

DECIZIA MEDICALA ASISTATA PROGRAMUL *BAYES*

INTRODUCERE

Bayes este un program de instruire asistata – ramane sa descoperiti voi insiva utilitatea si atractivitatea acestui mod de aprofundare a cunostintelor prezentate la curs.

Programul consta din doua parti: o parte de instruire asistata care este utila pentru a repeta concepte deja invatate privind probabilitati conditionate, regula lui Bayes, sensibilitatea si specificitatea unui clasificator etc. si o parte in care cunostintele studentului pot fi testate intr-o maniera interactiva.

Programul a fost dezvoltat initial la *Free University* in Amsterdam si apoi a fost rescris la *University of Limburg*, fiind rodul unei colaborari intre departamentele de informatica medicala de la universitatile din Maastricht, Freiburg si Gent, in cadrul unui program *Erasmus*.

UTILIZAREA PROGRAMULUI *BAYES*

lansarea *Bayes* Se tasteaza *bayese* la aparitia prompt-ului MS-DOS sau se utilizeaza pictograma de pe *Desktop* sub *Windows*.

MENIURILE *BAYES*

linia meniu Meniul principal se gaseste in partea de sus a ecranului, este de tip *pull-down* si are 5 optiuni:

Introduction	Distributions	Cutoff-point	ROC	Quit
Probabilities Bayes' Formula	Specification Modify overlap Draw sample		ROC display Vary overlap	

selectia optiunilor

<RET>	Fiecare dintre optiuni se alege prin pozitionarea corespunzatoare a barei de selectie (cu tastele sageti) si apasarea tastei Enter (indicata ca <RET>) sau sageata ↓.
<ESC>	
tastele functionale	Daca utilizati tastele sageti, "intoarcerea" dintr-o optiune se poate face mai usor; daca nu, puteti parasi un meniu cu tasta Escape .
<^RET>	Linia din partea de jos a ecranului contine descrierea functiilor pe care le indeplinesc unele dintre tastele functionale (F1, F2, F10, etc.), in functie de context.
F1	Combinatia <CTRL>+<Enter> se utilizeaza pentru confirmarea faptului ca toate datele numerice introduse sunt corecte si se poate trece la etapa urmatoare (= <i>Done</i>).
<ALT>+F1	Este tasta de <i>Help</i> - de ajutor suplimentar. Daca utilizati aceasta facilitate din meniul principal, aveti posibilitatea sa faceti "recapitularea" unor notiuni privind probabilitati conditionate, regula lui Bayes, etc.
F2	Combinatia <ALT>+F1 permite intoarcerea la etapa precedenta, in cadrul "navigarii" prin <i>Help</i> . Va ofera accesul la un mic "calculator de buzunar" - pentru calcule puteti utiliza tastele obisnuite: <ul style="list-style-type: none">• tastele numerice pentru introducerea valorilor numerice• tastele cu operatori aritmetici• M/R pentru introducere/rechemare date in/din memorie• C pentru resetare (<i>Clear</i>)• E pentru stergerea ultimei valori introduse (<i>Erase</i>)

Space	Tasta spatiu se utilizeaza pentru afisarea raspunsului corect (fara a se mai astepta raspunsul utilizatorului) - va indicam sa nu apelati la aceasta facilitate decat in caz de neaparata nevoie.
F10	Este tasta de stergere.
	MODULUL INTRODUCTION
	Reimprospatarea cunostintelor privind probabilitatile si formula lui Bayes
Probabilities	Probabilities va reaminteste calculul probabilitatilor intr-o tabela 2×2.
Bayes' Formula	Submeniul Bayes' Formula va reaminteste modul in care se calculeaza probabilitatile "posteroare" sau valorile prezise (sau calculate) - <i>predictive values</i> .
Probabilitati	Selectati aceasta optiune si incercati sa va reamintiti semnificatia notatiilor:
#	<ul style="list-style-type: none"> • numar
P(A)	<ul style="list-style-type: none"> • probabilitatea evenimentului A <p>P(D+) probabilitatea ca boala D sa fie prezenta</p>
P(A,B)	<ul style="list-style-type: none"> • probabilitatea evenimentelor A si B <p>P(D+,T-) probabilitatea combinatiei: boala prezenta si rezultat negativ al testului T</p>
P(A B)	<ul style="list-style-type: none"> • probabilitatea evenimentului A atunci cand este dat B <p>P(D+ T+) probabilitatea bolii atunci cand testul T este pozitiv</p>
	Raspundeti intrebarilor care se refera la tabelul prezentat.
Calc. prob. posteroare	Selectand optiunea <i>Bayes' Formula</i> se deschide o fereastra in care vi se cere sa precizati valorile pentru trei parametri:

