Στατιστική ΙΙ

Γιώργος Τσιρογιάννης

Τμήμα Διοίκησης Επιχειρήσεων Αγροτικών Προϊόντων και Τροφίμων,

Πανεπιστήμιο Πατρών

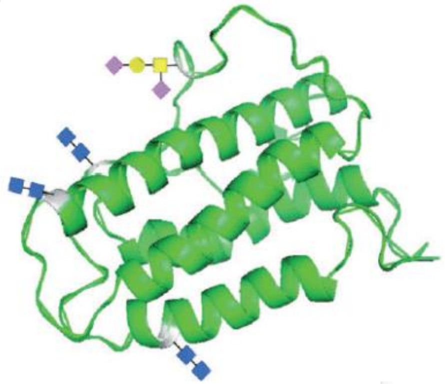
Ακαδημαϊκό έτος: 2021-2022

Διάλεξη 9η

- Έλεγχος t για δύο ανεξάρτητα δείγματα
  - Κατανομή δειγματοληψίας του $\bar{X}_1 - \bar{X}_2$
  - Έλεγχος t
  - p-τιμές
  - Στατιστικά σημαντικά αποτελέσματα
  - Διαστήματα εμπιστοσύνης

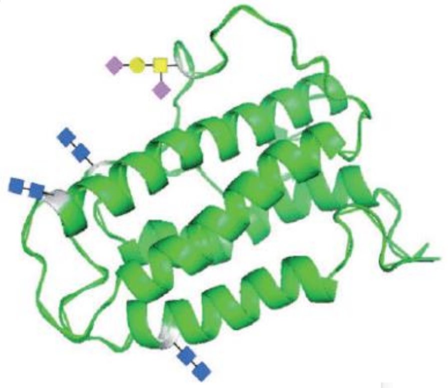
14ο κεφάλαιο

Παράδειγμα

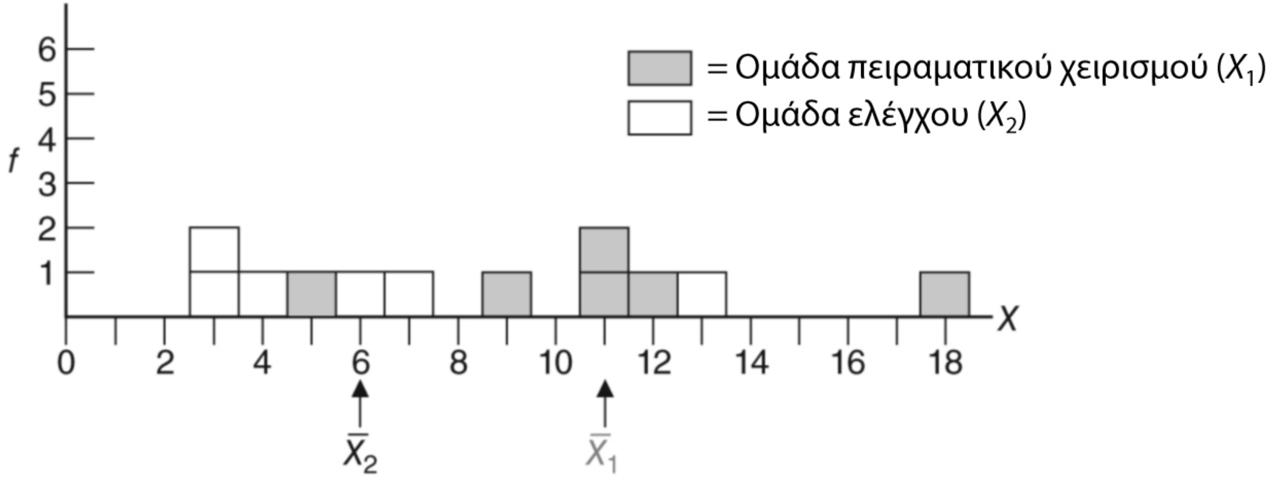


- Η Ερυθροποιητίνη (EPO) διεγείρει την παραγωγή ερυθρών αιμοσφαιρίων, που με την σειρά τους μεταφέρουν περισσότερο οξυγόνο
- Σε μια κλινική ένας ερευνητής θέλει να διαπιστώσει αν μπορεί να βοηθήσει ασθενείς στην σωματική τους αντοχή
- Σχηματίζει δύο ομάδες:
  - Λαμβάνουν EPO
  - Λαμβάνουν ένα άκακο σκεύασμα

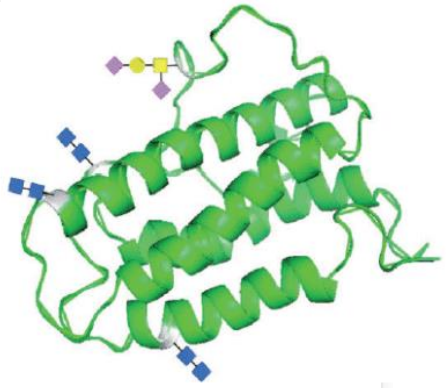
Παράδειγμα



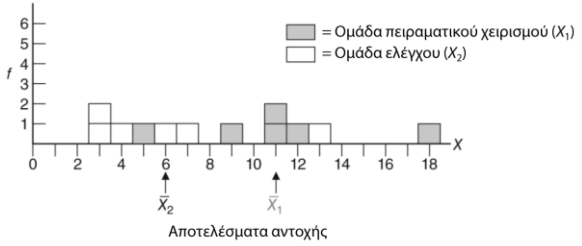
- Η αποτελεσματικότητα μετράται μέσω του χρόνου που ο ασθενής παραμένει σε ταχέως κινούμενο διάδρομο
- Στην πράξη θα κάναμε χρήση μεγαλύτερων δειγμάτων (μόνο 6 σε κάθε ομάδα για ευκολία στην παρουσίαση)



Παράδειγμα



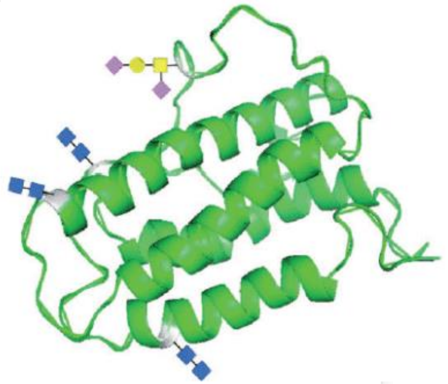
- Παρατηρούμε μια διαφορά περίπου 5 λεπτών μεταξύ των ομάδων
- Είναι πραγματική και αναμένεται να εμφανισθεί και σε άλλα παρόμοια πειράματα;
- Είναι αποτέλεσμα τυχαιότητας και ενδεχομένως σε άλλο παρόμοιο πείραμα αντιστραφεί η διαφορά (η επικάλυψη το κάνει πιθανό);



Ανεξάρτητα δείγματα

- Όταν το κάθε δείγμα είναι προϊόν τυχαίας δειγματοληψίας από δύο πληθυσμούς
- Επειδή προκύπτουν από διαφορετικούς πληθυσμούς δεν «ζευγαρόνονται»: δηλαδή τα σημεία δεν αντιστοιχίζονται 1-προς-1 μεταξύ των πληθυσμών

