-\bar{y}-\delta}{\sqrt{\dfrac{s_1^2}{n}+\dfrac{s_2^2}{m}}}$
		$\sigma_1 = \sigma_2 = \sigma$ άγνωστο n, m<30	$R = \{\lvert t\rvert > t_{n+m-2;\, a/2}\}$	$t = \dfrac{\bar{x}-\bar{y}-\delta}{s\sqrt{\dfrac{1}{n}+\dfrac{1}{m}}}$ $s = \sqrt{\dfrac{(n-1)s_1^2+(m-1)s_2^2}{n+m-2}}$

Υπενθύμιση (*συνέχεια*)

H_0	H_1	*Προϋποθέσεις*	*Απορρ. περιοχή R*	*Επεξηγήσεις*
		$\sigma_1^2 \neq \sigma_2^2$ άγνωστα n, m<30	$R = \{\lvert t\rvert > t_{v;\, a/2}\}$	$t = \dfrac{\bar{x}-\bar{y}-\delta}{\sqrt{\dfrac{s_1^2}{n}+\dfrac{s_2^2}{m}}}$, όταν n=m, $v=2(n-1)$ όταν $n \neq m$, $v = \dfrac{\left(\dfrac{s_1^2}{n}+\dfrac{s_2^2}{m}\right)^2}{\dfrac{(s_1^2/n)^2}{n-1}+\dfrac{(s_2^2/m)^2}{m-1}}$
		ζευγαρωτές παρατηρήσεις	$R = \left\{ \left\lvert \dfrac{\bar{z}\sqrt{n}}{s_z} \right\rvert > t_{n-1;\, a/2} \right\}$	όπου $\bar{z}$ και s_z η μέση τιμή και η τυπική απόκλιση των $x_i - y_i$

Πηγή: Κολυβά & Μπόρα-Σέντα (1995)

H_0: $\mu_1=\mu_2$

H_1: $\mu_1 \neq \mu_2$

Απάντηση

- Υπολογίζουμε από το δείγμα:

$$\boxed{\begin{aligned} &n_1 = 9, \quad n_2 = 10 \\ &\bar{Y}_1 = 121{,}9, \quad \bar{Y}_2 = 128{,}9 \\ &s_1^2 = 3{,}89, \quad s_2^2 = 8{,}19 \end{aligned}}$$

- Από την Εφαρμογή 5 έχουμε ότι οι δύο παραλλακτικότητες δεν διαφέρουν στατιστικά σημαντικά σε ε.σ. α=0,05 και συνεπώς έχει νόημα να υπολογίσουμε μια εκτίμηση της κοινής παραλλακτικότητας s^2 των δύο δειγμάτων.

Απάντηση (*συνέχεια*)

	$\mu_1-\mu_2 \neq \delta$	σ_1^2, σ_2^2 γνωστά	$R = \{\lvert z\rvert > z_{a/2}\}$	$z = \dfrac{\bar{x}-\bar{y}-\delta}{\sqrt{\dfrac{\sigma_1^2}{n}+\dfrac{\sigma_2^2}{m}}}$ συνήθως $\delta=0$
		σ_1^2, σ_2^2 άγνωστα n, m>30	$R = \{\lvert z\rvert > z_{a/2}\}$	$z = \dfrac{\bar{x}-\bar{y}-\delta}{\sqrt{\dfrac{s_1^2}{n}+\dfrac{s_2^2}{m}}}$
		$\sigma_1 = \sigma_2 = \sigma$ άγνωστο n, m<30	$R = \{\lvert t\rvert > t_{n+m-2;\, a/2}\}$	$t = \dfrac{\bar{x}-\bar{y}-\delta}{s\sqrt{\dfrac{1}{n}+\dfrac{1}{m}}}$ $s = \sqrt{\dfrac{(n-1)s_1^2+(m-1)s_2^2}{n+m-2}}$

Πηγή: Κολυβά & Μπόρα-Σέντα (1995)

H_0: $\mu_1=\mu_2$

H_1: $\mu_1\neq\mu_2$

Απάντηση (*συνέχεια*)

Απορριπτική Περιοχή

$$R = \left\{ |t| > t_{n+m-2;a/2} \right\} = \left\{ |t| > t_{17;0,025} = 2,110 \right\}$$

$$s = \sqrt{\frac{(n-1)s_1^2 + (m-1)s_2^2}{n+m-2}} \Rightarrow s = \sqrt{\frac{8 \times 3,89 + 9 \times 8,19}{9+10-2}} = \sqrt{6,17} = 2,48$$

$$t = \frac{\bar{Y}_1 - \bar{Y}_2}{s\sqrt{\frac{1}{n} + \frac{1}{m}}} \Rightarrow t = \frac{121,9 - 128,9}{2,48\sqrt{\frac{1}{9} + \frac{1}{10}}} = -6,140$$

Στατιστικό *t* του Ελέγχου

Κοινή Τυπική Απόκλιση

Απάντηση (συνέχεια)

$$s\sqrt{\frac{1}{n}+\frac{1}{m}}=\sqrt{\frac{s^2}{n}+\frac{s^2}{m}}=\sqrt{s^2\frac{n+m}{nm}}$$

