



UMFT

Universitatea de
Medicină și Farmacie
„Victor Babeș”
din Timișoara

Disciplina de Informatica si Biostatistica Medicala

Teste statistice

Dr. Mirela Frandes

mirela.frandes@umft.ro



Tipuri de teste statistice

	Comparare	De verificat
Conformitate	O valoare obtinuta cu o valoare data	
Omogenitate	2 Parametrii de dispersie	Cele 2 esantioane provin de la aceeasi populatie
Semnificatie	2 medii sau proportii	Cele 2 esantioane provin de la aceeasi populatie
Concordanta	O distributie experimentala cu o distributie teoretica sau e distributii experimentale	
Independenta	Independenta seriilor de valori	
Corelatie	Semnificatia parametrilor	Demonstrarea existentei unei legaturi intre 2 variabile



Teste statsitice

- Comparare:
 - O valoare de medie obtinuta cu o valoare teoretica
 - 2 valori medii pentru serii independente
 - > 2 valori medii
 - 2 dispersii
 - > 2 dispersii
 - O distributie experimentală cu o distributie teoretică



Teste statistice

- Etapele aplicării

Formularea ipotezelor medicale

Formularea ipotezelor statistice

Calcularea statisticii testului

Decizia statistica

Formularea concluziei



Teste statistice

- Semnificatia statistica
 - este un element foarte important
 - asigura ca rezultatul obtinut este real si nu a fost obtinut din intamplare
- Ipotezele statistice
 - postuleaza existenta unei diferente / asocieri semnificative intre parametrii statistici
- Semnificativ dpdv statistic $\neq$ semnificativ dpdv medical



Teste statistice

- Diferente statistice

- Semnificative

- Probabilitatea ca rezultatele să fi apărut din întâmplare sunt reduse
 - Diferențele pot fi atribuite altor cauze

- Nesemnificative

- Probabilitate mare ca rezultatele să fi apărut din întâmplare
 - O cauză poate fi variabilitatea de eșantionare



Ipoteze statistice

- Ipoteza initiala (H_0)

➤ Este o ipoteza de non-diferenta:

“**Nu exista** diferente semnificative...”

- Ipoteza alternativa (H_a)

➤ Este o ipoteza de diferenta:

“**Exista** diferente semnificative...”

➤ Este opusa ipotezei initiale



Teste statistice

- Pragul de semnificatie (not. α)
 - Riscul/probabilitatea de a comite o eroare statistica
 - Conventie: risc acceptabil $\alpha = 0,05$ (5%)
 - Probabilitatea incepand cu care vom considera ca diferentele observate sunt semnificative



UMFT

Universitatea de
Medicină și Farmacie
„Victor Babeș”
din Timișoara

Decizia statistica

p	Diferente	Diferente	Ipoteza initiala (H_0)
$> 0,05$	nesemnificative	N	acceptata
$< 0,05$	semnificative	S	rejectata
$< 0,01$	foarte semnificative	TS	rejectata
$< 0,001$	extrem de semnificative	ES	rejectata



Erorile statistice

Tip I

➤ α

- Eroarea care consta in a rejecta o ipoteza adevarata
- Ipoteza initiala H_0 -> este adevarata, dar nu este acceptata

Tip II

➤ β

- Eroarea care consta in a accepta o ipoteza falsa
- Ipoteza initiala H_0 -> este falsa, dar acceptata



Erorile statistice

		Situatia reala	
		H0 adevarata	H0 falsa
Decizia statistica	Acceptare H0	$p=1-\alpha$	$p=\beta$
	Rejectare H0	$p=\alpha$	$p=1-\beta$



UMFT

Universitatea de
Medicină și Farmacie
„Victor Babeș”
din Timișoara

Caracteristicile unui test statistic

Nivelul de incredere ($1 - \alpha$)

- Capacitatea de a accepta o ipoteza cand ea este adevarata

Puterea unui test ($1 - \beta$)

- capacitatea de rejecta o ipoteza cand ea este falsa



UMFT

Universitatea de
Medicină și Farmacie
„Victor Babeș”
din Timișoara

Teste parametrice și neparametrice

1. Teste parametrice

- Pentru variabile cu o **distributie cunoscuta (distributia normal)**
- Sunt mai ”puternice” decat testele neparametrice

2. Tests neparametrice

- Pentru o **distributie necunoscuta**
 - Sunt aplicate cand conditiile de aplicare a alor metode nu sunt satisfacute.
 - Probabilitatiile obtinute aplicand majoritatea testelor neparametrice sunt probabilitati exacte
 - Pentru esantioane de dimensiuni foarte reduse (pana la $n = 6$)
-



UMFT

Universitatea de
Medicină și Farmacie
„Victor Babeș”
din Timișoara

Teste statistice - natura esantioanelor

	Ce?	Exemple
Pentru esantioane independente	Observatiile sunt independente in interiorul grupelor si de la un grup la altul	TA la barbati si la femei
Pentru esantioane pereches	Masuri repetate pe aceeasi indivizi	TA inainte si dupa administrarea unui medicament/tratament