Παράδειγμα

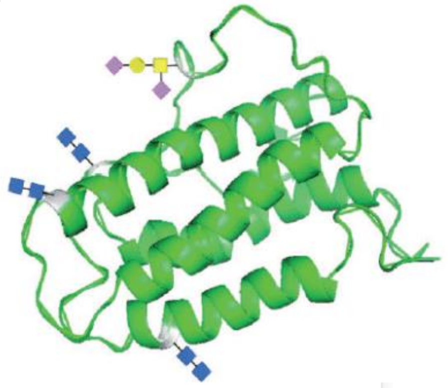


- Οι δύο πληθυσμοί:
  - Οι ασθενείς της συγκεκριμένης κλινικής που λαμβάνουν EPO
  - Οι ασθενείς της συγκεκριμένης κλινικής που λαμβάνουν άκακη ουσία
- Συνήθης υπόθεση:
  - Θεωρείται τυχαίο δείγμα από τους ασθενείς που πάσχουν από την ίδια ασθένεια και λαμβάνουν EPO
  - Θεωρείται τυχαίο δείγμα από τους ασθενείς που πάσχουν από την ίδια ασθένεια και λαμβάνουν άκακη ουσία
- Οι δύο αυτοί υποθετικοί πληθυσμοί αναφέρονται στην μηδενική και εναλλακτική υπόθεση

Επίδραση

- Η διαφορά μεταξύ των δύο πληθυσμών εκφράζει την επίδραση
  - Στο παράδειγμά μάς την επίδραση της EPO στην σωματική αντοχή των πληθυσμών
- Το μέγεθος και πρόσημο της διαφοράς ποσοτικοποιούν την επίδραση:
  - Πχ διαφορά +5 λεπτών στο διάδρομο υψηλής ταχύτητας

Παράδειγμα

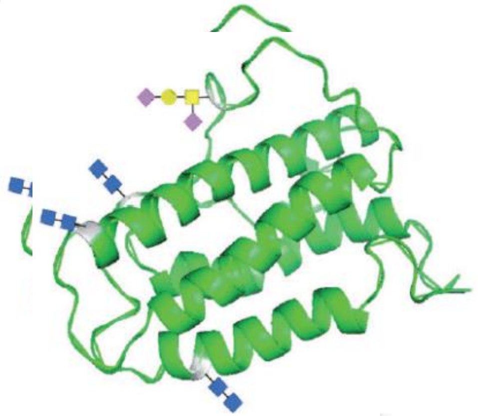


- Μηδενική υπόθεση:
  - Δεν συμβαίνει τίποτα το ιδιαίτερο με την λήψη EPO.
  - Δεν υπάρχει διαφορά μεταξύ των μέσων, είτε η EPO μειώνει την φυσική κατάσταση:

$$\mathrm{H}_0\colon \mu_1 - \mu_2 \leq 0$$

Όπου μ_1, μ_2 ο μέσος του πληθυσμού των ασθενών που λαμβάνου EPO και άκακη ουσία αντίστοιχα

Παράδειγμα



- Εναλλακτική υπόθεση:
  - Κάτι ιδιαίτερο συμβαίνει
    - Η EPO αυξάνει την αντοχή

$$H_1: \mu_1 - \mu_2 > 0$$

Μονόπλευρος έλεγχος αφού μας ενδιαφέρει η θετική επίδραση από την χορήγηση της EPO

Εναλλακτικές μορφές διατύπωσης

- $\mathrm{H}_0\colon \mu_1 - \mu_2 \geq 0, \mathrm{H}_1\colon \mu_1 - \mu_2 < 0$

- $\mathrm{H}_0\colon \mu_1 - \mu_2 = 0, \mathrm{H}_1\colon \mu_1 - \mu_2 \neq 0$

Κατανομή δειγματοληψίας $\bar{X}_1 - \bar{X}_2$

- Η κατανομή που ακολουθούν οι διαφορές μεταξύ των δειγματικών μέσων με βάση όλα τα πιθανά ζεύγη δύο δειγμάτων δύο πληθυσμών
- Αναπαριστά όλο το φάσμα των συνδυασμών δειγμάτων
- Μας βοηθάει να ποσοτικοποιήσουμε πόσο σπάνιο είναι το παρατηρούμενο αποτέλεσμα των δειγμάτων που έχουμε τυχαία συλλέξει
- Ένα κοινό αποτέλεσμα μπορεί να ερμηνευθεί μέσω της τυχαιότητας της δειγματοληψίας
- Ένα σπάνιο αποτέλεσμα μπορεί να ερμηνευθεί ως μια υπαρκτή διαφορά των μέσων πέραν της τυχαίας δειγματοληψίας

Ισχύουν

- Ο μέσος της κατανομής δειγματοληψίας $\bar{\mathrm{X}}_1 - \bar{\mathrm{X}}_2$ ισούται με την διαφορά των μέσων του πληθυσμού:

$$\mu_{\bar{\mathrm{X}}_1 - \bar{\mathrm{X}}_2} = \mu_1 - \mu_2$$

όπου μ_1, μ_2 ο μέσος του πρώτου και δεύτερου πληθυσμού αντίστοιχα

- Το τυπικό σφάλμα προκύπτει από τον συνδυασμό των τυπικών αποκλίσεων και των δύο πληθυσμών

$$\sigma_{\bar{\mathrm{X}}_1 - \bar{\mathrm{X}}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

όπου σ_1, n_1 και σ_2, n_2 η τυπική απόκλιση και το μέγεθος δείγματος του πρώτου και δεύτερου πληθυσμού αντίστοιχα

Έλεγχος t

- Βασίζεται στην τυποποιημένη κατανομή δειγματοληψίας t
- Η τυπική απόκλιση του πρώτου και δεύτερου πληθυσμού δεν θεωρείται γνωστή, αλλά εκτιμάται
- Βασίζεται στον λόγο

$$t = \frac{(\delta\iota\alpha\varphi o\rho\alpha\ \mu\varepsilon\tau\alpha\xi\acute{\upsilon}\ \tau\omega\nu\ \delta\varepsilon\iota\gamma\mu\alpha\tau\iota\kappa\acute{\omega}\nu\ \mu\varepsilon\sigma\omega\nu) - (\upsilon\pi o\theta\varepsilon\tau\iota\kappa\eta\ \delta\iota\alpha\varphi o\rho\alpha\ \tau\omega\nu\ \mu\varepsilon\sigma\omega\nu\ \pi\lambda\eta\theta\upsilon\sigma\mu o\upsilon)}{\varepsilon\kappa\tau\iota\mu\omega\mu\varepsilon\nu o\ \tau\upsilon\pi\iota\kappa\acute{o}\ \sigma\varphi\alpha\lambda\mu\alpha}$$

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{s_{\bar{X}_1 - \bar{X}_2}}$$