η ποσότητα

$$\frac{s^2}{n}+\frac{s^2}{m}$$

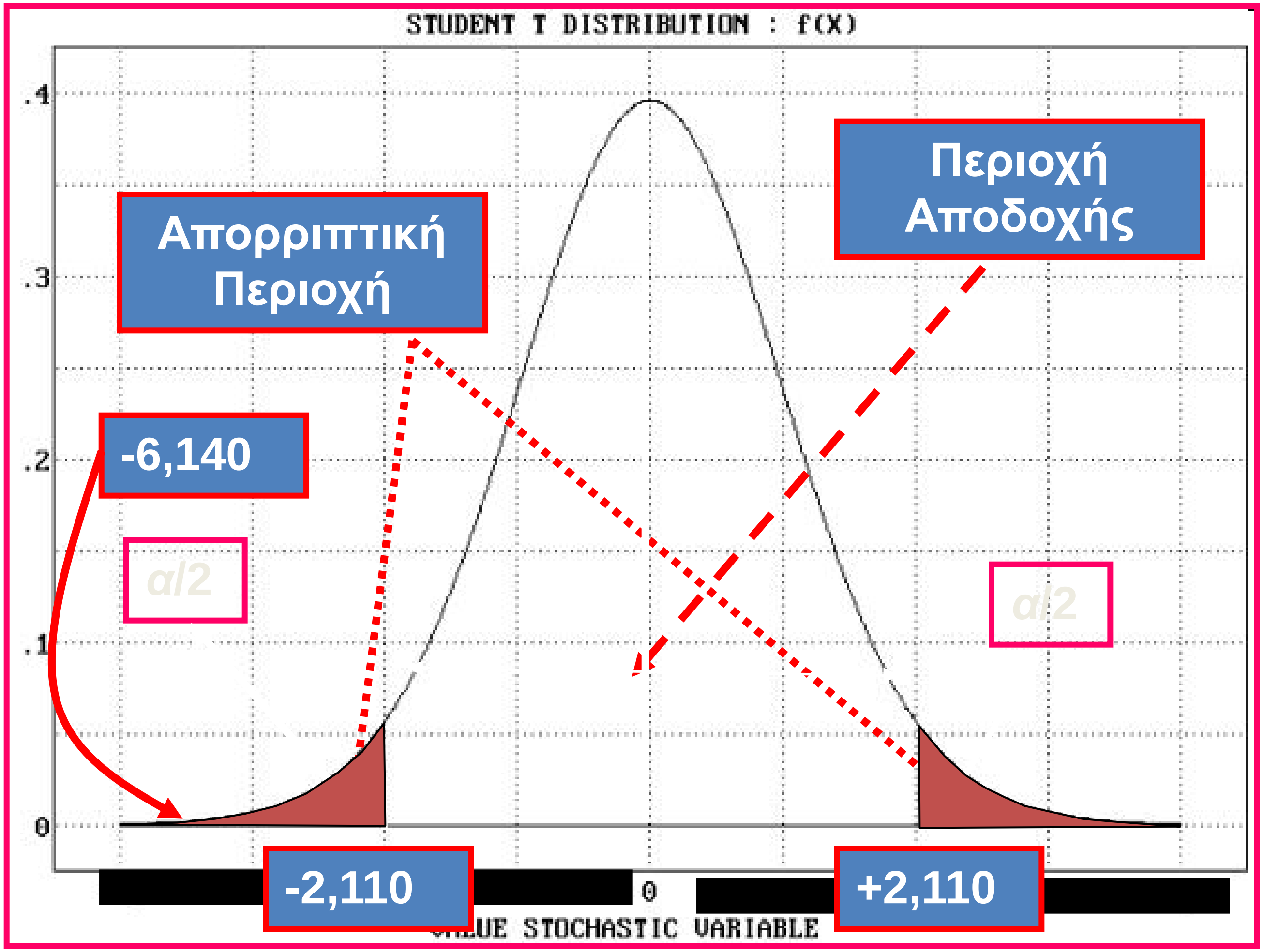
αποτελεί εκτίμηση της παραλλακτικότητας της διαφοράς των δύο μέσων όρων $\bar{Y}_1-\bar{Y}_2$ *και συμβολίζεται με:*

$$s^2_{\bar{Y}_1-\bar{Y}_2}=\frac{s^2}{n}+\frac{s^2}{m}=s^2\frac{n+m}{nm}$$

και το στατιστικό t γράφεται: $t=\dfrac{\bar{Y}_1-\bar{Y}_2}{s_{\bar{Y}_1-\bar{Y}}}$

Τυπικό Σφάλμα της Διαφοράς των δύο μέσων όρων

Περιοχή Αποδοχής και Απόρριψης



Απάντηση (*συνέχεια*)

- Η τιμή *t* που υπολογίστηκε από το δείγμα είναι -6,140 δηλ. πιο αριστερά από την κρίσιμη τιμή (θεωρητική) της *t* Κατανομής για 17 β.ε. σε ε.σ. α=0,05 η οποία είναι ίση -2,110.
- Συνεπώς, η τιμή *t* που υπολογίστηκε από το δείγμα βρίσκεται μέσα στην απορριπτική περιοχή. Άρα η μηδενική υπόθεση απορρίπτεται σε ε.σ. α=0,05.
- Οι δύο περίοδοι διαφέρουν στατιστικά σημαντικά σε ε.σ. α=0,05 σε ό,τι αφορά τη μέση περιεκτικότητα του γάλακτος σε Ca. Μεγαλύτερη (μέση) περιεκτικότητα σε Ca φαίνεται να υπάρχει τη χειμερινή περίοδο (γιατί;)

Απάντηση (*συνέχεια*)

- Ποιο είναι το τυπικό σφάλμα της διαφοράς των δύο μέσων όρων;

$$s^2_{\bar{Y}_1-\bar{Y}_2} = \frac{s^2}{n} + \frac{s^2}{m} = s^2 \frac{n+m}{nm} \Rightarrow$$

$$s^2_{\bar{Y}_1-\bar{Y}_2} = 6{,}17\frac{9+10}{9\times 10} = 1{,}30 \Rightarrow$$

$$s_{\bar{Y}_1-\bar{Y}_2} = \sqrt{1{,}30} = 1{,}14$$

Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης

Παρατήρηση

- Ένα Γενικό Συμπέρασμα:

$$\boxed{\Sigma\tau\alpha\tau\iota\sigma\tau\iota\kappa\acute{o}\ \mathrm{E}\lambda\acute{\varepsilon}\gamma\gamma\chi o\upsilon = \frac{\Delta\varepsilon\iota\gamma\mu\alpha\tau\iota\kappa\acute{\eta}\ \Delta\iota\alpha\varphi o\rho\acute{\alpha}}{\mathrm{T}\upsilon\pi\iota\kappa\acute{o}\ \Sigma\varphi\acute{\alpha}\lambda\mu\alpha\ \Delta\iota\alpha\varphi o\rho\acute{\alpha}\varsigma}}$$

Εφαρμογή 7

- Το ύψος των δένδρων ενός δάσους μετριέται συνήθως από το έδαφος. Η μέθοδος αυτή είναι ακριβής αλλά δαπανηρή.
- Σύμφωνα με μια άλλη μέθοδο το ύψος μπορεί να υπολογιστεί με βάση αεροφωτογραφίες.
- Για να μελετήσουμε τη δυνατότητα χρησιμοποίησης αεροφωτογραφιών, μετρήσαμε το ύψος 10 δένδρων και με τους δύο τρόπους και πήραμε τα παρακάτω αποτελέσματα:

Εφαρμογή 7 (*συνέχεια*)

	Ύψος δένδρων σε πόδια	
Α/Α Δένδρου	**Από το έδαφος**	**Από αεροφωτογραφία**
1	43	37
2	39	35
3	39	34
4	42	41
5	46	39
6	43	37
7	38	35
8	44	40
9	51	48
10	43	36

Μπορούμε να ισχυριστούμε ότι οι δύο μέθοδοι είναι ισοδύναμες σε ε.σ. α=0,05;

Απάντηση

- Πρόκειται για ζευγαρωτές παρατηρήσεις. Οι μετρήσεις και με τις δύο μεθόδους αφορούν στην ίδια πειραματική ή δειγματοληπτική μονάδα (το ίδιο δένδρο μετρήθηκε δύο φορές).

