



Autor: **Conf. Dr. Bondor Cosmina-Ioana**

Testarea ipotezelor



ALWAYS



SEEK



KNOWLEDGE

Obiective

Testarea ipotezelor

Testul Student T

Testul hi-pătrat

Erori în testarea ipotezelor și puterea testului

Intervale de încredere sau testarea ipotezelor

Exerciții

Compararea a două grupuri cu metoda intervalelor de încredere

- comparăm două populații P_1 și P_2 ,
 - mediile (necunoscute) μ_1 și μ_2 ,

eșantionare



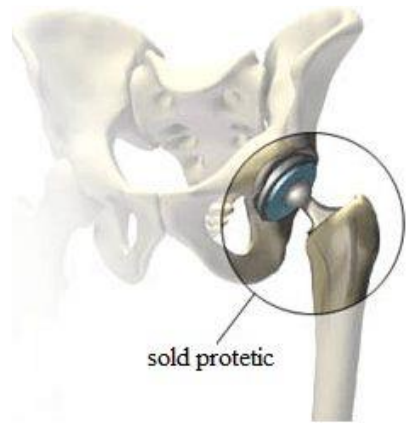
- extragem aleator două eșantioane cu n_1, n_2 subiecți
 - \bar{X}_1, \bar{X}_2 mediile
 - s_1, s_2 varianțele.

inferență



- estimăm (aproximăm) μ_1 și μ_2 media aritmetică necunoscută în populațiile P_1 și P_2 , cu \bar{X}_1 și \bar{X}_2 media aritmetică cunoscută din eșantion.

- Obiectiv: Compararea timpului de supraviețuire a protezei de șold în artroplastia totală de șold efectuată utilizând materialul A versus B versus C



când începem cercetarea nu cunoaștem supraviețuirea medie a protezei de șold realizată cu materialele A, B sau C.

scopul nostru: să găsim care material are cea mai îndelungată supraviețuire în artroplastia totală de șold

- Obiectiv: Să comparăm timpul mediu de supraviețuire al protezei de șold realizată cu materialele A, B sau C
- Dacă găsim o diferență de 8 luni între materiale putem declara clinic că supraviețuirea este mai îndelungată cu respectivul material
 - ne propunem clinic o diferență de minim 8 luni



- Obiectiv: Compararea timpului de supraviețuire a protezei de șold în artroplastia totală de șold efectuată utilizând materialul A versus B versus C
- ne propunem o diferență importantă clinic de 8 luni



- Metodă: Ne propunem să realizăm un studiu unde să demonstrăm că există o diferență de minim 8 luni între supraviețuirea protezei de șold realizată cu materialul A versus B versus C:
 - Calculăm dimensiunea necesară a eșantionului: selectăm aleator 3 grupuri de 100 de pacienți



material A



material B



material C



cu acest număr de pacienți putem demonstra o diferență ≥ 12 luni

Calculăm intervalele de încredere ale mediilor

- cu formula:

$$[\bar{X} - 1,96 \frac{s}{\sqrt{n-1}}, \bar{X} + 1,96 \frac{s}{\sqrt{n-1}}]$$

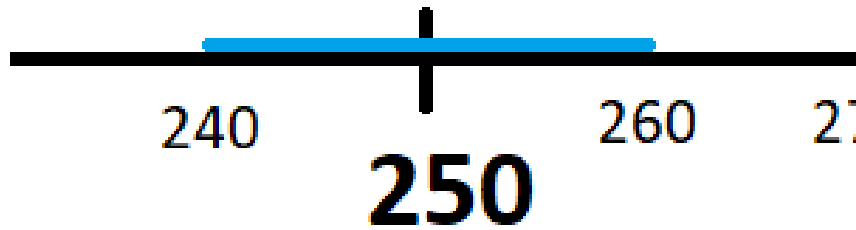
unde

\bar{X} - media aritmetică

s – deviația standard

n – numărul de indivizi luați în studiu

material C



- timpul mediu de supraviețuire: 250 luni
- 95% Interval de încredere 240 – 260 luni

Populație = toți
pacienții cu
artroplastie de șold

Interpretare: la întreaga populație timpul mediu de supraviețuire al protezei de șold executată cu materialul C va fi între 240 luni și 260 luni cu 5% eroare

!!! măsurăm timpul la toată populația - practic destul de greu!!!, dar să presupunem că facem asta

În cel mai rău caz, la toată populația timpul mediu de supraviețuire al protezei de șold executată cu materialul C poate fi 240 luni

În cel mai bun caz, la toată populația timpul mediu de supraviețuire al protezei de șold executată cu materialul C poate fi 260 luni

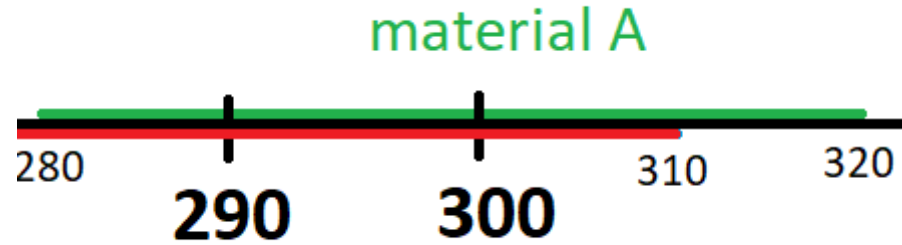


- timpul mediu de supraviețuire: 290 luni
- 95% Interval de încredere 270 – 310 luni

Interpretare: la întreaga populație timpul mediu de supraviețuire al protezei de șold executată cu materialul B va fi între 270 luni și 310 luni cu 5% eroare (nu știm exact cât va fi)

În cel mai rău caz, la toată populația timpul mediu de supraviețuire al protezei de șold executată cu materialul A poate fi 270 luni

În cel mai bun caz, la toată populația timpul mediu de supraviețuire al protezei de șold executată cu materialul A poate fi 310 luni



- timpul mediu de supraviețuire: 300 luni
- 95% Interval de încredere 280 – 320 luni