Εκτιμώμενο τυπικό σφάλμα του μέσου

$$SS_1 = \sum X_1^2 - \frac{(\sum X_1)^2}{n_1}$$

$$SS_2 = \sum X_2^2 - \frac{(\sum X_2)^2}{n_2}$$

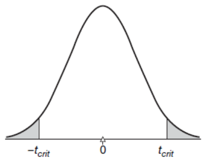
$$s_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{s_P^2}{n_1} + \frac{s_P^2}{n_2}}$$

$$s_p^2 = \frac{SS_1 + SS_2}{df} = \frac{SS_1 + SS_2}{n_1 + n_2 - 2}$$

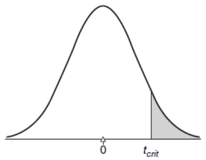
Ομαδοποιημένη εκτίμηση της διασποράς

Εύρεση των κρίσιμων τιμών

- Κάνουμε χρήση των πινάκων της κατανομής t
- Υπολογίζουμε τους βαθμούς ελευθερίας, από τον αριθμό που κάναμε χρήση στην ομαδοποιημένη εκτίμηση της διασποράς
- Αναζητάμε την κρίσιμη τιμή βάσει του επιπέδου σημαντικότητας του προβλήματος και τον τύπο υπόθεσης (μονόπλευρος ή αμφίπλευρος)



	$p > .05$	$p < .05$	$p < .01$	$p < .001$
df		**.05***	**.01****	**.001**
1	12.706	63.657	636.62	
2	4.303	9.925	31.598	
3	3.182	5.841	12.924	
4	2.776	4.604	8.610	
5	2.571	4.032	6.869	



	$p > .05$	$p < .05$	$p < .01$	$p < .001$
df	**.05**	**.01**	**.001**	
1	6.314	31.821	318.31	
2	2.920	6.965	22.326	
3	2.353	4.541	10.213	
4	2.132	3.747	7.173	
5	2.015	3.365	5.893	

Παράδειγμα

(α) αμφίπλευρος, α=0.05, n1=12, n2=11

±2.080

(β) μονόπλευρος δεξιάς πλευράς, α=0.05, n1=15, n2=13

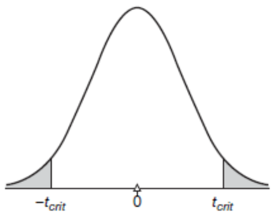
1.706

(γ) μονόπλευρος αριστερής πλευράς, α=0.01, n1=n2=25

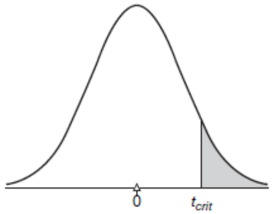
–2.423

(δ) αμφίπλευρος, α=0.01, n1=8, n2=10

±2.921



	p > .05	p < .05	p < .01	p < .001
df	.05*	.01**	.001	
1	12.706	63.657	636.62	
2	4.303	9.925	31.598	
3	3.182	5.841	12.924	
4	2.776	4.604	8.610	
5	2.571	4.032	6.869	
6	2.447	3.707	5.959	
7	2.365	3.499	5.408	
8	2.306	3.355	5.041	
9	2.262	3.250	4.781	
10	2.228	3.169	4.587	
11	2.201	3.106	4.437	
12	2.179	3.055	4.318	
13	2.160	3.012	4.221	
14	2.145	2.977	4.140	
15	2.131	2.947	4.073	
16	2.120	2.921	4.015	
17	2.110	2.898	3.965	
18	2.101	2.878	3.922	
19	2.093	2.861	3.883	
20	2.086	2.845	3.850	
21	2.080	2.831	3.819	
22	2.074	2.819	3.792	
23	2.069	2.807	3.767	
24	2.064	2.797	3.745	
25	2.060	2.787	3.725	
26	2.056	2.779	3.707	
27	2.052	2.771	3.690	
28	2.048	2.763	3.674	
29	2.045	2.756	3.659	
30	2.042	2.750	3.646	
40	2.021	2.704	3.551	
60	2.000	2.660	3.460	
120	1.980	2.617	3.373	
∞	1.960	2.576	3.291	



	p > .05	p < .05	p < .01	p < .001
df	.05	.01	.001	
1	6.314	31.821	318.31	
2	2.920	6.965	22.326	
3	2.353	4.541	10.213	
4	2.132	3.747	7.173	
5	2.015	3.365	5.893	
6	1.943	3.143	5.208	
7	1.895	2.998	4.785	
8	1.860	2.896	4.501	
9	1.833	2.821	4.297	
10	1.812	2.764	4.144	
11	1.796	2.718	4.025	
12	1.782	2.681	3.930	
13	1.771	2.650	3.852	
14	1.761	2.624	3.787	
15	1.753	2.602	3.733	
16	1.746	2.583	3.686	
17	1.740	2.567	3.646	
18	1.734	2.552	3.610	
19	1.729	2.539	3.579	
20	1.725	2.528	3.552	
21	1.721	2.518	3.527	
22	1.717	2.508	3.505	
23	1.714	2.500	3.485	
24	1.711	2.492	3.467	
25	1.708	2.485	3.450	
26	1.706	2.479	3.435	
27	1.703	2.473	3.421	
28	1.701	2.467	3.408	
29	1.699	2.462	3.396	
30	1.697	2.457	3.385	
40	1.684	2.423	3.307	
60	1.671	2.390	3.232	
120	1.658	2.358	3.160	
∞	1.645	2.326	3.090	

Βήματα υπολογισμού του t

Βήμα 1ο:

1. Εκχώρηση τιμής του n_1 (1)
2. Πρόσθεση όλων των μετρήσεων του X_1(2)
3. Υπολογισμός του $\bar{X}_1$(3)
4. Υπολογισμός των τετραγώνων όλων των μετρήσεων του X_1(4)
5. Πρόσθεση όλων των τετραγώνων (5)
6. Υπολογισμός του SS_1 (6)

Βήματα υπολογισμού του t

Βήμα 2ο : (εφαρμογή για το X_2)

1. Εκχώρηση τιμής του n_2 (1)
2. Πρόσθεση όλων των μετρήσεων του X_2(2)
3. Υπολογισμός του $\bar{X}_2$(3)
4. Υπολογισμός των τετραγώνων όλων των μετρήσεων του X_2(4)
5. Πρόσθεση όλων των τετραγώνων (5)
6. Υπολογισμός του SS_2 (6)

Βήματα υπολογισμού του t

Βήμα 3ο :

Εύρεση της ομαδοποιημένης διασποράς, s_p^2 (7)

$$s_p^2 = \frac{SS_1 + SS_2}{df} = \frac{SS_1 + SS_2}{n_1 + n_2 - 2}$$

Βήματα υπολογισμού του t

Βήμα 4ο :

Εύρεση του τυπικού σφάλματος (8)

$$s_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{s_P^2}{n_1} + \frac{s_P^2}{n_2}}$$

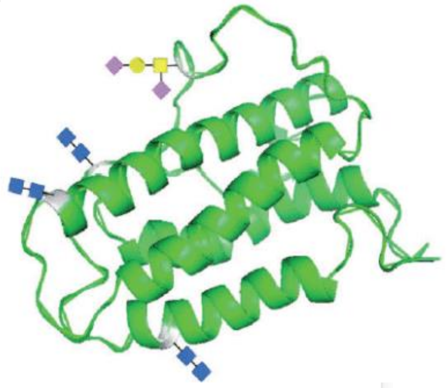
Βήματα υπολογισμού του t

Βήμα 5ο :

Υπολογισμός του λόγου t (9)