H_0	H_1	Προϋποθέσεις	Απορρ. περιοχή R	Επεξηγήσεις
		$\sigma_1^2 \neq \sigma_2^2$ άγνωστα n, m<30	$R = \{\lvert t \rvert > t_{\nu;\, a/2}\}$	$t = \dfrac{\bar{x}-\bar{y}-\delta}{\sqrt{\dfrac{s_1^2}{n}+\dfrac{s_2^2}{m}}}$, όταν n=m, $\nu=2(n-1)$ όταν $n \neq m$, $\nu = \dfrac{\left(\dfrac{s_1^2}{n}+\dfrac{s_2^2}{m}\right)^2}{\dfrac{(s_1^2/n)^2}{n-1}+\dfrac{(s_2^2/m)^2}{m-1}}$
		ζευγαρωτές παρατηρήσεις	$R = \left\{ \left\lvert \dfrac{\bar{z}\sqrt{n}}{s_z} \right\rvert > t_{n-1;\, a/2} \right\}$	όπου $\bar{z}$ και s_z η μέση τιμή και η τυπική απόκλιση των $x_i - y_i$

Πηγή: Κολυβά & Μπόρα-Σέντα (1995)

Απάντηση (*συνέχεια*)

- Ο Στατιστικός Έλεγχος είναι δίπλευρος:

$$\mathbf{H_0: \mu_1=\mu_2}$$

$$\mathbf{H_1: \mu_1 \neq \mu_2}$$

Σε ε.σ. α=0,05

Απάντηση (*συνέχεια*)

- Υπολογίζουμε τις διαφορές x_i-y_i

Α/Α Δένδρου	Από το έδαφος x_i	Από αεροφωτογραφία y_i	Διαφορές $x_i - y_i$
1	43	37	6
2	39	35	4
3	39	34	5
4	42	41	1
5	46	39	7
6	43	37	6
7	38	35	3
8	44	40	4
9	51	48	3
10	43	36	7

Απάντηση (*συνέχεια*)

- Υπολογίζουμε:

$$\overline{z} = 4,6$$
$$s_z^2 = 3,82$$
$$s_z = 1,96$$

- Απορριπτική Περιοχή του Ελέγχου:

$$R = \left\{ \left| \frac{\overline{z}\sqrt{n}}{s_z} \right| > t_{n-1;a/2} \right\} \Rightarrow$$
$$R = \left\{ \left| \frac{4,6\sqrt{10}}{1,96} \right| > t_{9;0,025} \right\}$$

$$\left| \frac{4,6\sqrt{10}}{1,96} \right| = 7,42 > t_{9;0,025} = 2,262$$

Απάντηση (*συνέχεια*)

- Η τιμή του στατιστικού ***t*** που υπολογίσαμε από το δείγμα **7,42** είναι μεγαλύτερη από την κρίσιμη (θεωρητική) τιμή **2,262** και συνεπώς η μηδενική υπόθεση απορρίπτεται σε ε.σ. α=0,05.
- Οι δύο μέθοδοι δεν είναι ισοδύναμες σε ε.σ. α=0,05.
- Οι δύο μέθοδοι διαφέρουν στατιστικά σημαντικά σε ε.σ. α=0,05.

Παρατηρήσεις

Ως ζευγαρωτές παρατηρήσεις μπορούν να θεωρηθούν οι μετρήσεις που γίνονται στις παρακάτω περιπτώσεις:

- Σε ζώα του ιδίου τοκετού.
- Σε γειτονικά εδαφοτεμάχια.
- Σε φυτά της ίδιας γλάστρας.
- Σε φύλλα του ιδίου φυτού.
- Από το ίδιο άτομο.
- Την ίδια μέρα.

Εφαρμογή 8

- Η απόδοση σε σύσπορο βαμβάκι (g/πειρ. τεμ.) τυχαίων γραμμών βαμβακιού ενός πληθυσμού μάρτυρα κι ενός από σπόρο ακτινοβολημένο με ακτίνες *γ* δίνεται στον πίνακα που ακολουθεί.
- Διαφέρουν τα δύο δείγματα ως προς τη μέση απόδοση σε ε.σ. α=0,05;

Εφαρμογή 8 (*συνέχεια*)

Μάρτυρας	Ακτινοβολημένο
980	1130
1310	1320
1090	490
1270	980
1240	760
1440	1020
1230	1290
1180	1250
1050	1120
1060	910
	530

Απάντηση

- Ο Στατιστικός Έλεγχος είναι δίπλευρος:

$$H_0: \mu_1=\mu_2$$

$$H_1: \mu_1 \neq \mu_2$$

Σε ε.σ. α=0,05

Απάντηση (*συνέχεια*)

- Υπολογίζουμε:

$$n_1 = 10 \qquad \overline{Y}_1 = 1.185,0 \qquad s_1^2 = 19.761,1$$

$$n_2 = 11 \qquad \overline{Y}_2 = 981,8 \qquad s_2^2 = 82.416,4$$

- Ελέγχουμε την ισότητα (ομοιογένεια) των παραλλακτικοτήτων των δύο πληθυσμών:

$$F = \frac{82.416,4}{19.761,1} = 4,17 > F_{10,9;0,025} = 3,96 \Rightarrow$$

Οι δύο παραλλακτικότητες διαφέρουν στατιστικά σημαντικά σε ε.σ. *α*=0,05

Απάντηση (*συνέχεια*)

H_0	H_1	Προϋποθέσεις	Απορρ. περιοχή R	Επεξηγήσεις
		$\sigma_1^2 \neq \sigma_2^2$ άγνωστα n, m<30	$R = \{\lvert t \rvert > t_{\nu;\, a/2}\}$	$t = \dfrac{\bar{x}-\bar{y}-\delta}{\sqrt{\dfrac{s_1^2}{n}+\dfrac{s_2^2}{m}}}$, όταν n=m, $\nu=2(n-1)$ όταν $n \neq m$, $\nu = \dfrac{\left(\dfrac{s_1^2}{n}+\dfrac{s_2^2}{m}\right)^2}{\dfrac{(s_1^2/n)^2}{n-1}+\dfrac{(s_2^2/m)^2}{m-1}}$
		ζευγαρωτές παρατηρήσεις	$R = \left\{ \left\lvert \dfrac{\bar{z}\sqrt{n}}{s_z} \right\rvert > t_{n-1;\, a/2} \right\}$	όπου $\bar{z}$ και s_z η μέση τιμή και η τυπική απόκλιση των $x_i - y_i$

Πηγή: Κολυβά & Μπόρα-Σέντα (1995)

Απάντηση (*συνέχεια*)

- Απορριπτική Περιοχή του Ελέγχου:

$$R = \left\{ |t| > t_{\nu;a/2} \right\}$$

$$t = \frac{\bar{Y}_1 - \bar{Y}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{\bar{Y}_1 - \bar{Y}_2}{s_{\bar{Y}_1 - \bar{Y}_2}}$$

$$\text{Αν} \quad n_1 = n_2 = n \quad \tau\acute{o}\tau\varepsilon \quad \nu = 2(n-1)$$

$$\text{Αν} \quad n_1 \neq n_2 \quad \tau\acute{o}\tau\varepsilon$$

$$\nu = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2} \right)^2}{\frac{\left(s_1^2 / n_1\right)^2}{n_1 - 1} + \frac{\left(s_2^2 / n_2\right)^2}{n_2 - 1}}$$

Απάντηση (*συνέχεια*)

- Κάνουμε υπολογισμούς:

Τυπικό Σφάλμα Διαφοράς των δύο μέσων όρων

$$s_{\bar{Y}_1-\bar{Y}_2} = \sqrt{\frac{s_1^2}{n} + \frac{s_2^2}{m}} \Rightarrow s_{\bar{Y}_1-\bar{Y}_2} = \sqrt{\frac{19.761{,}1}{10} + \frac{82.416{,}4}{11}} = \sqrt{9.468{,}5} = 97{,}3$$

$$t = \frac{1.185{,}0 - 981{,}8}{97{,}3} = 2{,}088$$

Βαθμοί Ελευθερίας

$$\nu = \frac{(19.761{,}1/10 + 82.416{,}4/11)^2}{\left[(19.761{,}1/10)^2/9\right] + \left[(82.416{,}4/11)^2/10\right]} = 14{,}8 \approx 15$$

Απάντηση (*συνέχεια*)

- Η κρίσιμη τιμή της $\boxed{t_{15;0,025} = 2,131}$
- Η τιμή του *t* που υπολογίσαμε από το δείγμα 2,088 είναι μικρότερη από την κρίσιμη-θεωρητική της *t* Κατανομής, δηλαδή πέφτει στην Περιοχή Αποδοχής του ελέγχου.
- Συνεπώς με βάση τα διαθέσιμα δεδομένα η μηδενική υπόθεση παραμένει (δεν απορρίπτεται) σε ε.σ. α=0,05.
- Τα δύο δείγματα δεν διαφέρουν στατιστικά σημαντικά ως προς τη μέση απόδοση σε ε.σ. α=0,05.

Απάντηση (*συνέχεια*)

- Ο μάρτυρας στο δείγμα φαίνεται να δίνει μεγαλύτερη απόδοση, αλλά η διαφορά δεν είναι αρκούντος μεγάλη ώστε να θεωρηθεί ως στατιστικά σημαντική σε ε.σ. α=0,05.
- Αυτή η διαφορά, με βάση τα διαθέσιμα δεδομένα, δεν φαίνεται να είναι συστηματική και θα μπορούσε να οφείλεται στην επίδραση τυχαίων παραγόντων και σφαλμάτων (τυχαία αίτια).

Παρατηρήσεις

- Στις Εφαρμογές 6 και 8 χρησιμοποιήθηκε το *t-test* για ανεξάρτητα δείγματα.
- Στην Εφαρμογή 7 χρησιμοποιήθηκε το *t-test* για ζευγαρωτές παρατηρήσεις.
- Στην Εφαρμογή 8 το στατιστικό ***t*** του ελέγχου ακολουθεί κατά προσέγγιση την *t*-Κατανομή.
- Αν οι μετρήσεις είναι στην πραγματικότητα ανεξάρτητες και εμείς τις θεωρήσουμε ως εξαρτημένες (ζευγαρωτές) τότε η παραδοχή αυτή μικρή επίδραση έχει στο επαγωγικό συμπέρασμα. Το αντίθετο, είναι "τραγικό" και η επίδρασή του είναι προς απρόβλεπτες κατευθύνσεις.

Προϋποθέσεις Εφαρμογής των Ελέγχων

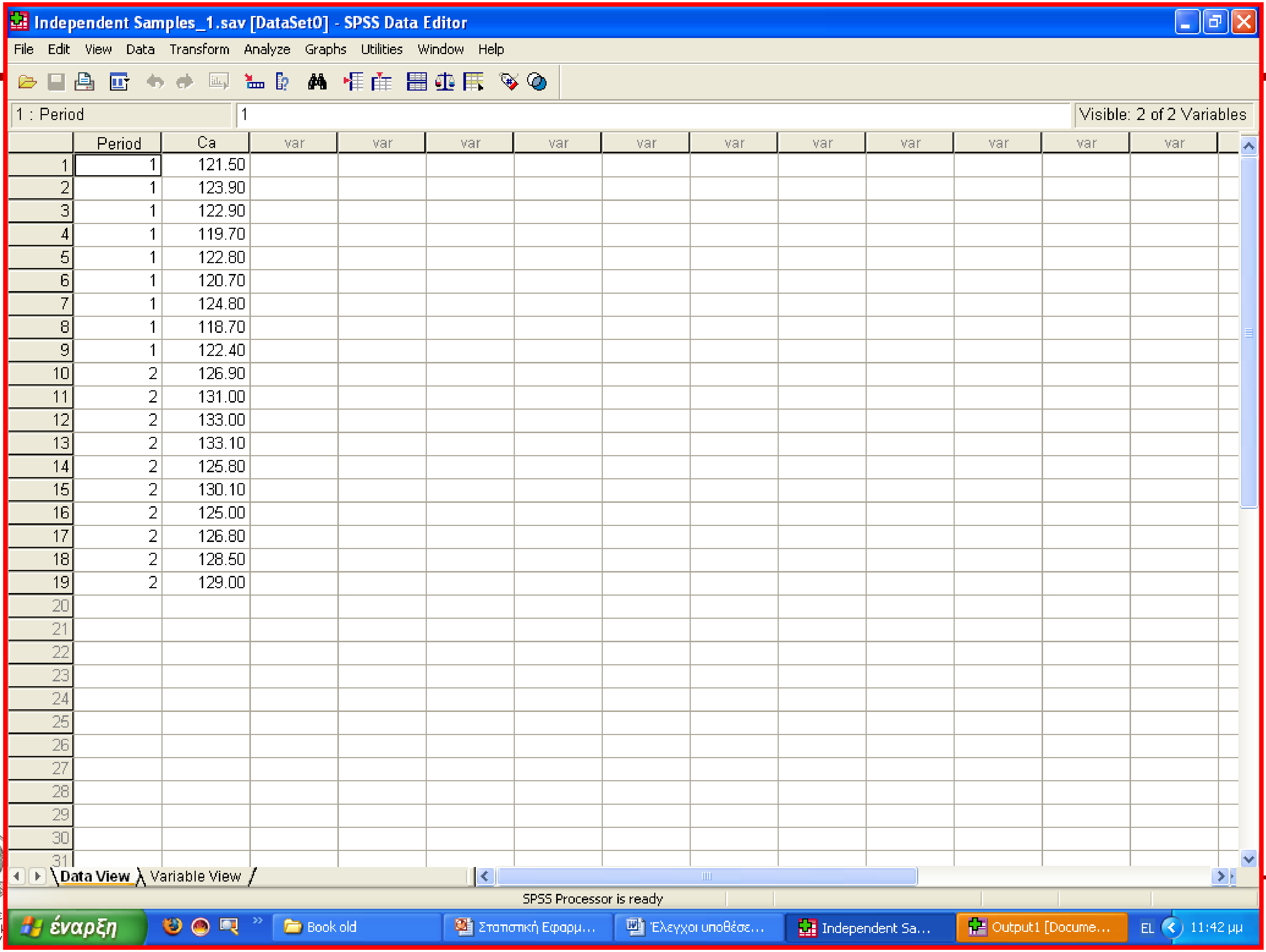
- Οι μετρήσεις προέρχονται από τυχαία δείγματα.
- Οι μετρήσεις είναι ανεξάρτητες μέσα σε κάθε δείγμα.
- Οι μετρήσεις προέρχονται από Κανονικές Κατανομές (εξαίρεση το *t-test* για ζευγαρωτές παρατηρήσεις, όπου οι διαφορές θα πρέπει να ακολουθούν Κανονική Κατανομή)

Γενικές Παρατηρήσεις

- Οι μονόπλευροι έλεγχοι δεν συνιστώνται πλέον.
- Σπάνια καταφεύγουμε στους ελέγχους υποθέσεων για συγκεκριμένες τιμές παραμέτρων αφού μπορούμε να έχουμε περισσότερη πληροφορία με τα αντίστοιχα
  (1-α)% διαστήματα εμπιστοσύνης.
- Υπάρχει στενή σχέση ανάμεσα στους ελέγχους υποθέσεων και στα διαστήματα εμπιστοσύνης.