Interpretare: la întreaga populație timpul mediu de supraviețuire al protezei de șold executată cu materialul A va fi între 280 luni și 320 luni cu 5% eroare (nu știm exact cât va fi)

În cel mai rău caz, la toată populația timpul mediu de supraviețuire al protezei de șold executată cu materialul A poate fi 280 luni

În cel mai bun caz, la toată populația timpul mediu de supraviețuire al protezei de șold executată cu materialul A poate fi 320 luni

Rezultate:



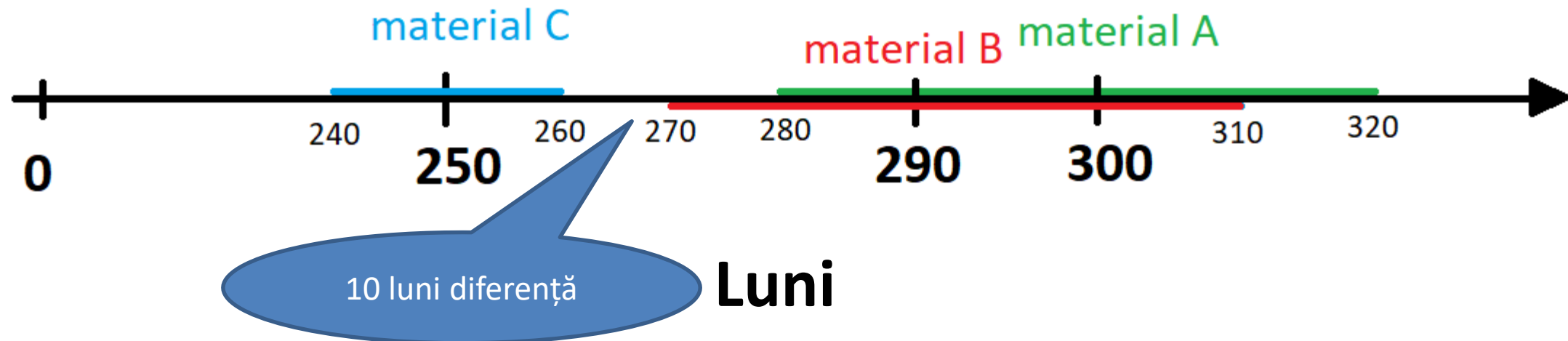
material C:
media: 250 luni,
95% Interval de încredere 240-260 luni

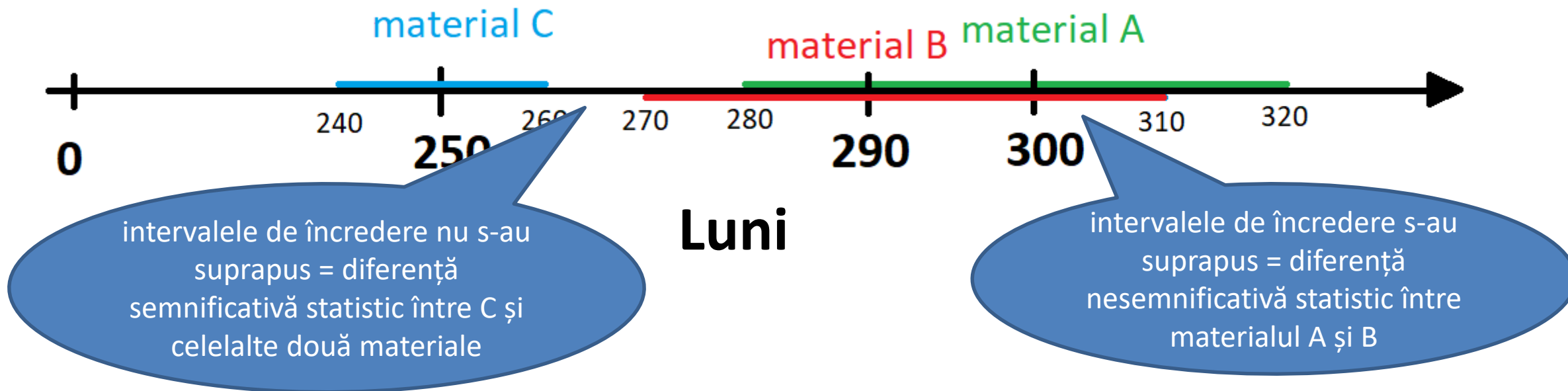


material B:
media: 290 luni,
95% Interval de încredere 270 -310 luni



material A:
media: 300 luni,
95% Interval de încredere 280 -320 luni





- **Concluzie 1:** Acest studiu a demonstrat că materialul C are o supraviețuire a protezei de șold semnificativ statistic mai mică decât materialul A sau B (95% IC pentru materialul C nu se suprapune cu 95% IC pentru materialul A sau B. Orice valoare din 95% IC pentru materialul C < orice valoare din 95% IC pentru materialele A sau B. Diferența de supraviețuire între materialul C și B este egală sau mai mare de 10 luni. Diferența de supraviețuire între materialul C și A este egală sau mai mare de 20 luni
- **Concluzie 2:** Nu am reușit să demonstrăm că materialul B are o supraviețuire a protezei de șold semnificativ statistic mai mică decât materialul A (95% IC pentru materialul B se suprapune cu 95% IC pentru materialul A).