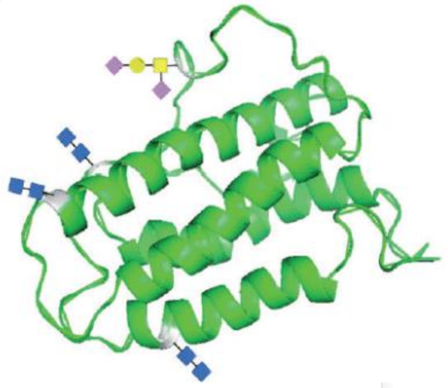
$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{s_{\bar{X}_1 - \bar{X}_2}}$$

Παράδειγμα

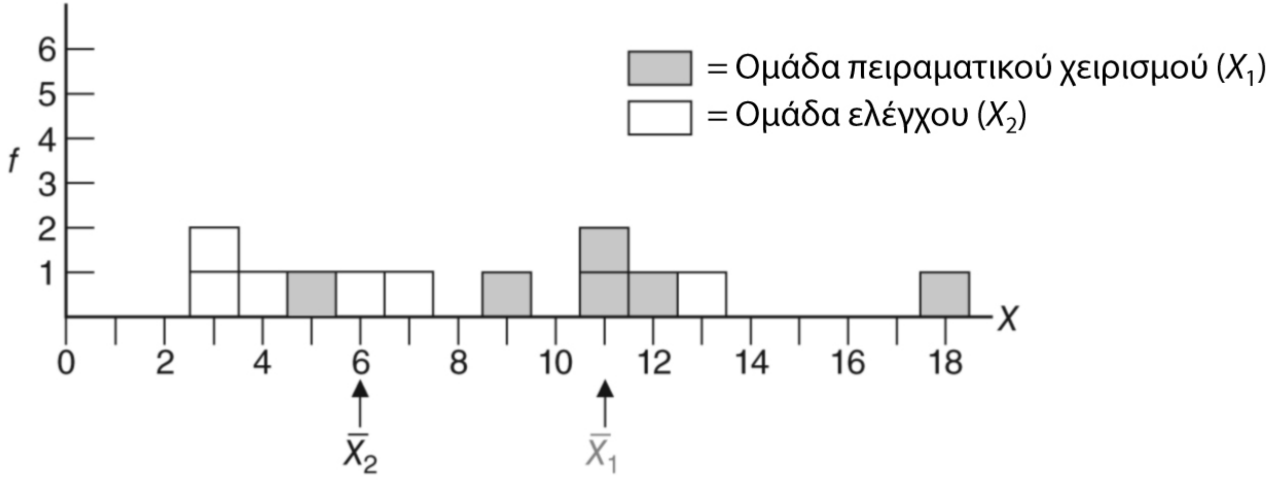


- Η Ερυθροποιητίνη (EPO) διεγείρει την παραγωγή ερυθρών αιμοσφαιρίων, που με την σειρά τους μεταφέρουν περισσότερο οξυγόνο
- Σε μια κλινική ένας ερευνητής θέλει να διαπιστώσει αν μπορεί να βοηθήσει ασθενείς στην σωματική τους αντοχή
- Σχηματίζει δύο ομάδες:
  - Λαμβάνουν EPO
  - Λαμβάνουν ένα άκακο σκεύασμα

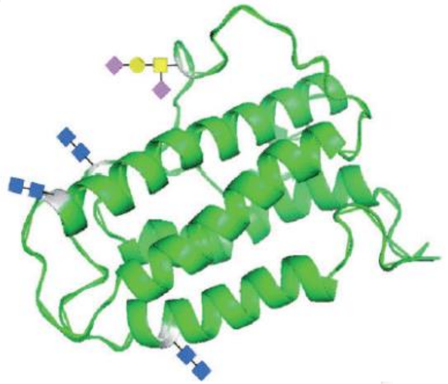
Παράδειγμα



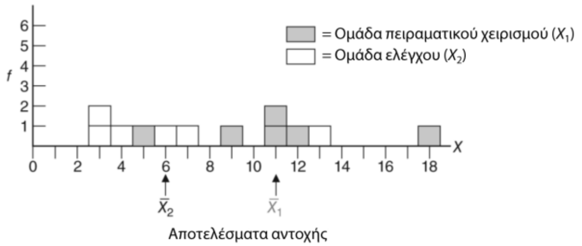
- Η αποτελεσματικότητα μετράται μέσω του χρόνου που ο ασθενής παραμένει σε ταχέως κινούμενο διάδρομο
- Στην πράξη θα κάναμε χρήση μεγαλύτερων δειγμάτων (μόνο 6 σε κάθε ομάδα για ευκολία στην παρουσίαση)



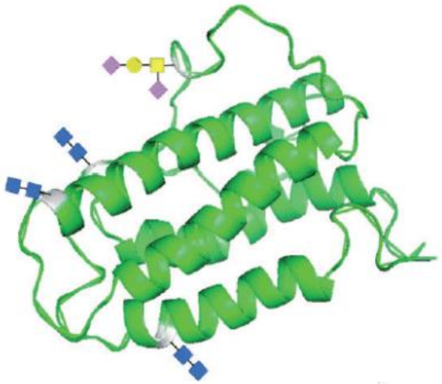
Παράδειγμα



- Παρατηρούμε μια διαφορά περίπου 5 λεπτών μεταξύ των ομάδων
- Είναι πραγματική και αναμένεται να εμφανισθεί και σε άλλα παρόμοια πειράματα;
- Είναι αποτέλεσμα τυχαιότητας και ενδεχομένως σε άλλο παρόμοιο πείραμα αντιστραφεί η διαφορά (η επικάλυψη το κάνει πιθανό);



Παράδειγμα



Βήματα υπολογισμού του t

Βήμα 1°:

1. Εκχώρηση τιμής του n_1 (1)
2. Πρόσθεση όλων των μετρήσεων του X_1(2)
3. Υπολογισμός του $\bar{X}_1$(3)
4. Υπολογισμός των τετραγώνων όλων των μετρήσεων του X_1(4)
5. Πρόσθεση όλων των τετραγώνων (5)
6. Υπολογισμός του SS_1 (6)

EPO group

X_1	X_1^2 (4)
12	144
5	25
11	121
11	121
9	81
18	324

1 $n_1 = 6$

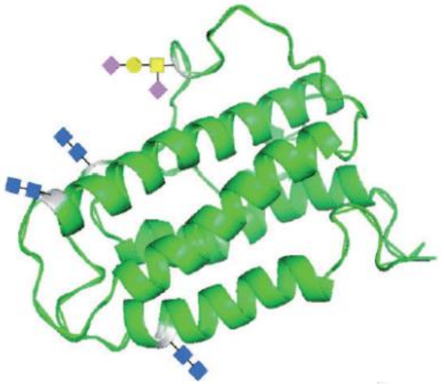
2 $\sum X_1 = 66$

3 $$\bar{X}_1 = \frac{\sum X_1}{n_1} = \frac{66}{6} = 11$$

5 $\sum X_1^2 = 816$

6 $$SS_1 = \sum X_1^2 - \frac{\left(\sum X_1\right)^2}{n_1} = 816 - \frac{(66)^2}{6} = 816 - 726 = 90$$