Παραδείγματα με το SPSS

Εφαρμογή 6



	Period	Ca
1	1	121.50
2	1	123.90
3	1	122.90
4	1	119.70
5	1	122.80
6	1	120.70
7	1	124.80
8	1	118.70
9	1	122.40
10	2	126.90
11	2	131.00
12	2	133.00
13	2	133.10
14	2	125.80
15	2	130.10
16	2	125.00
17	2	126.80
18	2	128.50
19	2	129.00

Εφαρμογή 6

Group Statistics

	Περίοδος	N	Mean	Std. Deviation	Std. Error Mean
Συγκέν ρωση Ca	Θεριν ή	9	121.9333	1.97294	.65765
	Χειμεριν ή	10	128.9200	2.86155	.90490

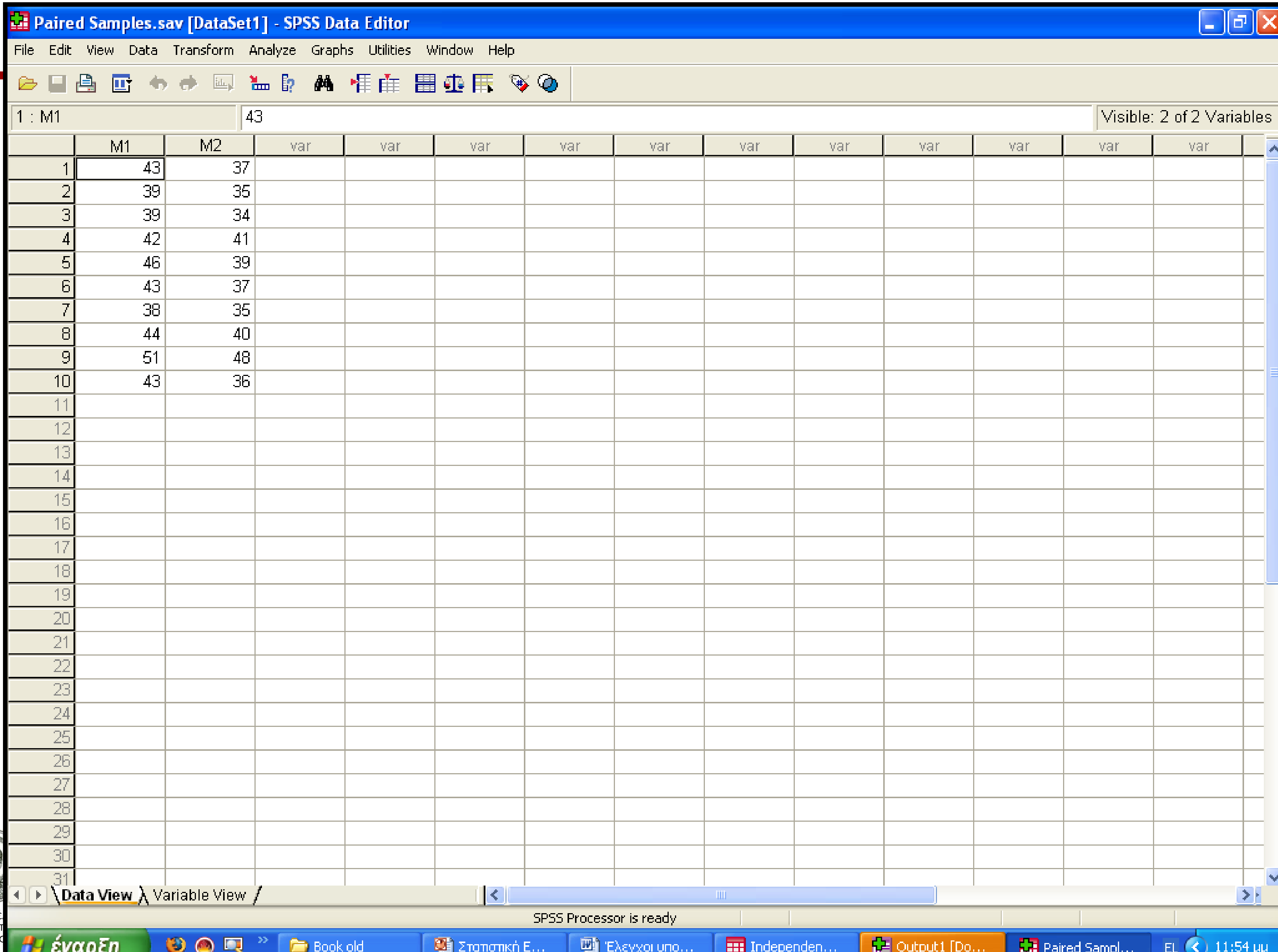
p-value

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Συγκέν ρωση Ca	Equal variances assumed	1.531	.233	-6.123	17	.000	-6.98667	1.14100	-9.39397	-4.57936
	Equal variances not assumed			-6.246	15.997	.000	-6.98667	1.11864	-9.35810	-4.61523

Αν $p<\alpha=0,05$ τότε Απορρίπτεται η Αντίστοιχη Μηδενική Υπόθεση. Αν $p\geq\alpha=0,05$ η H_0 παραμένει σε ε.σ. $\alpha=0,05$

Εφαρμογή 7



	M1	M2
1	43	37
2	39	35
3	39	34
4	42	41
5	46	39
6	43	37
7	38	35
8	44	40
9	51	48
10	43	36

Εφαρμογή 7

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Από το έδαφος	42.80	10	3.824	1.209
	Από αέρα	38.20	10	4.131	1.306

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Από το έδαφος & Από αέρα	10	.882	.001

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	95% Confidence Interval of the Difference Upper			
Pair 1	Από το έδαφος - Από αέρα	4.600	1.955	.618	3.201	5.999	7.440	9	.000

Η Μηδενική Υπόθεση Απορρίπτεται

Εφαρμογή 8

Independent Samples_2.sav [DataSet2] - SPSS Data Editor

	Group	Yield
1	1	980.00
2	1	1310.00
3	1	1090.00
4	1	1270.00
5	1	1240.00
6	1	1440.00
7	1	1230.00
8	1	1180.00
9	1	1050.00
10	1	1060.00
11	2	1130.00
12	2	1320.00
13	2	490.00
14	2	980.00
15	2	760.00
16	2	1020.00
17	2	1290.00
18	2	1250.00
19	2	1120.00
20	2	910.00
21	2	530.00