Comparația mediilor aritmetice a două eșantioane cu ajutorul testelor statistice

Compararea a două medii

- General:

populația P_1

populația P_2

\bar{X}_1 = media eșantion unu,

\bar{X}_2 = media eșantion doi,

s_1 = varianța eșantion unu

s_2 = varianța eșantion doi

- Exemplu:

= proteză de șold cu material A

= proteză de șold cu material B;

= media eșantion material A 300 luni,

= media eșantion material B 290 luni,

= varianța eșantion material A 100 luni

= varianța eșantion material B 100 luni



material A



material B

Compararea a două medii

- General:

populația P_1

populația P_2

\bar{X}_1 = media eșantion unu,

\bar{X}_2 = media eșantion doi,

s_1 = varianța eșantion unu

s_2 = varianța eșantion doi

- Exemplu:

= proteză de șold cu material A

= proteză de șold cu material B;

= media eșantion material A 300 luni,

= media eșantion material B 290 luni,

= varianța eșantion material A 100 luni

= varianța eșantion material B 100 luni



material A



material B

Pas 1 –test statistic: formularea ipotezelor

General:

- H_0 – ipoteza nulă: nu există o diferență semnificativă statistic între media grupului 1 și media grupului 2 – pe aceasta o presupunem prin absurd adevărată
- H_1 – ipoteza alternativă: există o diferență semnificativă statistic între media grupului 1 și media grupului 2 – pe aceasta dorim să o demonstrăm defapt

Exemplu:

- H_0 - ipoteza nulă: nu există o diferență semnificativă statistic între media timpului de supraviețuire a protezei de șold executată cu materialul A și media timpului de supraviețuire a protezei de șold executată cu materialul B
- H_1 – ipoteza alternativă: există o diferență semnificativă statistic între media timpului de supraviețuire a protezei de șold executată cu materialul A și media timpului de supraviețuire a protezei de șold executată cu materialul B

Pas 1 –test statistic: formularea ipotezelor

General:

- H_0 – ipoteza nulă: nu există o diferență semnificativă statistic între media grupului 1 și media grupului 2
- H_1 – ipoteza alternativă: există o diferență semnificativă statistic între media grupului 1 și media grupului 2

Exemplu:

- H_0 - ipoteza nulă: nu există o diferență semnificativă statistic între media timpului de supraviețuire a protezei de șold executată cu materialul A și media timpului de supraviețuire a protezei de șold executată cu materialul B
- H_1 - ipoteza alternativă: există o diferență semnificativă statistic între media timpului de supraviețuire a protezei de șold executată cu materialul A și media timpului de supraviețuire a protezei de șold executată cu materialul B

Pas 1 –test statistic: formularea ipotezelor

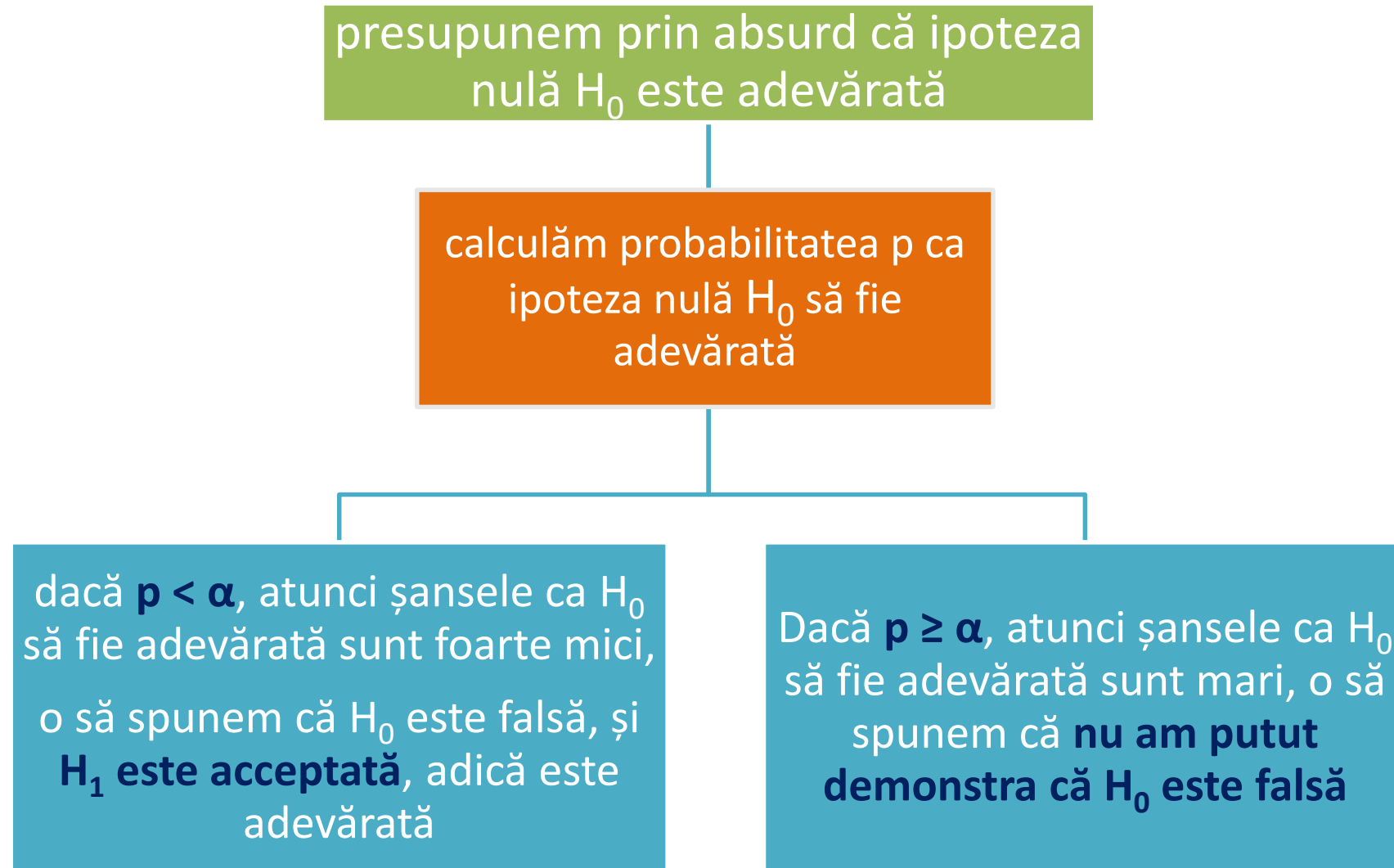
- H_0 – ipoteza nulă: nu există o diferență semnificativă statistic între media grupului 1 și media grupului 2
- H_1 – ipoteza alternativă: există o diferență semnificativă statistic între media grupului 1 și media grupului 2

“statistica” – estimăm statistic
mediile populației

numele ipoteza nulă H_0
vine de la diferența 0
între medii, adică medii
egale

Alternativa = cealaltă situație în
care diferența dintre medii nu
este 0

Metodă:



Pas 2 Test statistic – Alegerea testului statistic adecvat (din cele peste 100 de teste existente)

General:

- Ce comparăm?
 - medii,
 - varianțe,
 - procente
 -
- Cum este distribuția?
 - normală
 - nu este normală,
 - binomială
 - ...
- Cum sunt eșantioanele?
 - independente
 - dependente
- Cât de mulți pacienți avem în eșantioane?
 - puțini (<30) – eșantioane mici
 - mulți – eșantioane mari
- Cum sunt varianțele?
 - egale
 - neegale
- Câte eșantioane trebuie comparate?
 - unul,
 - două,
 - trei
 - mai multe

Independente sau dependente

- **Eșantioane independente**

- nu există nicio legătură între subiecții din cele două eșantioane

- **Eșantioane dependente**

- subiecții sunt

- perechi

- frați gemeni,
 - potriviți după
 - » vârstă,
 - » gen
 - » etc.,

- sunt aceeși subiecți evaluați de mai multe ori

- înainte și după tratament
 - înainte și după procedură

- Exemplu:

- distribuția normală

- de obicei ar trebui să verificăm distribuția,
 - dar acum facem aceste presupuneri că există în ambele grupuri,

- varianțe egale

- de obicei ar trebui să verificăm,
 - facem această presupunere

- sunt comparate două medii aritmetice,

- grupurile sunt independente

- vom alege:

- **Testul t Student pentru eșantioane independente în cazul varianțelor egale**

Pas 3. Selectia nivelului de semnificație

General:

- α – nivelul de semnificație
 - probabilitatea de a respinge ipoteza nulă atunci când aceasta este de fapt adevărată
- **Cel mai des α este ales să fie 0,05**
 - (corespunde $Z_{\alpha}=1.96$)

Exemplu:

- alegem $\alpha = 0,05$

Pas 3. Selectia nivelului de semnificație

General:

- α – nivelul de semnificație
 - probabilitatea de a respinge ipoteza nulă atunci când aceasta este de fapt adevărată
- **Cel mai des α este ales să fie 0,05**
 - (corespunde $Z_{\alpha}=1.96$)
- p-value
 - probabilitatea de a obține un rezultat mai extrem decât rezultatul obținut din datele observate.
 - Dacă $p < \alpha$, atunci vom respinge ipoteza nulă: demonstrăm că există diferență între mediile aritmetice și spunem că testul este semnificativ: avem diferență între medii
 - Dacă $p \geq \alpha$, atunci nu vom respinge ipoteza nulă: nu am reușit să demonstrăm că există diferență între mediile aritmetice și spunem că testul este nesemnificativ: nu avem diferență între medii

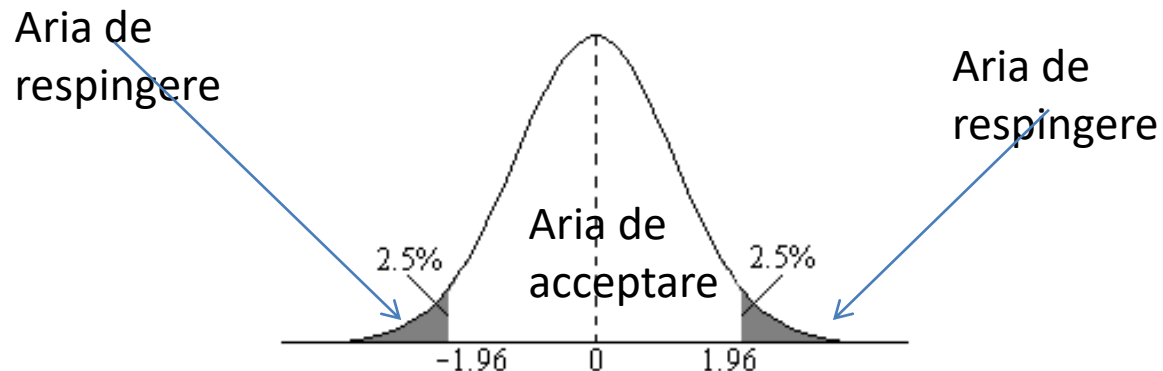
Exemplu:

- alegem $\alpha = 0,05$

Pas 4. Determinarea valorii pe care trebuie să o atingă parametrul de testare pentru a fi declarat semnificativ

General:

- **valoarea critică**
 - ex. pentru t-test: t_{α} , pentru Z test: Z_{α}
- Aria de respingere –intervalul critic
 - intervalul unde H_0 poate fi respinsă
 - diferit în funcție de distribuția testului și uneori de gradele de libertate
 - ex. $(-\infty, -Z_{\alpha}] \cup [Z_{\alpha}, \infty)$
- Intervalul de acceptare:
 - intervalul unde H_0 trebuie acceptată
 - ex. $(-Z_{\alpha}, Z_{\alpha})$



Exemplu:

test t –distribuția Student– **valoarea critică** este diferită în funcție de gradele de libertate

- $t_{\alpha} = 1,97$ = **valoarea critică pentru 100+100-2 grade de libertate**
- intervalul de respingere $(-\infty, -t_{\alpha}] \cup [t_{\alpha}, \infty) = (-\infty, -1,97] \cup [1,97, \infty)$
- intervalul de acceptare: $(-t_{\alpha}, t_{\alpha}) = (-1,97, 1,97)$

= un interval în jurul lui 0, dacă găsim o diferență în acest interval înseamnă că e o diferență mică, H_0 va fi acceptată (decidem că nu este diferență între medii)

Pas 4. Determinarea valorii pe care trebuie să o atingă parametrul de testare pentru a fi declarat semnificativ

General:

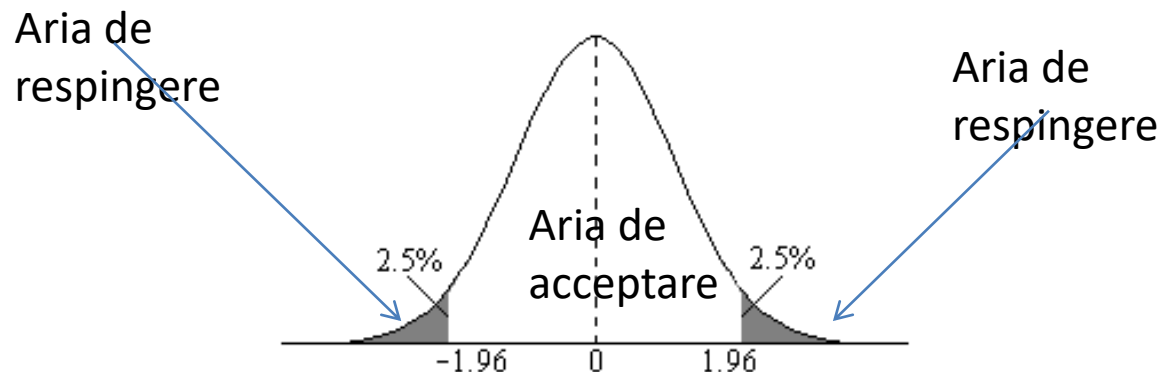
- **valoarea critică**
 - ex. pentru t-test t_{α} , pentru Z test Z_{α})
- Aria de respingere –intervalul critic
 - intervalul unde H_0 poate fi respinsă
 - ex. $(-\infty, -Z_{\alpha}] \cup [Z_{\alpha}, \infty)$
- Intervalul de acceptare:
 - intervalul unde H_0 trebuie acceptată
 - ex. $(-Z_{\alpha}, Z_{\alpha})$

Exemplu:

test t –distribuția Student– **valoarea critică** este diferită în funcție de gradele de libertate

- $t_{\alpha} = 1,97 =$ **valoarea critică pentru 100+100-2 grade de libertate**
- intervalul de respingere $(-\infty, -1,97] \cup [1,97, \infty)$
- intervalul de acceptare: $(-t_{\alpha}, t_{\alpha}) = (-1,97, 1,97)$

= un interval în jurul lui 0, dacă găsim o diferență în acest interval înseamnă că e o diferență mică, H_0 va fi acceptată (decidem că nu este diferență între medii)



Pas 5: calcularea parametrului

General:

- fiecare test statistic are o formulă a parametrului testului
- două medii, atunci $\bar{X}_1 - \bar{X}_2$ trebuie să fie în formulă

dacă diferența dintre \bar{X}_1 și \bar{X}_2 este mică, aproape de 0 atunci și parametrul testului este mic

Exemplu:

- parametrul testului:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}}}$$

diferența dintre medii.

material A: $\bar{X}_1 = 300$, $s_1 = 100$

material B: $\bar{X}_2 = 290$, $s_2 = 100$

$$t = \frac{300 - 290}{\sqrt{\frac{100^2}{100} + \frac{100^2}{100}}} = \frac{10}{\sqrt{200}} = \frac{10}{14,14} = 0,71,$$

p=0,48

p poate fi calculat numai cu calculatorul – este o arie

Pas 5: calcularea parametrului

General:

- fiecare test statistic are o formulă a parametrului testului
- două medii, atunci $\bar{X}_1 - \bar{X}_2$ trebuie să fie în formulă

dacă diferența dintre \bar{X}_1 și \bar{X}_2 este mică, aproape de 0
atunci și parametrul testului este mic

Exemplu:

- parametrul testului:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}}}$$

diferența dintre medii.

material A: $\bar{X}_1 = 300$, $s_1 = 100$

material B: $\bar{X}_2 = 290$, $s_2 = 100$

$$t = \frac{300-290}{\sqrt{\frac{100^2}{100} + \frac{100^2}{100}}} = \frac{10}{\sqrt{200}} = \frac{10}{14,14} = 0,71,$$

p=0,48

p poate fi calculat numai cu
calculatorul – este o arie

Pas 6. Concluzia

General:

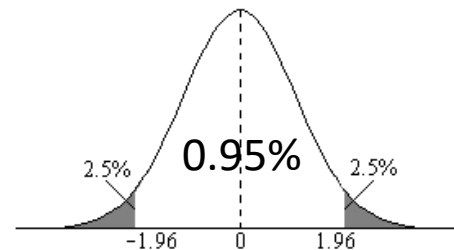
Dacă parametrul testului aparține intervalului de acceptare --> nu am reușit să respingem H_0 ($p \geq 0,05$)

Dacă parametrul testului aparține intervalului de respingere --> am reușit să respingem H_0 , acceptăm H_1 ($p < 0,05$)

la fel cu

Dacă $p \geq 0,05$ nu am reușit să respingem H_0

Dacă $p < 0,05$ am reușit să respingem H_0 , acceptăm H_1



(b) Two-tailed test

Exemplu:

Dacă $t \in (-t_\alpha, t_\alpha)$ – intervalul de acceptare - nu am reușit să respingem H_0 .

Dacă $t \in (-\infty, -t_\alpha] \cup [t_\alpha, \infty)$ – intervalului de respingere --> respingem H_0 , acceptăm H_1

la fel cu

If $p \geq 0,05$, nu am reușit să respingem H_0

If $p < 0,05$ respingem H_0 , acceptăm H_1

Pas 6. Concluzia

General:

Dacă parametrul testului \in intervalului de acceptare \rightarrow nu am reușit să respingem H_0

Dacă parametrul testului \in intervalului de respingere \rightarrow am reușit să respingem H_0 , acceptăm H_1

la fel cu

Dacă $p \geq 0,05$ nu am reușit să respingem H_0

Dacă $p < 0,05$ am reușit să respingem H_0 , acceptăm H_1

Exemplu:

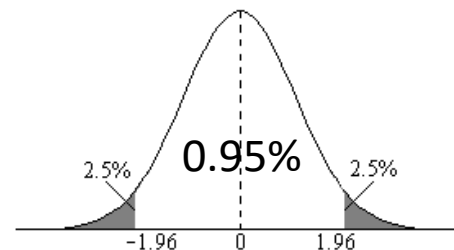
Dacă $t \in (-t_\alpha, t_\alpha)$ – intervalul de acceptare - nu am reușit să respingem H_0 .