Παράδειγμα



Control group

X_2	X_2^2
7	49
3	9
4	16
6	36
3	9
13	169

$n_2 = 6 \qquad \sum X_2 = 36 \qquad \sum X_2^2 = 288$

$$\bar{X}_2 = \frac{\sum X_2}{n_2} = \frac{36}{6} = 6$$

$$\begin{aligned} SS_2 &= \sum X_2^2 - \frac{\left(\sum X_2\right)^2}{n_2} \\ &= 288 - \frac{(36)^2}{6} \\ &= 288 - 216 \\ &= 72 \end{aligned}$$

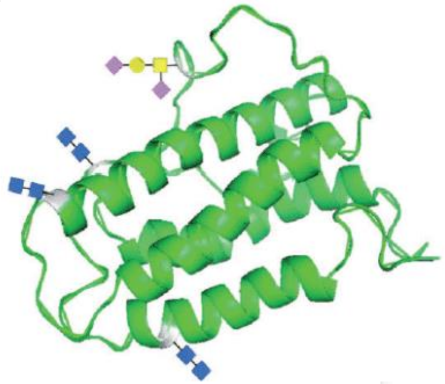
Βήματα υπολογισμού του t

Βήμα 2ο : (εφαρμογή για το X_2)

1. Εκχώρηση τιμής του n_2 (1)
2. Πρόσθεση όλων των μετρήσεων του X_2(2)
3. Υπολογισμός του $\bar{X}_2$(3)
4. Υπολογισμός των τετραγώνων όλων των μετρήσεων του X_2(4)
5. Πρόσθεση όλων των τετραγώνων (5)
6. Υπολογισμός του SS_2 (6)

19

Παράδειγμα



$$n_1 = 6 \qquad n_2 = 6$$

$$SS_1 = 90 \quad SS_2 = 72$$

Βήματα υπολογισμού του t

Βήμα 3ο :

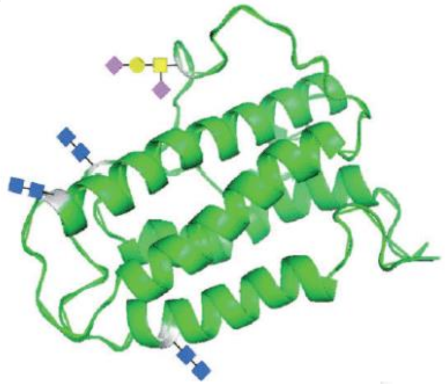
Εύρεση της ομαδοποιημένης διασποράς, s_p^2 (7)

$$s_p^2 = \frac{SS_1 + SS_2}{df} = \frac{SS_1 + SS_2}{n_1 + n_2 - 2}$$

20

$$7 \quad s_p^2 = \frac{SS_1 + SS_2}{n_1 + n_2 - 2} = \frac{90 + 72}{6 + 6 - 2} = \frac{162}{10} = 16.2$$

Παράδειγμα



Βήματα υπολογισμού του t

Βήμα 4ο :

Εύρεση του τυπικού σφάλματος (8)

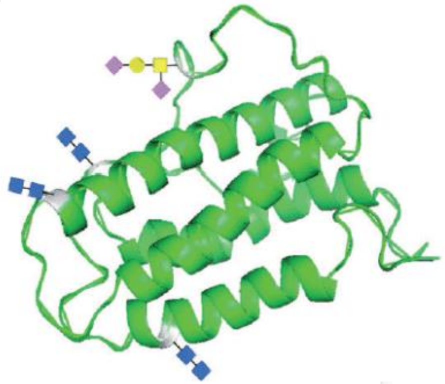
$$s_{\overline{X}_1-\overline{X}_2} = \sqrt{\frac{s_P^2}{n_1} + \frac{s_P^2}{n_2}}$$

21

$$s_P^2 = 16.2$$

$$s_{\overline{X}_1-\overline{X}_2} = \sqrt{\frac{s_P^2}{n_1} + \frac{s_P^2}{n_2}} = \sqrt{\frac{16.2}{6} + \frac{16.2}{6}} = \sqrt{\frac{32.4}{6}} = \sqrt{5.4} = 2.32$$

Παράδειγμα



Βήματα υπολογισμού του t

Βήμα 5ο :

Υπολογισμός του λόγου t (9)

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{s_{\bar{X}_1 - \bar{X}_2}}$$

22

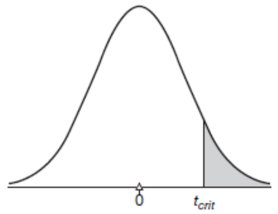
$$\bar{X}_1 = 11 \qquad \bar{X}_2 = 6$$

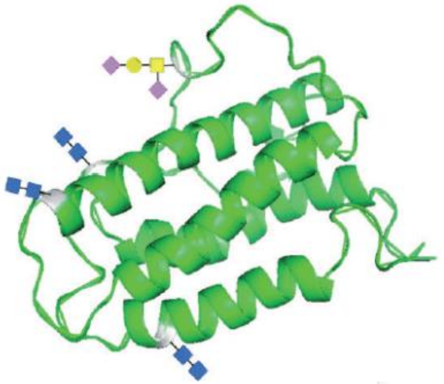
$$s_{\bar{X}_1 - \bar{X}_2} = 2.16$$

$$(\mu_1 - \mu_2)_{hyp} \; ? \longrightarrow 0$$

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)_{hyp}}{s_{\bar{X}_1 - \bar{X}_2}} = \frac{(11-6)-0}{2.32} = \frac{5}{2.32} = 2.16$$

Παράδειγμα





$\alpha=0.05$ $n_1 = 6$ $n_2 = 6$ $df=n_1+n_2-2=12-2=10$

$$t=\frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)_{hyp}}{s_{\bar{X}_1-\bar{X}_2}}=\frac{(11-6)-0}{2.32}=\frac{5}{2.32}=2.16 > $$