Εφαρμογή 8

Group Statistics

	Ομάδα	N	Mean	Std. Deviation	Std. Error Mean
Απόδοση	Μάρτυρας	10	1185.0000	140.57422	44.45347
	Ακτινοβολημένο	11	981.8182	287.08250	86.55863

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Απόδοση	Equal variances assumed	3.971	.045	2.025	19	.057	203.18182	100.33982	-6.83184	413.19548
	Equal variances not assumed			2.088	14.825	.054	203.18182	97.30626	-4.43529	410.79893

Η Μηδενική Υπόθεση Απορρίπτεται

Η Μηδενική Υπόθεση Δεν Απορρίπτεται

Εφαρμογή 8 (plus)

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Απόδοση	21	1078.5714	246.80530	53.85733

One-Sample Test

	Test Value = 1200					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
Απόδοση	-2.255	20	.036	-121.42857	-233.7730	-9.0841

Η Μηδενική Υπόθεση Απορρίπτεται

Εφαρμογή 8 (plus)

One-Sample Test

	Test Value = 1100					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference Lower	Upper
Απόδοση	-.398	20	.695	-21.42857	-133.7730	90.9159

Η Μηδενική Υπόθεση Δεν Απορρίπτεται

Παρατηρήσεις

- Στην Εφαρμογή 8 (plus) ελέγχθηκε αν η πραγματική μέση απόδοση **μ** (και για τα δύο δείγματα) είναι ίση με 1.200 και 1.100 αντίστοιχα. Στην πρώτη περίπτωση η μηδενική υπόθεση απορρίφθηκε ενώ στη δεύτερη όχι σε ε.σ. α=0,05.
- Το test του *Levene* είναι ένας άλλος έλεγχος ομοιογένειας-ισότητας παραλλακτικοτήτων.
- Η τιμή **p** (Sig.) εκφράζει την παρατηρούμενη στάθμη σημαντικότητας του αντίστοιχου στατιστικού ελέγχου.

Εισαγωγή στη Μη Παραμετρική Στατιστική

Μη Παραμετρικές Μέθοδοι

- Οι έλεγχοι υποθέσεων που εξετάσαμε στα προηγούμενα αφορούσαν παραμέτρους γνωστών κατανομών και κυρίως της Κανονικής Κατανομής.
- Συχνά τα δεδομένα μας δεν ακολουθούν καμιά από τις γνωστές κατανομές.
- Για την ανάλυση τέτοιων δεδομένων έχουν αναπτυχθεί μέθοδοι η εφαρμογή των οποίων δεν προϋποθέτει συγκεκριμένη κατανομή του πληθυσμού από τον οποίο προέρχεται το δείγμα.
- Οι μέθοδοι αυτές χαρακτηρίζονται ως ελεύθερες κατανομών.

Μη Παραμετρικές Μέθοδοι (*συνέχεια*)

- Με τις μεθόδους αυτές συγκρίνονται κατανομές και όχι παράμετροι (άμεσα) και για το λόγο αυτό ονομάζονται Μη Παραμετρικές Μέθοδοι.
- Είναι γενικότερες μέθοδοι και απαιτούν λιγότερες τεχνικές-μεθοδολογικές και πιθανοθεωρητικές προϋποθέσεις.
- Απαιτούν λιγότερο υπολογιστικό έργο αλλά συχνά δεν έχουν την ίδια αποτελεσματικότητα με αυτές των Παραμετρικών Μεθόδων (π.χ. έχουν μικρότερη ισχύ-*power*)

Μη Παραμετρικές Μέθοδοι (*συνέχεια*)

- Δεν χρησιμοποιούν τις πρωτογενείς τιμές αλλά τις τάξεις-βαθμίδες (*ranks*) των τιμών.
- Είναι κατάλληλες και για δεδομένα όπου οι αντίστοιχες μεταβλητές είναι μετρημένες σε κλίμακες διάταξης ή διαβάθμισης (*ordinal scales*).

Τυπολογία Κλιμάκων Μέτρησης (Ι)

- Ονομαστικές
- Διάταξης
- Διαστήματος
- Αναλογικές
- Δίτιμες-Δυαδικές
- Μονοδιάστατες
- Πολυδιάστατες

Ποιοτικές-Κατηγορικές

Ποσοτικές

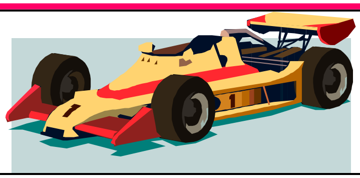
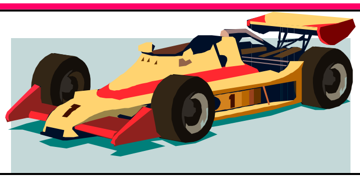
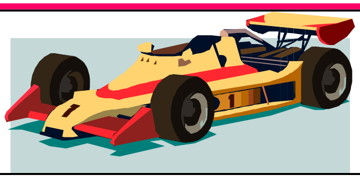
???

Στάσεις, Απόψεις, Διαθέσεις, Συμπεριφορές, Γνώσεις

Τυπολογία Κλιμάκων Μέτρησης (ΙΙ)

- Ονομαστικές (*nominal*): Το σύνολο τιμών τους δηλώνει μόνο διαφοροποίηση (π.χ. χρώμα ματιών, τόπος γέννησης, φύλο).
- Διάταξης (*ordinal*): Στο σύνολο τιμών τους μπορούμε να ορίσουμε σχέση διάταξης (π.χ. σειρά κατάταξης σε ένα αγώνισμα, μορφωτικό επίπεδο, κλάσεις ηλικιών, κλάσεις εισοδήματος, βαθμίδες ιεραρχίας).
- Διαστήματος (*interval*): Ίσες διαφορές μεταξύ των τιμών τους συνεπάγονται και ίσες διαφορές για το αντίστοιχο χαρακτηριστικό που μετρά η κλίμακα (π.χ. θερμοκρασία, ηλικία). Το μηδέν δεν είναι καλά ορισμένο.
- Αναλογίας (*ratio*): Οι τιμές τους αντιστοιχούν αναλογικά στην ποσότητα του χαρακτηριστικού που μετρούν (π.χ. ταχύτητα, χρόνος, μήκος, βάρος, εισόδημα). Το μηδέν είναι καλά ορισμένο.