- sensitivity*
 - specificity*
 - prevalence*
- sensibilitatea testului la care va referiti
 - specificitatea testului
 - prevalenta unei anumite boli = probabilitatea bolii respective in randul populatiei

Puteti pastra valorile implicite (*Enter*) sau puteti introduce alte valori ori de cate ori doriti acest lucru.

Valorile prezise pentru un rezultat pozitiv, respectiv negativ vor fi:

- predictive value of a positive result*
 - predictive value of a negative result*
- $P(D+|T+)=\dots\dots\dots$
 - $P(D-|T-)=\dots\dots\dots$

MODULUL DISTRIBUTIONS

Specificarea distributiilor si generarea esantioanelor de valori

In cele prezentate anterior, am considerat raspunsul unui test ca fiind pozitiv sau negativ. Teste cu acest tip de raspuns sunt utilizate frecvent, dar deseori iesirea prima a unui test este un numar (valoarea unei variabile cu un domeniu continuu de valori) si trebuie sa stabilim o valoare de prag = “granita” intre pozitiv si negativ. Alegerea acesteia influenteaza calitatea deciziei si, de aceea, este foarte importanta.

Considerand cazul ipotetic in care s-ar aplica un test pentru toti membrii unei populatii, s-ar putea asocia o distributie pentru multimea valorilor posibile ale variabilei-test. Aceasta este situatia pe care incearca sa o reprezinte optiunea *Distributions* din program.

Un anumit test va fi cu atat mai bun (aplicarea lui cu atat mai relevanta) cu cat distributiile variabilei-test pentru cele doua situatii ce trebuiesc discriminate (persoane bolnave, respectiv sanatoase) se vor suprapune mai putin - in cazul **ideal** ele sunt disjuncte.

	In practica, exista intotdeauna o zona a intervalului de valori pentru variabila-test in care cele doua distributii se suprapun.
Specification	Puteti specifica distributii normale (<i>Gaussian distributions</i>) sau distributii arbitrare, definite de utilizator (<i>Free distributions</i>).
Gaussian distributions	Datele care se cer introduse sunt urmatoarele:
<i>boundaries</i>	<ul style="list-style-type: none"> • marginea inferioara/superioara pentru valori - intervalul definit de acestea va reprezenta $media \pm 2 * deviatia_standard$ (adica intervalul in care se vor gasi 95% dintre valori) <i>Daca introducem 100 pentru limita inferioara si 200 pentru cea superioara, distributia astfel specificata va avea media 150 si deviatia standard 25.</i>
<i>prevalence</i>	<ul style="list-style-type: none"> • prevalenta - va determina aria relativa (si, de aici, inaltimea relativa) a celor doua distributii (bolnavi si sanatosi) O valoare mica pentru prevalenta va conduce la un numar mic de persoane bolnave si aria corespunzatoare a distributiei acestora va fi mai mica.
Modify overlap	Aceasta optiune va permite sa stabiliti intervalul in care cele 2 distributii se suprapun.
Draw sample	Aceasta optiune va permite generarea unor esantioane de valori ce corespund distributiilor definite in prealabil. Situatiia este analoaga celei in care, in practica medicala, culegeti date pentru a analiza performantele unui test.
<i>sample size</i>	Acest parametru stabileste marimea esantionului (numarul de valori) pe care doriti sa-l generati (valoarea max. 10000).