$$\mathrm{H_0}: \mu_1-\mu_2 \le 0$$

Απορρίπτεται

$p > .05$ $p < .05$ $p < .01$ $p < .001$

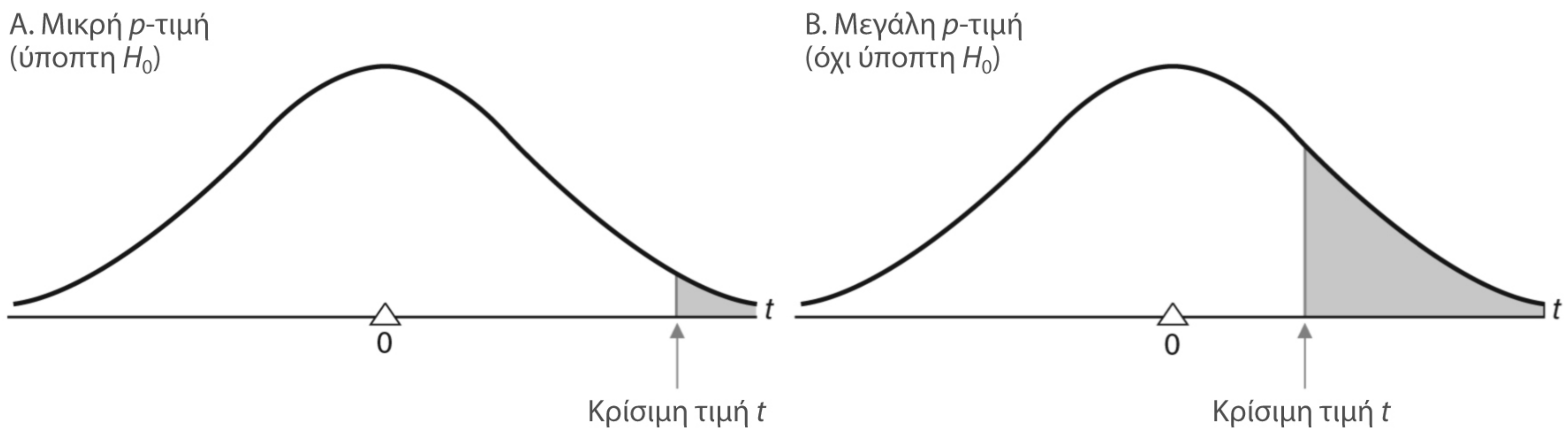
df	.05	.01	.001
1	6.314	31.821	318.31
2	2.920	6.965	22.326
3	2.353	4.541	10.213
4	2.132	3.747	7.173
5	2.015	3.365	5.893
6	1.943	3.143	5.208
7	1.895	2.998	4.785
8	1.860	2.896	4.501
9	1.833	2.821	4.297
10	1.812	2.764	4.144
11	1.796	2.718	4.025
12	1.782	2.681	3.930
13	1.771	2.650	3.852
14	1.761	2.624	3.787
15	1.753	2.602	3.733
16	1.746	2.583	3.686
17	1.740	2.567	3.646
18	1.734	2.552	3.610
19	1.729	2.539	3.579
20	1.725	2.528	3.552
21	1.721	2.518	3.527
22	1.717	2.508	3.505
23	1.714	2.500	3.485
24	1.711	2.492	3.467
25	1.708	2.485	3.450
26	1.706	2.479	3.435
27	1.703	2.473	3.421
28	1.701	2.467	3.408
29	1.699	2.462	3.396
30	1.697	2.457	3.385
40	1.684	2.423	3.307
60	1.671	2.390	3.232
120	1.658	2.358	3.160
∞	1.645	2.326	3.090

p-τιμές (p-values)

- Ποσοτικοποιεί την σπανιότητα εμφάνισης του δείγματος, υπό την συνθήκη ότι η μηδενική-υπόθεση είναι αληθής
- Συνήθως η H0 ούτε απορρίπτεται ούτε διατηρείται, αλλά εξετάζεται με τον βαθμό υποψίας
- Μικρές τιμές χρησιμοποιούνται στην πράξη για την αναιρέσουν την H0 και να υποστηρίξουν την υπόθεση έρευνας

p-τιμές (p-values)

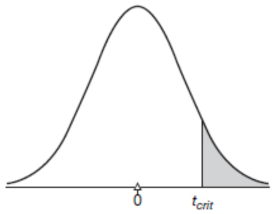
- Αναπαριστά την αναλογία εμβαδού, πέρα από το παρατηρούμενο αποτέλεσμα στην ουρά κατανομής της δειγματοληψίας

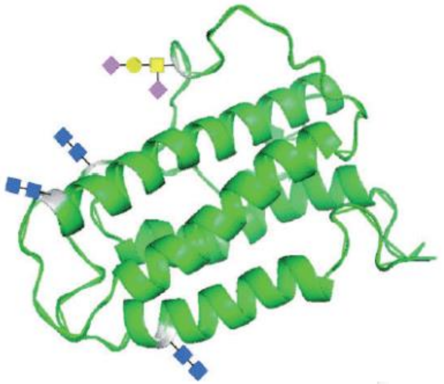


p-τιμές (p-values) - προσεγγίσεις

- Δουλεύουμε με τους πίνακες (πχ κατανομής t)
- Επιλέγουμε αμφίπλευρο ή μονόπλευρο (ανάλογα με το πρόβλημα)
- Εντοπίζουμε την γραμμή με τον σωστό βαθμό ελευθερίας
- Κινούμαστε κάθετα ώστε να εντοπίσουμε το κάτω οριό του

Παράδειγμα





$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)_{hyp}}{s_{\bar{X}_1 - \bar{X}_2}} = \frac{(11-6)-0}{2.32} = \frac{5}{2.32} = 2.16$$

p-τιμή < 0.05

	$p > .05$	$p < .05$	$p < .01$	$p < .001$
df	.05	.01	.001	
1	6.314	31.821	318.31	
2	2.920	6.965	22.326	
3	2.353	4.541	10.213	
4	2.132	3.747	7.173	
5	2.015	3.365	5.893	
6	1.943	3.143	5.208	
7	1.895	2.998	4.785	
8	1.860	2.896	4.501	
9	1.833	2.821	4.297	
10	1.812	2.764	4.144	
11	1.796	2.718	4.025	
12	1.782	2.681	3.930	
13	1.771	2.650	3.852	
14	1.761	2.624	3.787	
15	1.753	2.602	3.733	
16	1.746	2.583	3.686	
17	1.740	2.567	3.646	
18	1.734	2.552	3.610	
19	1.729	2.539	3.579	
20	1.725	2.528	3.552	
21	1.721	2.518	3.527	
22	1.717	2.508	3.505	
23	1.714	2.500	3.485	
24	1.711	2.492	3.467	
25	1.708	2.485	3.450	
26	1.706	2.479	3.435	
27	1.703	2.473	3.421	
28	1.701	2.467	3.408	
29	1.699	2.462	3.396	
30	1.697	2.457	3.385	
40	1.684	2.423	3.307	
60	1.671	2.390	3.232	
120	1.658	2.358	3.160	
∞	1.645	2.326	3.090	

Επίπεδο σημαντικότητας ή p-τιμή;

- Το επίπεδο σημαντικότητας προαποφασίζεται και συνεισφέρει σε μια πιο δομημένη έρευνα
- Η p-τιμή αποτυπώνει την σπανιότητα σύμφωνα με την κατανομή δειγματοληψίας
- Η p-τιμή όταν χρησιμοποιείται ορθά, μπορεί να παίξει τον ρόλο απόφασης απόρριψης ή όχι της H0

Παράδειγμα

- Επιλέξτε το σπανιότερο αποτέλεσμα μεταξύ των ζευγών από τις παρακάτω p-values

(a_1) $p > .05$	**(a_2)** $p < .05$
(b_1) $p < .001$	**(b_2)** $p < .01$
(c_1) $p < .05$	**(c_2)** $p < .01$
(d_1) $p < .10$	**(d_2)** $p < .20$
(e_1) $p = .04$	**(e_2)** $p = .02$

$a_2, b_1, b_2, c_1, c_2, e_1, e_2$

Παράδειγμα

- Ποια από τις παρακάτω p-values μας κάνει να απορρίψουμε την H0 σε επίπεδο σημαντικότητας 0.05

(a₁) $p > .05$	**(a₂)** $p < .05$
(b₁) $p < .001$	**(b₂)** $p < .01$
(c₁) $p < .05$	**(c₂)** $p < .01$
(d₁) $p < .10$	**(d₂)** $p < .20$
(e₁) $p = .04$	**(e₂)** $p = .02$