Τυπολογία Κλιμάκων Μέτρησης (III) Παράδειγμα

Ονομαστική	Χώρα προέλευσης οδηγού			
		Αγγλία	Γαλλία	Βραζιλία
Διάταξης	Κατάταξη στο αγώνισμα	3ος	2ος	1ος
Διαστήματος	Βαθμολογία απόδοσης στο παγκόσμιο πρωτάθλημα (κλίμακα 0-100)	80	90	98
Αναλογίας	Επίδοση-Χρόνος Τερματισμού σε hours	1,4	1,3	1,2

Παραδείγματα με το SPSS

Εφαρμογή 6

Ranks

	Περίοδος	N	Mean Rank	Sum of Ranks
Συγκέν ρωση Ca	Θεριν ή	9	5.00	45.00
	Χειμεριν ή	10	14.50	145.00
	Total	19		

Mann-Whitney test

Test Statistics[c]

			Συγκέν ρωση Ca
Mann-Whitney U			.000
Wilcoxon W			45.000
Z			-3.674
Asy mp. Sig. (2-tailed)			.000
Exact Sig. [2*(1-tailed Sig.)]			.000[a]
Monte Carlo Sig. (2-tailed)	Sig.		.000[b]
	99% Confidence Interv al	Lower Bound	.000
		Upper Bound	.000
Monte Carlo Sig. (1-tailed)	Sig.		.000[b]
	99% Confidence Interv al	Lower Bound	.000
		Upper Bound	.000

a. Not corrected f or ties.

b. Based on 10000 sampled tables with starting seed 2000000.

c. Grouping Variable: Περίοδος

Η Μηδενική Υπόθεση Απορρίπτεται

Εφαρμογή 6 (*συνέχεια*)

- **Kolmogorov-Smirnov Test**

Frequencies

	Περίοδος	N
Συγκέν ρωση Ca	Θεριν ή	9
	Χειμεριν ή	10
	Total	19

Test Statistics[b]

			Συγκέν ρωση Ca
Most Extreme Dif ferences	Absolute		1.000
	Positiv e		.000
	Negativ e		-1.000
Kolmogorov-Smirnov Z			2.176
Asy mp. Sig. (2-tailed)			.000
Monte Carlo Sig. (2-tailed)	Sig.		.000[a]
	99% Confidence Interv al	Lower Bound	.000
		Upper Bound	.000

a. Based on 10000 sampled tables with starting seed 2000000.

b. Grouping Variable: Περίοδος

Η Μηδενική Υπόθεση Απορρίπτεται

Εφαρμογή 7

Ranks

		N	Mean Rank	Sum of Ranks
Από αέρα - Από το έδαφος	Negative Ranks	10[a]	5.50	55.00
	Positive Ranks	0[b]	.00	.00
	Ties	0[c]		
	Total	10		

a. Από αέρα < Από το έδαφος

b. Από αέρα > Από το έδαφος

c. Από αέρα = Από το έδαφος

Test Statistics[b,c]

			Από αέρα - Από το έδαφος
Z			-2.810[a]
Asymp. Sig. (2-tailed)			.005
Monte Carlo Sig. (2-tailed)	Sig.		.002
	99% Confidence Interval	Lower Bound	.001
		Upper Bound	.003
Monte Carlo Sig. (1-tailed)	Sig.		.001
	99% Confidence Interval	Lower Bound	.000
		Upper Bound	.002

a. Based on positive ranks.

b. Wilcoxon Signed Ranks Test

c. Based on 10000 sampled tables with starting seed 926214481.

- **Wilcoxon Signed Ranks Test**

Η Μηδενική Υπόθεση Απορρίπτεται

Εφαρμογή 8

Ranks

	Ομάδα	N	Mean Rank	Sum of Ranks
Απόδοση	Μάρτυρας	10	13.15	131.50
	Ακτινοβολημένο	11	9.05	99.50
	Total	21		

Test Statistics[c]

			Απόδοση
Mann-Whitney U			33.500
Wilcoxon W			99.500
Z			-1.514
Asymp. Sig. (2-tailed)			.130
Exact Sig. [2*(1-tailed Sig.)]			.132[a]
Monte Carlo Sig. (2-tailed)	Sig.		.136[b]
	99% Confidence Interval	Lower Bound	.127
		Upper Bound	.145
Monte Carlo Sig. (1-tailed)	Sig.		.068[b]
	99% Confidence Interval	Lower Bound	.062
		Upper Bound	.075

a. Not corrected for ties.

b. Based on 10000 sampled tables with starting seed 1314643744.

c. Grouping Variable: Ομάδα

- **Mann-Whitney Test**

Η Μηδενική Υπόθεση Απορρίπτεται

Παράδειγμα Προσαρμογής στην Κανονική Κατανομή

- Από την Εφαρμογή 8: Έλεγχος αν η Απόδοση Ακολουθεί Κανονική Κατανομή

One-Sample Kolmogorov-Smirnov Test

			Απόδοση
N			21
Normal Parameters[a,b]	Mean		1078.5714
	Std. Deviation		246.80530
Most Extreme Differences	Absolute		.154
	Positive		.116
	Negative		-.154
Kolmogorov-Smirnov Z			.707
Asymp. Sig. (2-tailed)			.699
Monte Carlo Sig. (2-tailed)	Sig.		.650[c]
	99% Confidence Interval	Lower Bound	.637
		Upper Bound	.662

a. Test distribution is Normal.

b. Calculated from data.

c. Based on 10000 sampled tables with starting seed 624387341.

Η Μηδενική Υπόθεση Δεν Απορρίπτεται

Ερώτηση

- Ποια είναι η Μηδενική Υπόθεση στον προηγούμενο έλεγχο;

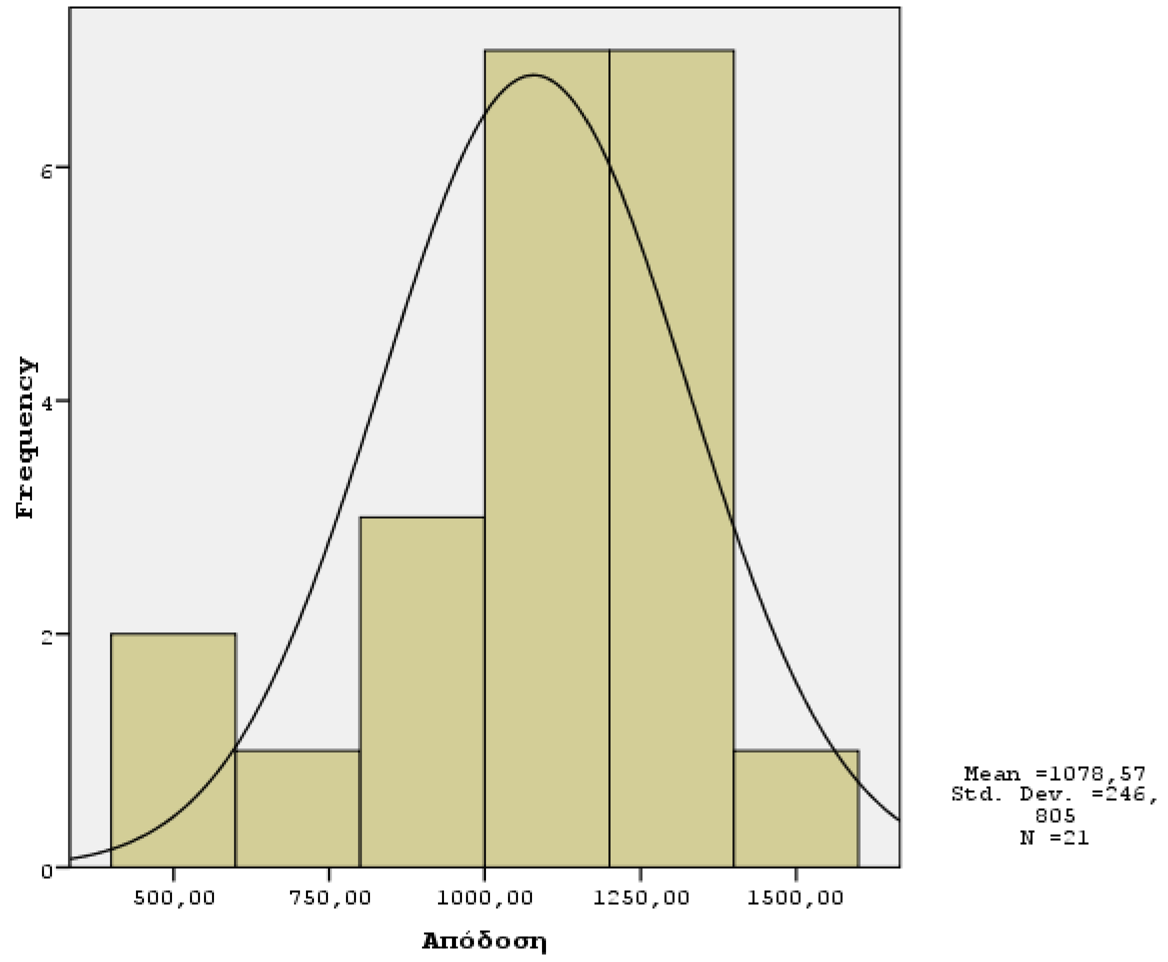
H_0: Η απόδοση ακολουθεί Κανονική Κατανομή