MODULUL CUTOFF-POINT

Stabilirea pragului de decizie pentru un test

Cutoff-point	<p>Pragul care stabileste limita de la care un test va fi considerat pozitiv este valoarea care va influenta in mod decisiv performantele unui test si "calitatea" deciziilor bazate pe acesta.</p> <p>In practica, un test pozitiv conduce la urmatoarea actiune: se presupune ca persoana in cauza este bolnava si se administreaza un anumit tratament. Trebuie insa sa realizam ca in situatia in care probabilitatea posterioara $P(D+ T+)$ este mica testul nu este prea bun. De aceea, un alt factor care trebuie neaparat luat in considerare la alegerea unui anumit test este <i>positive predictive value</i>.</p>
Organizarea ecranului	<p>Pe ecran apar reprezentate grafic cele doua distributii definite anterior si valorile efective existente in esantionul de lucru.</p> <p>Deasupra axei orizontale apare distributia persoanelor sanatoase (D+), iar sub axa a celor bolnave (D-).</p>
tastele ← →	<p>Pragul de separatie intre categoriile test pozitiv (T+) si test negativ (T-) apare ca o linie verticala ce poate fi mutata stanga-dreapta. Valoarea de prag apare afisata pe ecran.</p> <p>Notatiile utilizate au urmatoarea semnificatie:</p> <ul style="list-style-type: none">TP • adevarat-pozitiviTN • adevarat-negativiFP • fals-pozitiviFN • fals-negativi
<RET>	<p>Observati ca tabela 2×2 din partea de jos a ecranului "urmareste" miscarea pragului de separatie: atunci cand o noua valoare este selectata prin apasarea tastei Enter.</p>

- F3 Permite modificarea prevalentei.
- F9 Permite ca pe coloana din dreapta sa apara afisate fie formula lui Bayes, fie versiunea precedenta a tabelii 2×2 de calcul.
- F4 Permite extragerea unui nou esantion de valori pe baza acelorasi distributii definite anterior.

MODULUL *ROC*

Curba ROC **Curba Receiver-Operating Characteristic si functia de cost**

Curba ROC sintetizeaza specificitatea si sensibilitatea (cei 2 indicatori de baza ai performantelor unui test) intr-un singur grafic, ce reprezinta potentialul unui test pentru toate valorile posibile ale pragului de decizie (*cutoff-point*).

Curbele ROC sunt utilizate pentru a compara performantele diverselor teste si pentru a determina pragul optim pentru un anumit test.

ROC display La alegerea acestei optiuni vi se cere sa introduceti costurile pentru deciziile gresite:

- Costuri
- fals negativi
 - fals pozitivi

Pentru simplificare, in acest program s-a ales situatia in care doar clasificarile incorecte ale testului au asociate costuri; intr-o descriere mai generala, deciziile corecte ar avea asociate costuri negative. Putem deci considera ca deciziile corecte au asociate costuri zero.

In general, se prefera ca un test sa fie aplicat pentru acea valoare de prag pentru care costurile sunt minime.

	Costurile nu se refera doar la aspectele de ordin material sau financiar - ele implica si discomfortul persoanei afectate de o decizie incorecta. De aceea, se pot indica valori diferite pentru cele doua tipuri de decizii incorecte.
Organizarea ecranului	In partea dreapta jos a ecranului apare reprezentarea grafica a costurilor functie de valoarea de prag. Cele doua curbe folosesc urmatoarea conventie: <ul style="list-style-type: none">• linie continua pentru populatie• linie punctata pentru esantionul curent de valori
Taste functionale	Rolul tastelor functionale (F1, F2, F3, F4) ca si al celorlalte taste indicate (<RET>, <ESC>) este similar cu situatia de la modulul anterior.
Quit	PARASIREA PROGRAMULUI <i>BAYES</i>

**IN ACEASTA SEDINTA ATI INVATAT DESPRE
*DECIZIA MEDICALA:***

- probabilitati conditionate
- regula lui Bayes
- specificitate/sensibilitate ca evaluatori ai calitatii unui clasificator
- costuri in decizia medicala
- alegerea pragului de decizie (*cutoff-point*) si influenta lui asupra calitatii unui clasificator