Στατιστικά σημαντικά αποτελέσματα

- Οι έλεγχοι υποθέσεων (ΕΥ) αναφέρονται ως έλεγχοι σημαντικότητας
- Τα αποτελέσματα των ΕΥ περιγράφονται ως στατιστικά σημαντικά αν η Η0 έχει απορριφθεί
- Τα αποτελέσματα των ΕΥ περιγράφονται ως μη-στατιστικά σημαντικά αν η Η0 διατηρείται
- Δηλαδή ένα στατιστικά σημαντικό αποτέλεσμα δηλώνει ότι είναι απίθανο να οφείλεται στην τύχη
- Αναφέρεται στο δείγμα
- Δεν δείχνει ότι η υποκείμενη επίδραση (effect) είναι σημαντική
- Μεταξύ δειγμάτων, δείχνει μόνο ότι η Η0 είναι ψευδής και δεν καθορίζει το μέγεθος της διαφοράς των μέσων

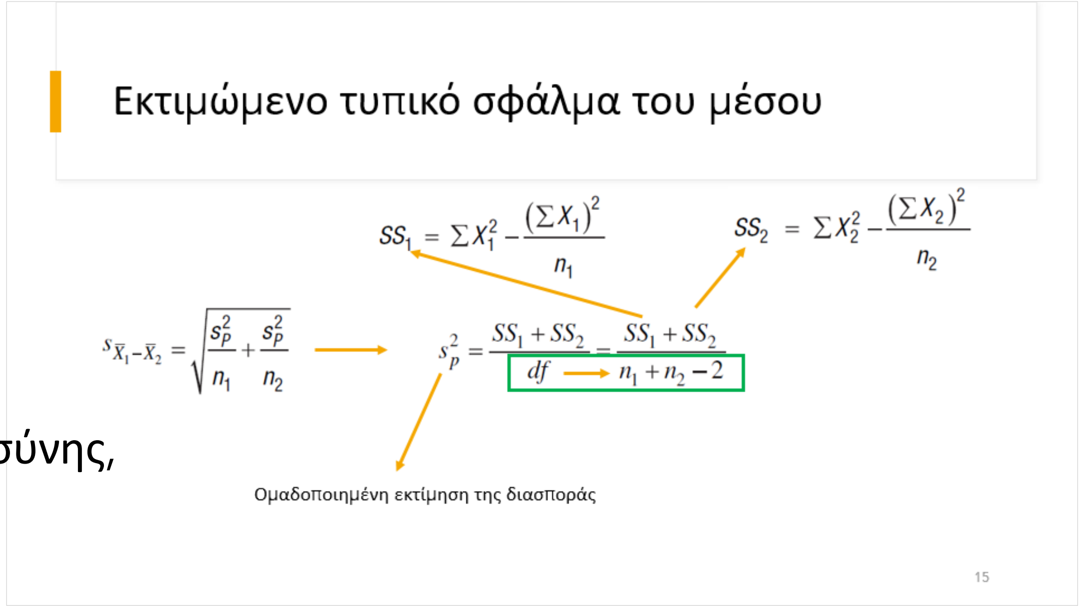
Εκτιμήσεις διαστημάτων εμπιστοσύνης

Καθορίζει τα εύρη τιμών της διαφοράς $\mu_1 - \mu_2$ που μακροπρόθεσμα θα περιλαμβάνουν την άγνωστη επίδραση.

$$\bar{X}_1 - \bar{X}_2 \pm \left(t_{conf}\right)\left(s_{\bar{X}_1-\bar{X}_2}\right)$$

$$\bar{X}_1 = \frac{\sum X_1}{n_1} \qquad \bar{X}_2 = \frac{\sum X_2}{n_2}$$

Τιμή που εξαρτάται από το επίπεδο εμπιστοσύνης,
και τα μεγέθη των δειγμάτων



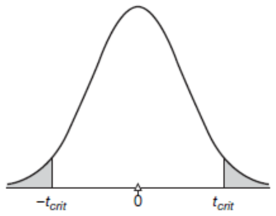
Παράδειγμα

X_1	X_2
12	7
5	3
11	4
11	6
9	3
18	13

n1=n2=6

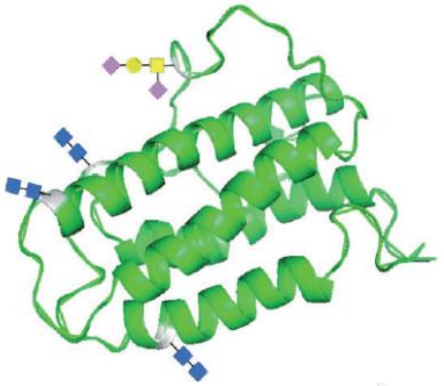
$$s_{\bar{X}_1-\bar{X}_2} = \sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}$$

$$s_p^2 = \frac{SS_1 + SS_2}{df} = \frac{SS_1 + SS_2}{n_1 + n_2 - 2}$$



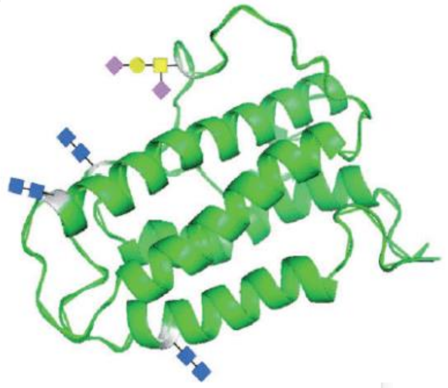
$p > .05$ $p < .05$ $p < .01$ $p < .001$

df	.05*	.01**	.001
1	12.706	63.657	636.62
2	4.303	9.925	31.598
3	3.182	5.841	12.924
4	2.776	4.604	8.610
5	2.571	4.032	6.869
6	2.447	3.707	5.959
7	2.365	3.499	5.408
8	2.306	3.355	5.041
9	2.262	3.250	4.781
10	2.228	3.169	4.587
11	2.201	3.106	4.437
12	2.179	3.055	4.318
13	2.160	3.012	4.221
14	2.145	2.977	4.140
15	2.131	2.947	4.073
16	2.120	2.921	4.015
17	2.110	2.898	3.965
18	2.101	2.878	3.922
19	2.093	2.861	3.883
20	2.086	2.845	3.850
21	2.080	2.831	3.819
22	2.074	2.819	3.792
23	2.069	2.807	3.767
24	2.064	2.797	3.745
25	2.060	2.787	3.725
26	2.056	2.779	3.707
27	2.052	2.771	3.690
28	2.048	2.763	3.674
29	2.045	2.756	3.659
30	2.042	2.750	3.646
40	2.021	2.704	3.551
60	2.000	2.660	3.460
120	1.980	2.617	3.373
∞	1.960	2.576	3.291



$$\bar{X}_1 - \bar{X}_2 \pm (t_{conf})(s_{\bar{X}_1-\bar{X}_2})$$

Παράδειγμα



$$\bar{X}_1 - \bar{X}_2 \pm (t_{conf})(s_{\bar{X}_1-\bar{X}_2})$$

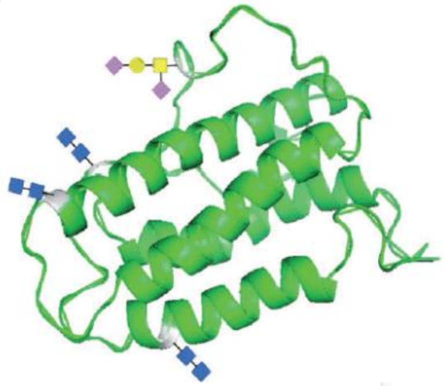
$$\bar{X}_1 = \frac{\sum X_1}{n_1} = \frac{66}{6} = 11 \qquad \bar{X}_2 = \frac{\sum X_2}{n_2} = \frac{36}{6} = 6$$

2.228

$$s_{\bar{X}_1-\bar{X}_2} = \sqrt{\frac{s_P^2}{n_1} + \frac{s_P^2}{n_2}} = \sqrt{\frac{16.2}{6} + \frac{16.2}{6}} = \sqrt{\frac{32.4}{6}} = \sqrt{5.4} = 2.32$$

$$s_P^2 = \frac{SS_1 + SS_2}{n_1 + n_2 - 2} = \frac{90+72}{6+6-2} = \frac{162}{10} = 16.2$$

$$SS_1 = \sum X_1^2 - \frac{(\sum X_1)^2}{n_1} \qquad SS_2 = \sum X_2^2 - \frac{(\sum X_2)^2}{n_2}$$

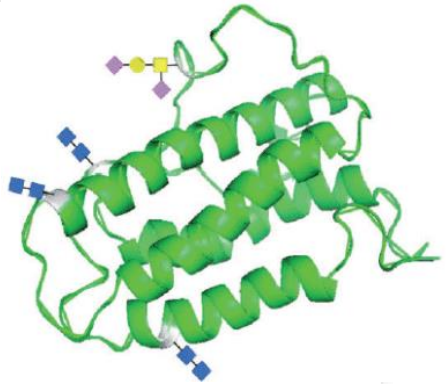


Παράδειγμα

$$\bar{X}_1 - \bar{X}_2 \pm \left(t_{conf}\right)\left(s_{\bar{X}_1-\bar{X}_2}\right)$$

$$5 \pm (2.228)(2.32) = 5 \pm 5.17 = \begin{cases} 10.17 \\ -0.17 \end{cases}$$

Παράδειγμα



$$5 \pm (2.228)(2.32) = 5 \pm 5.17 = \begin{cases} 10.17 \\ -0.17 \end{cases}$$

- Μακροπρόθεσμα το 95% όλων των διαστημάτων εμπιστοσύνης, παρομοίων με αυτά που υπολογίσαμε, θα περιλαμβάνει την άγνωστη διαφορά των μέσων του πληθυσμού.
- Η πραγματική επίδραση (διαφορά των μέσων) δεν αναμένεται να είναι μικρότερη των -0.17 λεπτών, ούτε μεγαλύτερη από 10.17 λεπτά
- Το πάνω όριο δηλώνει πλεονέκτημα της EPO
- Το κάτω όριο δηλώνει πιθανή αρνητική επίδραση στην αντοχή
- Χαρακτηρίζεται αρκετά ανακριβές (ενδεχομένως λόγω του μικρού μεγέθους των δειγμάτων)
- Προτείνεται αύξηση του μεγέθους των δειγμάτων

Σύνδεσμοι σε online συγγράμματα

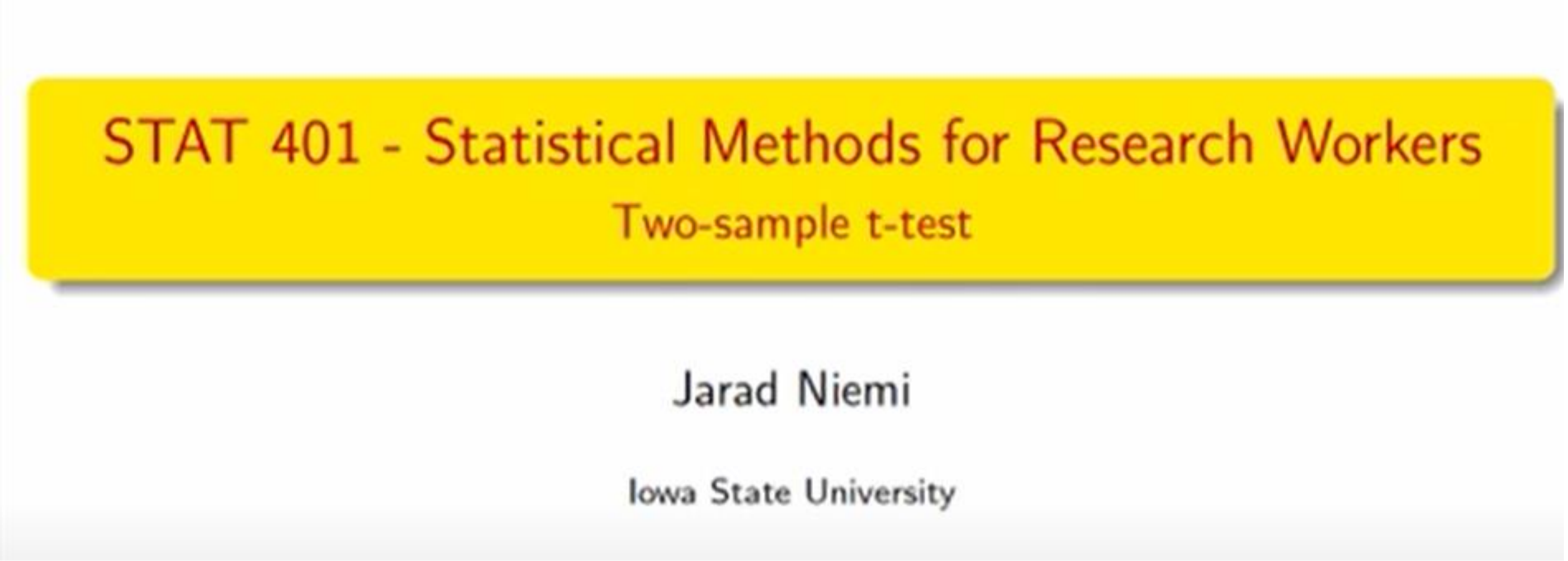
- https://www.westga.edu/academics/research/vrc/assets/docs/TwoSampleProblems_LectureNotes.pdf

Two-Sample Problems

Diana Mindrila, Ph.D.
Phoebe Balentyne, M.Ed.

Σύνδεσμοι σε online video

- https://youtu.be/anu13FU4Gow



R code

- t.test(δείγμα1, δείγμα2)

Back up

Τέσσερα δυνατά αποτελέσματα:

Τι ισχύει πραγματικά

Τι εγώ πιστεύω

	Κατάσταση του H0	
Απόφαση	Αληθής H0	Ψευδής H0
Διατήρηση του H0	1) Σωστή απόφαση	3) Σφάλμα τύπου II (αστοχία)
Απόρριψη του H0	2) Σφάλμα τύπου I (εσφαλμένος συναγερμός)	4) Σωστή απόφαση