Dacă $t \in (-\infty, -t_\alpha] \cup [t_\alpha, \infty)$ – intervalului de respingere \rightarrow respingem H_0 , acceptăm H_1

la fel cu

If $p \geq 0,05$, nu am reușit să respingem H_0

If $p < 0,05$ respingem H_0 , acceptăm H_1



(b) Two-tailed test

Exemplu:

$$-1,97 < t = 0,71 < 1,97$$

nu am reușit să respingem H_0

nu există o diferență semnificativă statistic între media timpului de supraviețuire a protezei de șold executată cu materialul A și media timpului de supraviețuire a protezei de șold executată cu materialul B

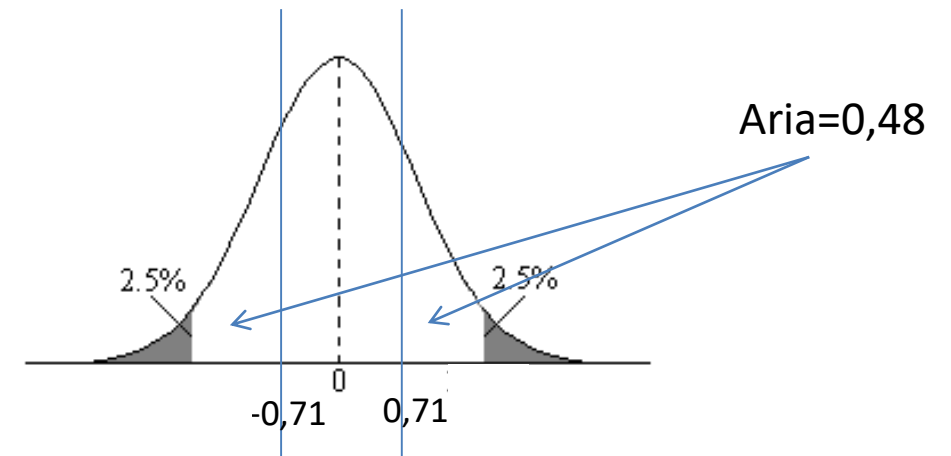
Aria de sub curbă pentru $t = 0,71$ este 0,48, $p = 0,48 > 0,05$

Nu e corect să spunem

~~Acceptăm ipoteza nulă~~

E corect să spunem:

Nu am reușit să respingem ipoteza nulă



Exemplu 2

material C comparat cu material B

Pas 1 –test statistic: formularea ipotezelor

General:

- H_0 – ipoteza nulă: nu există o diferență semnificativă statistic între media grupului 1 și media grupului 2
- H_1 – ipoteza alternativă: există o diferență semnificativă statistic între media grupului 1 și media grupului 2

Exemple:

- H_0 - ipoteza nulă: nu există o diferență semnificativă statistic între media timpului de supraviețuire a protezei de șold executată cu materialul B și media timpului de supraviețuire a protezei de șold executată cu materialul C
- H_1 – ipoteza alternativă: există o diferență semnificativă statistic între media timpului de supraviețuire a protezei de șold executată cu materialul B și media timpului de supraviețuire a protezei de șold executată cu materialul C

- Exemplu:

- distribuția normală

- de obicei ar trebui să verificăm distribuția,
 - dar acum facem această presupunere că există în ambele grupuri,

- varianțe egale

- de obicei ar trebui să verificăm,
 - facem această presupunere că sunt egale

- sunt comparate două medii aritmetice,

- grupurile sunt independente

- vom alege:

- **Testul t Student pentru eșantioane independente în cazul varianțelor egale**

Pas 3. Selectia nivelului de semnificație

General:

- α – nivelul de semnificație
 - probabilitatea de a respinge ipoteza nulă atunci când aceasta este de fapt adevărată
- **Cel mai des α este ales să fie 0,05**
 - (corespunde $Z_{\alpha}=1.96$)
- p-value
 - probabilitatea de a obține un rezultat mai extrem decât rezultatul obținut din datele observate.
 - Dacă $p < \alpha$, atunci vom respinge ipoteza nulă: demonstrăm că există diferență între mediile aritmetice și spunem că testul este semnificativ: avem diferență între medii
 - Dacă $p \geq \alpha$, atunci nu vom respinge ipoteza nulă: nu am reușit să demonstrăm că există diferență între mediile aritmetice și spunem că testul este nesemnificativ: nu avem diferență între medii

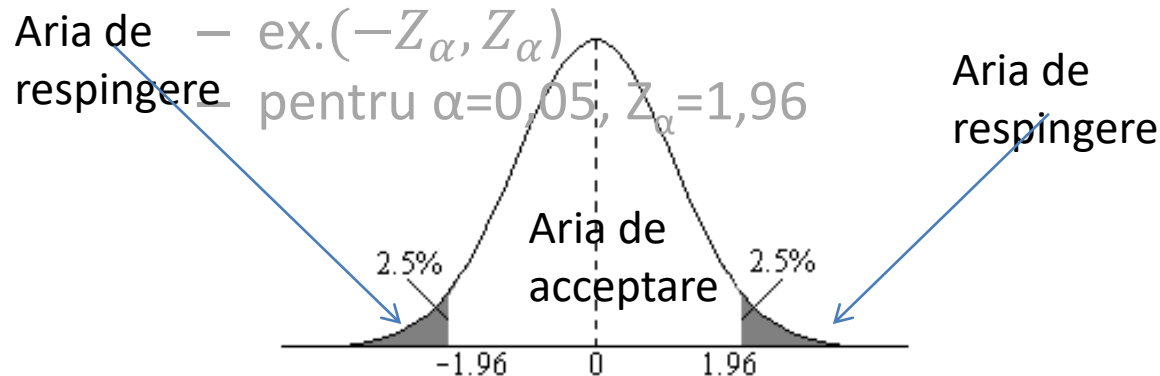
Exemplu:

- alegem $\alpha = 0,05$

Pas 4. Determinarea valorii pe care trebuie să o atingă parametrul de testare pentru a fi declarat semnificativ

General:

- **valoarea critică**
 - ex. pentru t-test t_α , pentru Z test Z_α)
- **Aria de respingere –intervalul critic**
 - intervalul unde H_0 poate fi respinsă
 - diferit în funcție de distribuția testului și uneori de gradele de libertate
 - ex. $(-\infty, -Z_\alpha] \cup [Z_\alpha, \infty)$
- **Intervalul de acceptare:**
 - intervalul unde H_0 trebuie acceptată



Exemplu:

- $t_\alpha = 1,97$ = **valoarea critică pentru 100+100-2 grade de libertate**
- intervalul de respingere $(-\infty, -t_\alpha] \cup [t_\alpha, \infty) = (-\infty; -1,97] \cup [1,97; \infty)$
- intervalul de acceptare: $(-t_\alpha, t_\alpha) = (-1,97; 1,97)$

Pas 5: calcularea parametrului

General:

- fiecare test statistic are o formulă a parametrului testului
- două medii, atunci $\bar{X}_1 - \bar{X}_2$ trebuie să fie în formulă

Exemplu:

- parametrul testului:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}}}$$

material A: $\bar{X}_1 = 250$, $s_1 = 50$

material B: $\bar{X}_2 = 290$, $s_2 = 100$

$$t = \frac{250-290}{\sqrt{\frac{50^2}{100} + \frac{100^2}{100}}} = \frac{40}{\sqrt{125}} = \frac{10}{11,18} = 3.58$$

p=0,0004

Pas 6. Concluzia

General:

Dacă parametrul testului \in intervalului de acceptare \rightarrow nu am reușit să respingem H_0

Dacă parametrul testului \in intervalului de respingere \rightarrow am reușit să respingem H_0 , acceptăm H_1

la fel cu

Dacă $p \geq 0,05$ nu am reușit să respingem H_0

Dacă $p < 0,05$ am reușit să respingem H_0 , acceptăm H_1

Exemplu:

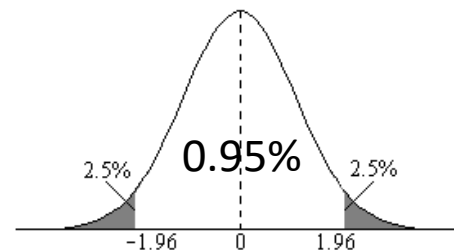
Dacă $t \in (-t_\alpha, t_\alpha)$ – intervalul de acceptare - nu am reușit să respingem H_0 .

Dacă $t \in (-\infty, -t_\alpha] \cup [t_\alpha, \infty)$ – intervalului de respingere \rightarrow respingem H_0 , acceptăm H_1

la fel cu

If $p \geq 0,05$, nu am reușit să respingem H_0

If $p < 0,05$ respingem H_0 , acceptăm H_1



(b) Two-tailed test

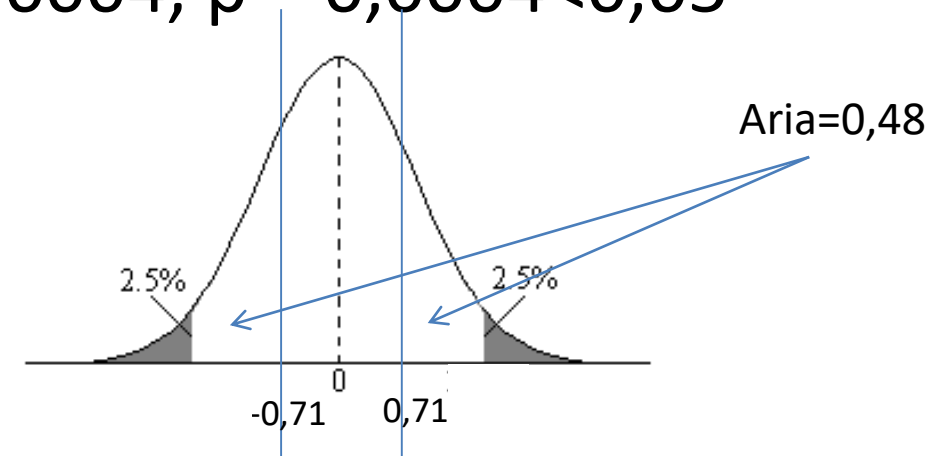
Exemplu:

$$3,58 \in (-\infty; -1,97] \cup [1,97; \infty)$$

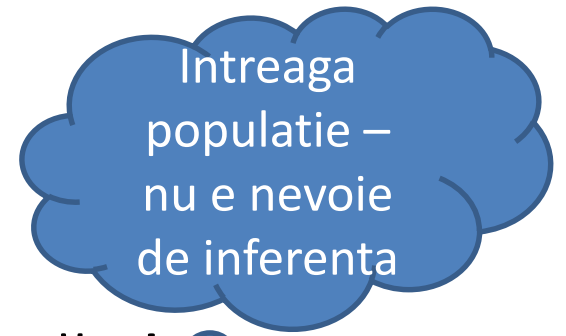
respingem H_0 , acceptăm H_1

Există diferență semnificativă statistic între media timpului de supraviețuire a protezei de șold executată cu materialul B și media timpului de supraviețuire a protezei de șold executată cu materialul C

Aria de sub curbă pentru $t = 3,58$ este 0,0004, $p = 0,0004 < 0,05$



Scenariu – Amputații

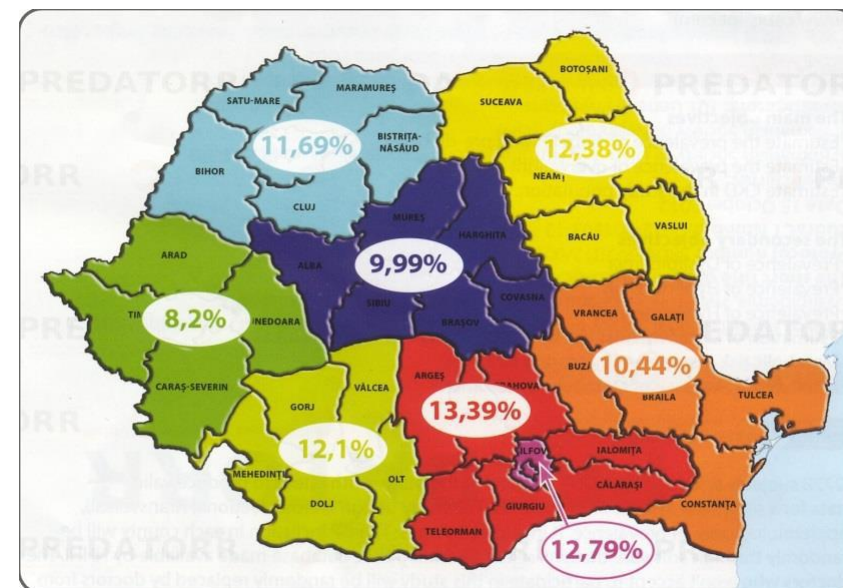


- Scopul studiului a fost de a evalua frecvența amputărilor datorate diabetului în România. Au fost analizate datele din toate spitalele din țară în 5 ani.
- Numărul total de amputări datorate diabetului zaharat a fost de 24.312, 70,3% cu diabet zaharat de tip 2.
- La fiecare 21 de minute există o amputare datorată diabetului
- Veresiu IA, Iancu SS, Bondor C. Trends in diabetes-related lower extremities amputations in Romania—A five year nationwide evaluation. Diabetes Research and Clinical Practice. August 2015;109(2):293–8.

Scenariu – studiul Predatorr

- Scop: primul studiu de măsurare a prevalenței diabetului în România
- Metoda de eșantionare:
 - stratificat,
 - transversal,
 - clusterizat
 - Selecție aleatorie.
- 2728 de participanți
- Prevalența diabetului a fost de 11,6%

Un esantion –
e nevoie de
inferenta



Intervale de incredere

- Vrem să estimăm prevalența diabetului în România.
- Într-un studiu cu 2728 de participanți, frecvența diabetului zaharat 11,6%

$$f=0,116, Z_{\alpha}=1,96$$

$$\left[0,116 - 1,96 \sqrt{\frac{0,116(1 - 0,116)}{2728}}; 0,116 + 1,96 \sqrt{\frac{0,116(1 - 0,116)}{2728}} \right]$$

$$[0,116-0,006;0,116+0,006]$$

[0,110;0,122] – intervalul de încredere de 95%

(frecvența diabetului în populația română va fi între 11%-12,2% cu o eroare de 5%, cel puțin acolo sunt 95% dintre frecvențe dacă repetăm studiul pe toate esantioanele posibile de 2728 de indivizi cu diabet)

$$\left[f - Z_{\alpha} \sqrt{\frac{f(1-f)}{n}}; f + Z_{\alpha} \sqrt{\frac{f(1-f)}{n}} \right]$$

Scenariu

Întrebare: La persoanele cu infarct miocardic, colesterolul mare este asociat cu repetarea infarctului miocardic?

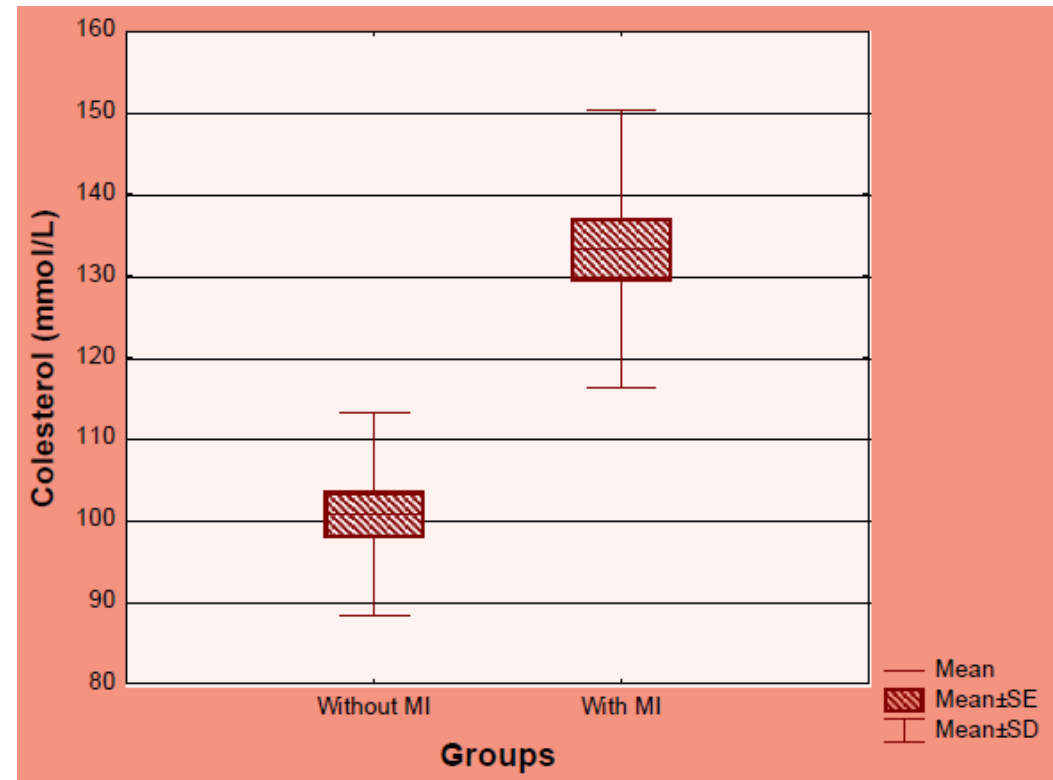
Selectăm aleator 40 de persoane.

Colesterolul la pacienții **cu repetare** a infarctului miocardic

($n_1=20$): $133 \pm 17 \text{ mmol/L}$ ($\bar{X}_1 \pm s_1$) medie \pm dev.st.

Colesterolul la pacienții **fără repetare** a infarctului

Miocardic ($n_2=20$): $101 \pm 12 \text{ mmol/L}$ ($\bar{X}_2 \pm s_2$)



Compararea
a doua
medii
implica un
test statistic

Multumesc!!!