UMFT

Universitatea de
Medicină și Farmacie
„Victor Babeș”
din Timișoara

Teste statistice

Stabilirea unui nivel de semnificație (α)

α este probabilitatea de a rejecta incorect ipoteza initiala H_0 cand ea este adevarata

Valori uzuale:

- $\alpha = 0,05$ (5% probabilitate de eroare)
- $\alpha = 0,01$ (1% probabilitate de eroare)
- $\alpha = 0,001$ (0.1 % probabilitate de eroare)

Alegerea testului potrivit pentru a testa ipoteza initiala (H_0)

daca esantionul este:

De mari dimensiuni ($n \geq 30$ (100)) se utilizeaza testul Z

De mici dimensiuni ($n < 30$) se utilizeaza testul statistic t (test t-Student) cu $n-1$ grade de libertate



UMFT

Universitatea de
Medicină și Farmacie
„Victor Babeș”
din Timișoara

Interpretarea rezultatului

- **Luand in considerare *p-value***
 - *p-value* este probabilitatea de a obtine, cand H_0 este adevarata, un rezultat pentru statistica testului mai extrem decat rezultatul observat
 - Daca $p\text{-value} < \alpha \Rightarrow$ se accepta H_a si se rejecteaza H_0
 - Daca $p\text{-value} \geq \alpha \Rightarrow$ se accepta H_0
- **$p = 0.42390353 > \alpha (=0,05) \Rightarrow$ se accepta H_0 : nu exista diferente semnificative intre valorile de colesterol inainte si dupa tratament**



Teste statistice uzuale (2 valori de medii)

- 2 serii independente
 - $H_0 : m_1 = m_2$
 - condition: $s_1 = s_2$
 - **Testul t-Student pentru esantioane independente**
(test de semnificatie, parametric)
- 2 serii pare
 - $H_0 : m_1 = m_2$
 - **Testul t-Student pentru esantioane pereche**
(test de semnificatie, parametric)



UMFT

Universitatea de
Medicină și Farmacie
„Victor Babeș”
din Timișoara

Testul t-Student

(2 esantioane independente)

- Conditii de aplicare:
 - variantele esantioanelor sunt cunoscute
 - $n_1 \neq n_2$
- Ipoteza initiala (H_0) – Nu exista diferente semnificative intre cele doua medii (la nivel de populatie)
- Ipoteza alternativa (H_1) – Exista diferente semnificative intre cele doua medii (la nivel de populatie)



UMFT

Universitatea de
Medicină și Farmacie
„Victor Babeș”
din Timișoara

Teste statistice – decizia statistica

- p = probabilitatea ca :
 - H_0 este adevarata
 - Diferentele nu sunt semnificative

7. Formularea deciziei (in functie de p)

- $p \geq \alpha$ - Se accepta H_0
- $p < \alpha$ - Se rejecteaza H_0



UMFT

Universitatea de
Medicină și Farmacie
„Victor Babeș”
din Timișoara

ANOVA (ANALISIS OF VARIANCE) ($n > 2$)

- 2 serii independente

a) One way analysis (unifactoriale)

- $H_0 : m_1 = m_2 = \dots = m_n$
- **Test KRUSKAL – WALLIS** (test neparametric)

b) Two ways analysis (bifactoriale)

- $H_0 : m_1 = m_2 = \dots = m_n$
 - **Test FRIEDMAN** (test neparametric)
-



UMFT

Universitatea de
Medicină și Farmacie
„Victor Babeș”
din Timișoara

Test ANOVA

- Permite verificarea ca mai multe esantioane provin de la aceeași populație (esantioanele sunt diferite sau nu)
 - Condițiile de aplicare:
 - populațiile sunt distribuite normal
 - variantele sunt egale
 - esantioanele și observațiile sunt independente
-



UMFT

Universitatea de
Medicină și Farmacie
„Victor Babeș”
din Timișoara

Testul χ^2 de independență

Utilitate

- Pentru a testa asocierea / independența între variabile calitative
 - variabile calitative cu două sau mai multe valori (categorii)
 - incompatibile și independente

Exemples:

- Relația între
 - un factor de risc și apariția unei afecțiuni
 - un tratament în comparație cu un alt tratament
-



UMFT

Universitatea de
Medicină și Farmacie
„Victor Babeș”
din Timișoara

Testul χ^2 de independenta

Conditii de aplicare:

- Observatii independente
- Mai mult de ($>$) 80% dintre frecventele **teoretice** trebuie sa fie mai mari decat 5

Daca cel putin 20% dintre frecventele teoretice sunt inferioare lui 5, testul nu se poate aplica (se aplica testul exact a lui Fisher).

Ipoteza initiala (exemple):

- Actiunea factorului de risc si aparitia afectiunii sont independente
 - Echivalent cu:
- Nu exista o relatie/asociere/legatura intre factorul de risc si aparitia afectiunii
- Tratamentul nu este util pentru trata o anumita afectiune
 - Echivalent cu:
- Tratamentul si afectiune sunt independente: Nu exista o relatie/ asociere /legatura intre tratament si afectiune