H_0: Η κατανομή της απόδοσης δεν διαφέρει από την Κανονική

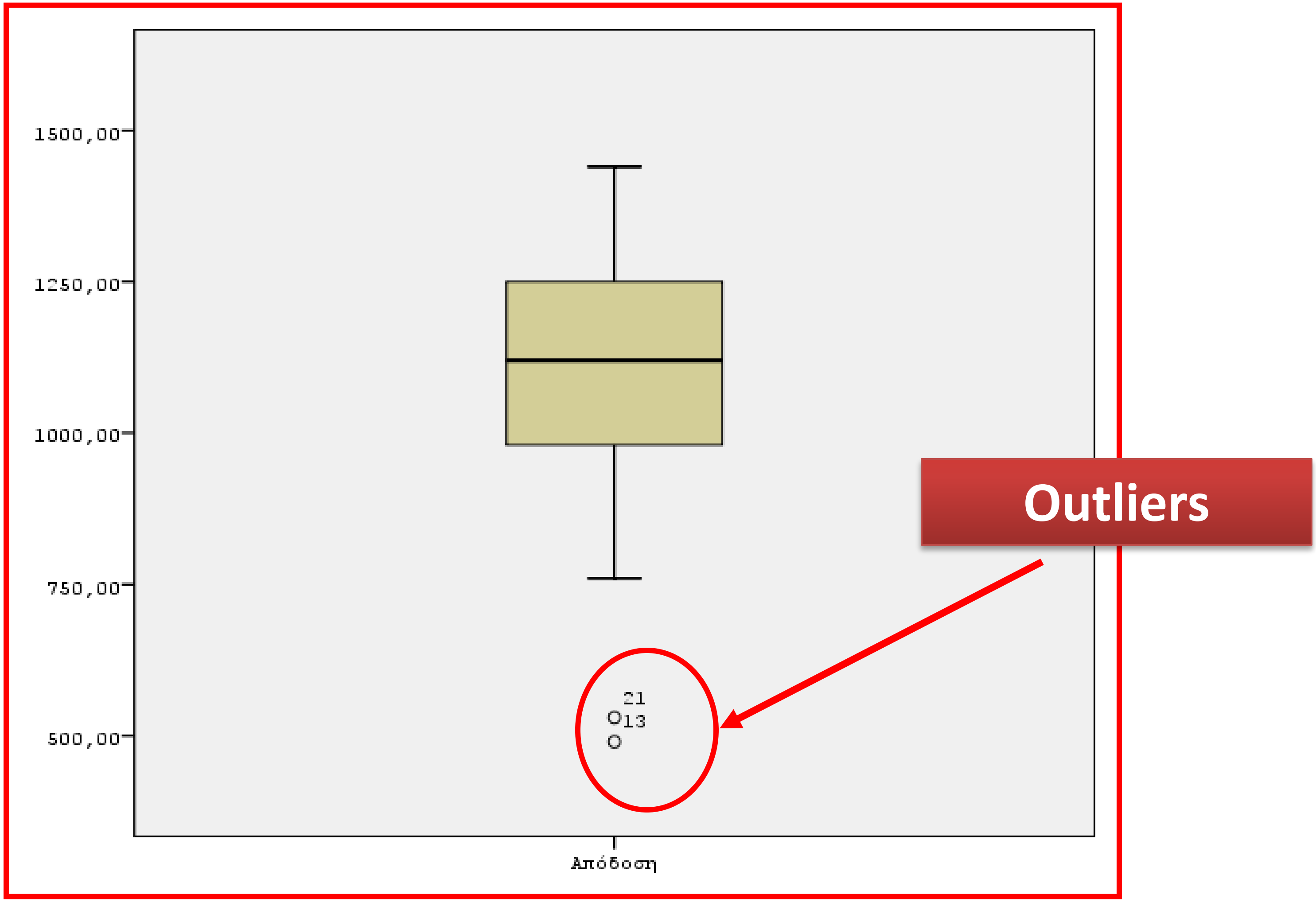
H_1: Η απόδοση δεν ακολουθεί Κανονική Κατανομή

H_1: Η κατανομή της απόδοσης διαφέρει από την Κανονική.

Ιστόγραμμα της Απόδοσης



Box plot της Απόδοσης



Περιγραφικά Μέτρα της Απόδοσης

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std.	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
Απόδοση	21	490.00	1440.00	1078.5714	246.80530	-1.118	.501	1.055	.972
Valid N (listwise)	21								

Αντιστοιχίες

Παραμετρικές Μέθοδοι	Μη Παραμετρικές
T-test για ανεξάρτητα δείγματα	Mann-Whitney test
T-test για ζευγαρωτές παρατηρήσεις	Wilcoxon test
One-way ANOVA	Kruskal-Wallis test
Pearson’s Correlation	Spearman’s Correlation

Βιβλιογραφία

- **Φωτιάδης, Ν. (1995).** *Εισαγωγή στη Στατιστική για Βιολογικές Επιστήμες*. Θεσσαλονίκη: University Studio Press.
- **Κολυβά, Φ. και Μπόρα-Σέντα, Ε. (1995).** *Στατιστική: Θεωρία-Εφαρμογές*. Θεσσαλονίκη: Εκδόσεις ΖΗΤΗ.
- **Φασούλας, Α. Κ. (ανατ. 2008).** *Στοιχεία Πειραματικής Στατιστικής*. Θεσσαλονίκη: Άγις-Σάββας Δ. Γαρταγάνης.

Σημείωμα Αναφοράς

Σημείωμα Αδειοδότησης

Τέλος Ενότητας

Επεξεργασία: Μαρία Αλεμπάκη

Θεσσαλονίκη, Εαρινό Εξάμηνο 2014-2015

Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο
ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ
ΕΣΠΑ 2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ
Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης