MOISE ALINA ELENA

INFERENȚA STATISTICĂ BAYESIANĂ CU APLICAȚII

ISBN 978-973-0-31158-7

PLOIEȘTI

2020

INFERENȚA STATISTICĂ BAYESIANĂ CU APLICAȚII/ **ALINA ELENA MOISE-2020, Ploiești**

Lucrare publicată online http://cevmpl.ro/

ISBN 978-973-0-31158-7

CUPRINS

Introducere

In aceasta lucrare sunt abordate cateva aspecte ale Inferentei Bayesiene ,pentru problemele de estimare a parametrilor si de testare a unor ipoteze statistice. In Inferenta Bayesiana se combina informatia data de probabilitatile aprorice cu cea obtinuta printr-o selectie realizata, iar informatia finala este exprimata prin probabilitati aposteriorice.

In primul capitol sunt prezentate notiuni teoretice ale analizei bayesiene, precum si notiuni din teoria deciziilor care intervin in inferenta bayesiana.

Ultimul capitol este unul aplicativ. Prima parte a capitolului cuprinde modelarea Bayesiana a raspunsurilor unui test grila. In cea de-a doua parte este descris un process de repartizare a elevilor de liceu pe specializari cu ajutorul functiilor de decizie Bayes.

1. ANALIZA BAYESIANA

1.1. INTRODUCERE

Analiza bayesiana are la baza principiul bayes si presupune cunoasterea functiei pierdere U si a distributiei apriori π. Evaluarea unei decizii se face pe baza riscului $r(\pi,\delta) = E^{\pi}[R(\theta,\delta)]$.Cu toate criticile legate de existent distributiei apriori,analiza bayes furnizeaza instrumentul rational de alegere in conditii de incertitudine.Justificarea acestui fapt rezulta din urmatoarele considerente: daca D este multimea functiilor de decizie pe care avem o ordine preferentiala (data de principiul de decizie utilizat , de exemplu Bayes) , exista o functie utilitate u: D → R care se acorda cu ordinea preferintiala data. In termeni de functie risc , pentru δ ∈D definim $r(\delta) = -u(\delta)$.(Avem in vedere ca si functia pierdere U se obtine in acelasi mod dintr- o functie utilitate).

Acum sa consideram o problema de decizie in care alegem δ ∈D cand θ este presupus cunoscut. In aceasta situatie, ordinea preferentiata pe D ,stiind pe θ ,ne conduce ca mai sus ,la o functie risc $R_{\theta}(\delta)$.

Relatia intre $r(\delta)$ si $R_{\theta}(\delta)$ este data de teorema fundamentala a analizei bayesiene care,in anumite conditii de regularitate, afirma ca $r(\delta)$ este o functionala liniara pozitiva de $R_{\theta}(\delta)$ iar din torema de prezentare a lui Riesz avem ca exista o masura π pe (Θ,F) pentru care avem $r(\delta) = E^{\pi}[R_{\theta}(\delta)]$ pentru orice δ ∈D .

Sa observam ca atunci cind U este obtinut dintr- o functie utilitate, cantitatea $R_{\theta}(\delta)$ trebuie sa fie interpretata ca o pierdere medie a lui δ in ipoteza ca starea este θ, adica $R_{\theta}(\delta)$ este tocmai $R(\theta,\delta) = E_{\theta}[U(\theta,\delta(X))]$. Prin urmare $r(\delta) = E^{\pi}[R(\theta,\cdot)]$ iar π trebuie sa fie o masura finita. (Practic, cind π disponibil nu este o masura finita, se poate considera un spatiu trunchiat al parametrilor).

Prin urmare ,orice metoda rationala de alegere a deciziilor corespunde la o ordonare indusa de $r(\pi, \delta)$ in raport cu o anumita distributie apriori π.

1.2. FORMA EXTENSIVA A ANALIZEI BAYESIENE

Pentru un π dat ,in analiza bayesiana, se cere un $\delta \in$D care minimizeaza $r(\pi, \delta)$. Aceasta forma a analizei bayesiene poarta numele de forma normala. Cum determinarea lui δ ca mai sus , nu este o chestiune simpla pentru multe probleme de decizie, s-a cautat o alta directie.

Daca ,in cele ce urmeaza ,presupunem ca putem schimba ordinea de integrare iar $\delta \in$D ,atunci,

$$r(\pi, \delta) = E^{\pi}[R(\theta, \delta)] \int_{\Theta} R(\theta, \delta) \cdot \pi(\theta) d\nu$$

$$= \int_{\Theta} E_{\theta}[U(\theta, \delta(X)] \cdot \pi(\theta) d\nu = \int_{\Theta} \int_{\mathfrak{X}} U\big(\theta, \delta(x)\big) f(x|\theta) \cdot \pi(\theta) d\mu \, d\nu$$

$$= \int_{\mathfrak{X}} [\int_{\Theta} U\big(\theta, \delta(x)\big) f(x|\theta) \cdot \pi(\theta) \, d\nu] d\mu$$

$$= \int_{\mathfrak{X}} m(x) [\int_{\Theta} U\big(\theta, \delta(x)\big) f(x|\theta) \cdot \pi(\theta) \, d\nu] d\mu,$$

unde μ este o masura finita pe $\mathfrak{X}$ in raport cu care $\{P_{\theta}\}_{\theta}$ este absolut continua iar ν o masura finita pe Θ in raport cu care $\pi(\theta)$ este o densitate a unei distributii apriori.

Acum a alege δ care minimizeaza $r(\pi, \delta)$ revine la a alege, pentru orice $x \in \mathfrak{X}$ un element $\delta(x) \in A$ care minimizeaza cantitatea

$$\int_{\Theta} U\big(\theta, \delta(x)\big) \cdot \pi\,(\theta|x) \, d\nu$$

adica minimizeaza pierderea medie aposteriori (pierderea medie in raport cu densitatea aposteriori $\pi(x|\theta)$ a lui θ cunoscind x) .

In concluzie ,o decizie Bayes δ in raport cu π se poate determina luind, pentru orice $x \in \mathfrak{X}$,actiunea $\delta(x) \in A$ care minimizeaza pierderea medie aposteriori.

Forma analizei bayesiene in care se determina deciziile Bayes in acest mod poarta numele de forma extensiva a analizei bayesiene.

Deoarece in multe situatii, $m(x)$ este greu de gasit ,determinarea unei decizii Bayes δ in raport cu π revine la alegerea, pentru orice $x \in \mathfrak{X}$ a unei actiuni $\delta(x) \in A$ care minimizeaza expresia:

$$\int_{\Theta} U(\theta, \delta(x)) f(x|\theta) \cdot \pi(\theta) dv.$$

Este interesant sa gindim o decizie Bayes in modul urmator : Daca alegem o actiune fara a considera procesul de experimentare,atunci optima va fi acea actiune care Bayes in raport cu π dat. Cind procesul de experimentare este utilixat ,indata ce X a fost observata, alegerea unei decizii optime este in principal de aceasi natura cu cazul in care nu avem proces de experimentare.Exista o diferenta numai prin faptul ca distributia lui θ se schimba dintr-o distributie apriori intr-o distributie aposteriori.Prin urmare ,daca $X = x_0$, optima este actiunea $\theta(x_0)$ care este Bayes relativ la $\pi(\theta|x_0)$.

Din aceasta discutie concluzionam ca actiunea $\delta(x_0)$ care este specificata de o decizie Bayes δ cind $X = x_0$ este observata, se poate determina ,fara a calcula efectiv decizia δ ,si pentru alte valori $x \in \mathfrak{X}$ .In plus ,se poate determina o decizie Bayes δ in raport cu π fara a calcula riscul Bayes $r(\pi, \delta)$.

Exemplul 1 :

Fie $\Theta = \{\theta_1, \theta_2\}, A = \{a_1, a_2\}, U(\theta_1, a_1) = U(\theta_2, a_2) = 0$, $U(\theta_2, a_1) = 2\, U(\theta_1, a_2) = 10$, $\mathfrak{X} = \{0,1\}$, $f(0|\theta_1) = 1 - f(1|\theta_1) = \frac{1}{4}$, $f(0|\theta_2) = 1 - f(1|\theta_2) = \frac{2}{3}$

Daca $\pi = (p, 1 - p)$ cu $0 \leq p \leq 1$ este o distributie apriori, atunci $\pi(\theta|x)$ este dat de:

$$\pi(\theta_1|0) = 1 - \pi(\theta_2|0) = 3p/(8 - 5p)$$

$$\pi(\theta_1|1) = 1 - \pi(\theta_2|1) = 9p/(5p + 4)$$

Indata ce valoarea x a variabilei aleatoare X este observata,se pune problema alegerii ce trebuie facuta intre cele doua actiuni a_1 *si* a_2.

Se verifica usor ca pentru $x = 1$ expresia din 1 este minima cand: $\delta(1) = a_2$ daca $p < 8/17$; $\delta(1) = a_1$ daca $p > 8/17$;$\delta(1) = a_1$ sau $\delta(1) = a_2$ daca $p = 8/17$.

Pentru $x = 0$,expresia din 1 este minima cand: $\delta(0) = a_2$ daca $p < 16/19$; $\delta(0) = a_1$ daca $p > 16/19$;$\delta(0) = a_1$ sau $\delta(0) = a_2$ daca $p = 16/19$.

In concluzie δ este o decizie Bayes relativ la π daca:

$\delta(x) = a_2$ pentru $0 \le p \le 8/17$; $\delta(0) = a_2$ si $\delta(1) = a_1$ pentru $8/17 < p < 16/19$; $\delta(x) = a_1$ pentru $16/19 \le p \le 1$.

Corespunzator acestor decizii Bayes riscul Bayes este:

$$r(\pi,\delta) = \begin{cases} 5p, & 0 < p < 8/17 \\ \dfrac{40-25p}{12}, & 8/17 < p < 16/19 \\ 10(1-p), & 16/19 \le p \le 1 \end{cases}$$

Acest risc Bayes este functie concave.De astfel, in general daca $\mathcal{H}$ este o multime convexa de distributii apriori, atunci $r^*(\pi) = \inf_\delta r(\pi,\delta)$ este o functie concava.

Forma extensive descrisa este mai simpla decat forma normal a analizei bayesiene si va fi folosita pentru a determina deciziile Bayes.Multi decidenti recunosc numai forma extensive a analizei pentru a determina deciziile Bayes deoarece $\pi(\theta|x)$ descrie credinta post experimentala (posibil subiectiva) cu privire la starea adevarata θ iar pierdereamedie in raport cu $\pi(\theta|x)$ .Aici gradul preciziei finale pentru o actiune $a \in A$ este tocmai pierderea medie aposteriori iar $r(\pi,\delta)$ este o masura a preciziei initiale in raport cu a .

Aceasta forma extensive permite studiul problemelor de decizie in care $r(\pi,\delta)$ este infinit pentru orice $\delta \in$D ,deoarece pierderea medie aposteriori este intotdeauna finite.Pentru asemenea probleme de decizie forma normal a analizei bayesiene nu are nici-o semnificatie.Cand $r(\pi,\delta)$ este infinit o decizie Bayes definita in cadrul analizei extensive poarta numelede decizie Bayes formala.

Se observa ca o decizie Bayes in cadrul analizei extensive verifica si principiul verosimilitatii deoarece espresia

$$\int_{\Theta} U(\theta, \delta(x)) f(x|\theta) \cdot \pi(\theta) d\nu$$

contine informatia de selectie numai prin intermediul functiei de verosimilitate $f(x|\theta)$

1.3. INFERENTA BAYESIANA

Problemele de decizie privind inferenta cu privire la θ pot fi tratate utilizand analiza bayesiana. Acest lucru se datoreaza faptului ca distributia aposteriori contine toata informatia cu privire la θ atat informatia de selectie cat si informatia apriori, iar orice inferenta cu privire la θ se face utilizand aceasta distributie.

In statistica matematica clasica in cadrul abordarii bayesiene, au fost definite mai multe tehnici cu privire la estimarea parametrului θ, tehnici ce se aplica la distributia aposteriori. Cateva din aceste tehnici vor fi ilustrate prin exemple in cele ce urmeaza.

Definitia 1. *Estimatia de maxima verosimilitate generalizata a lui θ este valoarea θ care maximizeaza distributia aposteriori $\pi(\theta|x)$ considerate ca functie de θ.*

De aici observam ca aceasta tehnica reprezinta analogul bayesian al tehnicii clasice de maxima verosimilitate si care consta in a alege drept estimatie pentru θ pe $\hat{\theta}$ care maximizeaza functia de verosimilitate $f(x|\theta)$ .Valoarea $\hat{\theta}$ din definitie, de fapt, corespunde la modul distributiei aposteriori $\pi(\theta|x)$, la cea mai posibila valoare a lui θ dand informatia apriori si selectia x .Totusi, experienta arata ca in cele mai multe cazuri media si mediana aposteriori conduc la estimatii bayesiene mai bune.

Exemplul 1 **:**

Fie $X \sim N(\theta,\sigma^2)$ cu $\theta \in \Theta \subseteq \mathbb{R}^1$ necunoscut iar $\pi(\theta)$ destitatea normal de tip $N(\mu,\tau^2)$ cu μ si τ^2cunoscute. $r(\theta|x)$ este o densitate de tip $N(\mu(x),\rho^{-1})$ unde $\mu(x) = \sigma^2 \cdot \frac{\mu}{\sigma^2+\tau^2} + \tau^2 \cdot \frac{x}{\sigma^2+\tau^2}$. Dar $\pi(\theta|x)$ isi atinge maximul in media repartitiei normale $N(\mu(x),\rho^{-1})$ .Deci $\theta = \mu(x)$ este estimatia de maxima verosimilitate generalizata a lui θ si coincide cu media si mediana lui $\pi(\theta|x)$.

Exemplul 2:

Fie $\Theta = (-\infty, +\infty)$, $\pi(\theta) = [\pi \cdot (1 + \theta^2)]^{-1}$, $f(x|\theta) = e^{-(x-\theta)}$ pentru $x > \theta$ si zero in rest. Atunci deducem usor ca $\pi(\theta|x) = e^{-(x-\theta)} \cdot \left[m(x) \cdot \pi \cdot \left(1 + \theta^2\right)\right]^{-1}$ pentru $x > 0$ si zero in rest. Deoarece

$$\frac{d\pi(\theta|x)}{d\theta} = e^{-(x-\theta)} \cdot (\theta - 1)^2 \left[m(x) \cdot \pi \cdot \left(1 + \theta^2\right)^2\right]^{-1}$$

pentru orice $\theta \leq x$ deducem ca $\pi(\theta|x)$ este strict crescatoare si prin urmare θ care maximizeaza $\pi(\theta|x)$ este dat de $\hat{\theta} = x$ care difera de media si mediana distributiei $\pi(\theta|x)$.

In statistica clasica un alt mod de a estima pe θ este legat de notiunea de regiune de incredere. Analogul Bayesian al acestui concept este regiunea credibila.

Definitia 2. *Fie* $0 < \alpha < 1$ *.Multimea* $C \subseteq \Theta$ *este o* $(1-\alpha)-$ *regiune credibila pentru* θ *daca* $P(C|x) \geq 1 - \alpha$*, unde* $P(C|x) = \int_C dF^{\pi(\theta|x)}$.

Daca in statistica clasica regiunile de incredere erau interpretate in sensul probabilitatii de acoperire, probabilitatea ca selectia sa fie astfel incat regiunea de incredere obtinuta sa contina pe θ, si care reprezinta o masura a preciziei initiale, in cadrul inferentei bayesiene probabilitatea de acoperire bayesiana este o masura a preciziei finale deoarece dupa ce se observa x, probabilitatea ca θ sa se afle in C este de cel putin $1 - \alpha$.

In multimea regiunilor credibile un rol important il au regiunile credibile de densitate aposteriori mare. Asemenea regiuni contin cele mai probabile valori ale lui θ. O $1-\alpha-$ regiune credibila de cea mai mare densitate aposteriori este o multime de foarma $C=\{\theta \in \Theta \,|\pi(\theta|x) \geq k(\alpha)\}$, unde $k(\alpha)$ este cea mai mare constanta reala pentru care $P(C|x) \geq 1-\alpha$.

***Exemplul* 1.** (continuare):

Deoarece $\pi(\theta|x)$ este o densitate $N(\mu(x),\rho^{-1})$ care este unimodala si simetrica in raport cu $\mu(x)$, deducem ca o $1-\alpha-$ regiune credibila de cea mai mare densitate aposteriori este data de interval $C=\left(\mu(x)+z\left(\frac{\alpha}{2}\right)\cdot\sqrt{\rho^{-1}}, \mu(x)-z\left(\frac{\alpha}{2}\right)\cdot\sqrt{\rho^{-1}}\right)$,unde $z(\alpha)$ este $\alpha-$cuantila distributiei $N(0,1)$.

Un neajuns al acestor regiuni credibile de cea mai mare desitate aposteriori consta in faptul ca uneori ele pot fi formate din reuniuni de multimi disjuncte. Luand , de exemplu $f(x|\theta)$ de tip normal iar $\pi(\theta)$de tip Cauchy obtinem $\pi(\theta|x)$ ca fiind o densitate bimodala iar regiunea credibila apare ca o regiune de doua interval disjuncte. In asemenea cazuri este indicat sa utilizam regiuni credibile de probabilitate exact $1-\alpha$.

In finalul acestei sectiuni consideram analogul Bayesian al testarii ipotezelor statistice. Fie ipoteza nula $H_0\colon \theta \in \Theta_0$ si ipoteza alternativa $H_1\colon \theta \in \Theta_1$. Atunci in cadrul clasic testul statistic de a decide intre H_0.siH_1 este definit in functie de erorile de tip I si II. Aceste erori, care sunt niste probabilitati, reprezinta sansa ca selectia ce o obtinem sa ne conduca, prin intermediul testului statistic, la acceptarea unei ipoteze false. Din acest motiv ele reprezinta masuri ale preciziei initiale.

O analiza simpla arata ca in cadrul Bayesian, a decide intre cele doua ipoteze revine la a compara probabilitatile aposteriori : $\alpha_0=P(\Theta_0\,|x)=P^{\pi(\cdot|x)}(\Theta_0\,)$ si $\alpha_1=P(\Theta_1|x)$. Deoarece $\alpha_0$si α_1 sunt probabilitatile ca H_0 respectiv H_1 sa fie adevarate, ele devin masuri ale preciziei finale.

Daca am presupune ipotezele $H_0\colon \theta=\theta_0$, $H_1\colon \theta \neq \theta_0$.atunci pentru cazul cand θ este un parametru continuu, cadrul bayesian conduce la $\alpha_0=0$ si prin urmare H_0 niciodata nu poate fi acceptata. Totusi acest neajuns poate fi remediat printr-o reformulare a problemei. De regula rareori avem, de testat, cu adevarat ca $\theta=\theta_0$ .De aceea ipoteza H_0 se reformuleaza astfel : $H_0\colon \theta \in (\theta_0-\varepsilon,\theta_0+\varepsilon)$, unde ε reflecta acuratetea cu care este cunoscut θ_0 .Un alt mod de a evita neajunsul de mai sus, in cadrul Bayesian, consta in a acorda o probabilitate apriori, $\pi_0 > 0$ pentru θ_0 iar pentru valorile $\theta \neq \theta_0$, probabilitatea $1 - \pi_0$ .Ilustram acest lucru printr-un exemplu.

Exemplul 2:

Fie $X \sim N(\theta, \sigma^2)$ cu σ^2 cunoscut. Scopul este de a testa $H_0: \theta = \theta_0$, in raport cu ipoteza alternativa $H_1: \theta \neq \theta_0$, unde $\theta \in \Theta = (-\infty, +\infty)$. Presupunem ca probabilitatea apriori pentru θ_0 este π_0 iar pentru $\theta \neq \theta_0$ avem desitatea apriori $(1 - \pi_0) \cdot \pi_1(\theta)$, unde $\pi_1(\theta)$ este o densitate de tip $N(\mu, \tau^2)$.

Astfel ca densitatea pentru (X, θ) este

$$h(x, \theta) = \begin{cases} f(x|\theta) \cdot \pi_0 , & \theta = \theta_0 \\ f(x|\theta) \cdot (1 - \pi_0) \cdot \pi_1(\theta) , & \theta \neq \theta_0 \end{cases}$$

Prin urmare :

$$m(x) = f(x|\theta) \cdot \pi_0 + (1 - \pi_0) \cdot \int_{\Theta - \{\theta_0\}} f(x|\theta) \cdot \pi_1 \, d\theta = f(x|\theta_0) \cdot \pi_0 + (1 - \pi_0) \cdot g(x) ,$$

unde $g(x)$ este o densitate de tip $N(\mu, \sigma^2 + \tau^2)$.

Acum densitatea aposteriori este :

$$\pi(\theta|x) = \begin{cases} f(x|\theta_0) \cdot \pi_0 \cdot [f(x|\theta_0) + (1 - \pi_0) g(x)]^{-1}, & \theta = \theta_0 \\ f(x|\theta) \cdot (1 - \pi_0) \cdot \pi_1(\theta)[[f(x|\theta_0)\pi_0 + (1 - \pi_0) g(x)]^{-1} , & \theta \neq \theta_0 \end{cases}$$

Daca notam $\Theta = \{\theta_0\}$, $\Theta_1 = \Theta - \Theta_0$ obtinem usor ca

$$P(\theta_0|x) = \pi(\theta_0|x) , P(\Theta_1|x) = 1 - \pi(\theta_0|x).$$

Deoarece

$$\frac{P(\Theta_0|x)}{P(\Theta_1|x)} = \frac{\pi_0}{1-\pi_0}\sqrt{\frac{(\sigma^2+\tau^2)}{\sigma^2}} \times \exp\left\{-\frac{(x-[\theta_0+(\theta_0-\mu)\sigma^2\cdot\tau^{-2}])^2}{2\cdot\sigma^2(1+\sigma^2\cdot\tau^{-2})} + \frac{(\theta_0-\mu)^2}{2\tau^2}\right\}$$

deducem ca $P(\theta_0|x) > 2^{-1}$ daca si numai daca $P(\Theta_0|x) > P(\Theta_1|x)$ daca si numai daca:

$$\left|\frac{x-\theta_0}{\sigma} + \frac{(\mu-\theta_0)\sigma}{\tau^2}\right| < \sqrt{\frac{(\sigma^2+\tau^2)}{\tau^2}}\left\{\log\left[\left(\frac{\pi_0}{1-\pi_0}\right)^2 \times \frac{(\sigma^2+\tau^2)}{\sigma^2}\right] + \frac{(\theta_0-\mu)^2}{\tau^2}\right\}^{1/2}$$

In acest caz se accepta H_0 iar in caz contrar H_1.

1.4. EXISTENTA FUNCTIILE DE DECIZIE BAYES

Fie modelul de decizie $(\Theta, \ , \Re)$ in care $\Theta = \{\theta_1, \dots, \theta_k\}$ este o multime finita. Definim multimea risc ca fiind multimea

$$Y = \{y = (y_1, y_2, \dots, y_k | y = \Re(\theta_i, \delta), \delta \in \ , 1 \le i \le k\}.$$

Consideram Y ca o submultime a spatiului Euclidian k-dimensional. Daca

$$\pi = (p_1, \dots, p_k), p_i \ge 0, 1 \le i \le k, \sum_{i=1}^{k} p_i = 1$$

este o distributie apriori pe Θ, atunci punctele din Y pentru care $\sum_{i=1}^{k} p_i y_i = \alpha$, cu α numar real dat, pot fi considerate echivalente in raport cu ordonarea data de principiul Bayes relative la distributia π. Orice asemenea hiperplan, $\sum_{i=1}^{k} p_i y_i = \alpha$

este ortogonal cu vectorul ce uneste originea spatiului Euclidian k-dimensional cu punctual dat de componentele lui $\pi, (p_1, \dots, p_k)$ .Pentru a gasi functiile de decizie

Bayes relative la π trebuie sa gasim cele mai mici valori ale lui α pentru care hiperplanul , $\sum_{i=1}^{k} p_i y_i = \alpha$ intersecteaza multimea Y. E clar ca daca Y nu contine puncte frontiera, atunci nu exista functii de decizie Bayes. Daca $\varepsilon > 0$ este fixat, atunci functiile de decizie $\varepsilon -$ Bayes relativ la π corespund la acele puncte din Y care se afla pe si sub hiperplanul, $\sum_{i=1}^{k} p_i y_i = \alpha + \varepsilon$.

Pentru a stabili rezultatul principal din aceasta sectiune avem nevoie de cateva definitii.

Definitia 1. *Fie* $z = (z_1, \dots, z_k) \in \mathbb{R}^k$. *Definim*

$$(\leq z] = \{(y_1, y_2, \dots, y_k) \in \mathbb{R}^k | y_i \leq z_i, 1 \leq i \leq k \}.$$

Definitia 2. *Fie* $C \subseteq \mathbb{R}^k$ *multime convexa.Definim multimea*

$$\underline{\partial} C = z \in \mathbb{R}^k \left| (\leq z] \bigcap \overline{C} = \{z\} \right\},$$

unde $\overline{C}$ *este inchiderea lui C.*

Un punct $z \in \underline{\partial} C$ nu este altceva decat un punct de frontiera inferioara pentru C.

De exemplu, daca $C = \{z = (z_1, z_2) \in \mathbb{R}^2 \mid (z_1)^2 + (z_2)^2 \leq \rho \}$ cu $\rho \leq 0$, atunci $\underline{\partial} C = \{z = (z_1, z_2) \in \mathbb{R}^2 | (z_1)^2 + (z_2)^2 = \rho, z_1 \leq 0, z_2 \leq 0\}$. In cazul multimii $C = \{z = (z_1, z_2) \in \mathbb{R}^2 | z_1 \in [a, b], z_2 \in [c, d]$, unde a < b, c < d, avem $\underline{\partial} C = \{(a, c)\}$.

Definitia 3. *Multimea convexa C din* $\mathbb{R}^k$*este inferior inchisa daca* $\underline{\partial} C \subseteq C$.

Se constata usor ca orice multime convexa inchisa este si o multime inferior inchisa. In general, inversa nu este adevarata.

Acum avem :

Teorema 1. *Fie* $\Theta = \{\theta_1, \dots, \theta_k\}$ *iar Y multime marginita inferior si inchisa inferior. Atunci pentru orice distributie apriori* $\pi = (p_1, \dots, p_k)$ *cu* $p_i > 0$, $1 \leq i \leq k$ *, existat cel putin o functie de decizie Bayes relativ la* π .

Demonstratie. Fie distributia π ca in enunt. Definim multimea

$$V = \left\{ \alpha \in \mathbb{R}^1 \middle| \alpha = \sum_{i=1}^{k} p_i y_i \,, (y_1, y_2, \dots , y_k) \in Y \right\}$$

Deoarece Y este marginita inferior, rezulta usor ca si V este marginita inferior in $\mathbb{R}^1$. Prin urmare exista $\alpha_0 = \inf V$. Deci exista un sir de puncte in Y, $\{y^n\}_n$ astfel ca

$$\lim_{n\to\infty} \sum_{i=1}^{k} p_i {y_i}^n \quad = \alpha_0 \ .$$

De aici, deoarece $p_i > 0,\ 1 \le i \le k$, rezulta ca sirurile din $\mathbb{R}^1$ sunt marginite si superior. Exista deci y^0 punct limita pentru $\{y^n\}_n$ astfel ca $\sum_{i=1}^{k} p_i {y_i}^{o} = \alpha_0$ .Aratam ca $y^0 \in \underline{\partial} V$ .Mai intai observam ca Y este o multime convexa. Pentru $\lambda \in (0,1)$ si y^1 , $y^2 \in Y$ exista $\delta^1, \delta^2 \in \mathcal{D}$ astfel incat $y_i^j = \Re(\theta_i \,, \delta^j), 1 \le i \le k$,$1 \le j \le 2$ ..Definim $\delta_\lambda = \lambda\delta^1 + (1-\lambda)\delta^2$ ca fiind functia de decizie care pentru orice $x \in \mathfrak{X}$ asociaza distributia de probabilitate pe A de forma $\lambda P^1 + (1-\lambda)P^2$ unde $P^j = \delta^j(x)$,$1 \le j \le 2$. Acum se vede usor ca $\Re(\theta, \delta_\lambda) = \lambda\Re(\theta, \delta^1) + (1-\lambda)\Re(\theta, \delta^2)$ si prin urmare

$$y(\lambda) = (\Box(\theta_1, \delta_\lambda), \Re(\theta_2, \delta_\lambda), \dots , \Re(\theta_k, \delta_\lambda)) = \lambda y^1 + (1-\lambda) y^2 \in Y.$$

Acum aratam ca $y^0 \in \underline{\partial} V$ .Deoarece y^0 este un punct limita pentru Y avem $y^0 \in \overline{Y}\, si\, \{\ y^0\} \subseteq (\le \ y^0] \cap \overline{Y}$. Pe de alta parte, $(\le \ y^0] \cap \overline{Y} \subseteq \{\ y^0\}$ deoarece daca $y' \neq y^0, y' \in (\le \ y^0]$ avem : $\sum_{i=1}^{k} p_i y_i' < \alpha_0$. Astfel ca daca $y' \in \overline{Y}$, atunci exista $y \in Y$ pentru care: $\sum_{i=1}^{k} p_i y_i \ < \alpha_0$, ceea ce contrazice faptul ca α_0, este o margine iferioara pentru V. In concluzie avem ca $(\le \ y^0] \cap \overline{Y} = \{\ y^0\}$. Aceasta implica $y^0 \in \underline{\partial} V$.

Cum Y este inferior inchisa obtinem ca $y^0 \in V$ si prin urmare valoarea minima pentru $\sum_{i=1}^{k} p_i \Re(\theta_i, \delta)$ este atinsa de un punct din Y. Prin urmare, orice δ din pentru care $\Box(\theta_i, \delta) = y_i$, $1 \le i \le k$,este o functie de decizie Bayes relativ la π .

Astfel teorema este demonstrata.

In teorema de mai sus ipoteza $p_i > 0,\ 1 \le i \le k$ nu poate fi suprimata deoarece daca $k = 2$, $Y = \{(y_1, y_2) \in \mathbb{R}^2 | y_1 \cdot y_2 \ge 1, y_1 > 0\}$ este convexa, inferior marginita si inferior inchisa iar pentru $\pi = (0,1)$ avem $\sum_{i=1}^{k} p_i y_i = y_1$. Astfel ca

$inf_{y \in Y}(\sum_{i=1}^{k} p_i y_i) = 0$ nu e atins pe Y. Deci nu exista functie de decizie Bayes relativa la π.

Teorema 2.$(H, \mathcal{F}_{\pi_0})$ *este un spatiu compact daca si numai daca exista* $\delta' \in H$*astfel incat* $r(\pi_0, \delta') = inf_\delta\, r(\pi_0, \delta)$

In final sa observam ca Teorema 2 furnizeaza o conditie necesara si suficienta ca o furnctie de decizie sa fie Bayes relativ la o distributie apriori. Deci exista o decizie Bayes relativ la π daca si numai daca spatial $(\ , \mathcal{F}_\pi)$ este un spatiu compact.

1.5. CAZUL ESTIMARII PARAMETRILOR

In general problema de estimare este o problema de decizie statistica, in care decizia luata de statistician este estimatia sa cu privire la parametrul θ ale carui valori apartin la o submultime Θ a spatiului Euclidian $\mathbb{R}^q$ $(q \geq 1)$ .Deoarece trebuie estimate valoarea lui θ , spatiul actiunilor A de regula concide cu Θ .Acum pierderea $U(\theta, a)$reprezinta discrepanta intre θ si estimatia sa a. Din acest motiv, deseori, pentru aceste problem se presupune ca functia pierdere are forma :$U(\theta, a) = w(\theta)u(\theta - a)$, unde $u(\theta - a) \leq 0$ pentru $\theta \in \Theta, a \in A$ si $u(0) = 0$, iar w este o functie nenegativa care indica abaterea relative a elementului eroare $\theta - a$ pentru diferitele valori ale lui θ. Totusi pentru multe aplicatii, fara a pierde in generalitate, se poate presupune ca $w$este constanta pe Θ.

Cand θ este un parametru unidimensional, deseori se presupune $U(\theta, a) = \rho \cdot |\theta - a|^\alpha$ unde $\rho > 0$ si $\alpha > 0$. Pentru $\alpha = 2$ obtinem pierderea patratica care este frecvent utilizata datorita elegantei aparatului matematic care se aplica. In acest caz o decizie nealeatoare poate fi considerata ca o estimatie a starii adevarate, dar necunoscuta, θ .

Acum decizia Bayes, numita estimatie Bayes, este data de d care minimizeaza $R(\theta, d) = \rho \cdot E_\theta\big(\theta - d(X)\big)^2$, adica minimizeaza eroarea medie patratica a lui $d(X)$.

Teorema 1: *Fie* $U(\theta, a) = \rho(\theta - a)^2$. *Atunci media aposteriori pentru* $\pi(\theta|x)$ *este o estimatie pentru* θ.

Estimatia Bayes se obtine minimizand pierderea medie aposteriori cand $X = x$, adica minimizand $E^{\pi(\theta|x)}(\theta - a)^2$. Egaland cu zero derivata in raport cu "a" a acestei expresii (presupunand ca sunt finite toate integralele care apar) obtinem $-2E^{\pi(\theta|x)}[\theta] + 2a = 0$, adica $a = E^{\pi(\theta|x)}[\theta]$. Prin urmare estimatia Bayes este data de media distributiei aposteriori a lui θ dand $X = x$. Aici am considerat ca π este o distributie apriori pe Θ .

Observatie: Folosind acelasi procedeu, in cazul pierderii patratice ponderate, $U(\theta, a) = w(\theta)(\theta - a)^2$, deducem ca estimatia Bayes este de forma :

$$d(x) = E^{\pi(\theta|x)}[\theta\, w(\theta)] \cdot (E^{\pi(\theta|x)}[w(\theta)])^{-1}.$$

***Exemplul* 1**:

Fie $\Theta = \mathrm{A} = (0, +\infty), U(\theta, a) = \rho \cdot (\theta - a)^2 , \rho > 0.$

Presupunem ca statisticianul observa valoarea unei variabile aleatoare X avand o distributie uniforma pe $(0, \theta)$ cu densitatea $f(x|\theta) = \theta^{-1}$ pentru $x \in (0, \theta)$ si zero in rest. Daca $\pi(\theta) = \theta \exp(-\theta)$, $\theta > 0$, atunci $\pi(\theta|x) = \exp(x - \theta)$ daca $\theta > x$ si zero in rest. Prin urmare estimatia Bayes este data de

$$d(x) = E^{\pi(\theta|x)}[\theta] = \int_x^{+\infty} \exp(x - \theta)\, d\theta = x + 1$$

Deci daca se observa $X = x$, atunci estimatia Bayes a lui θ va fi $x + 1$.

Exemplul 2:

Fie $X = (X_1, \dots X_n)\, cu\, X_i \sim \mathcal{P}(\theta)$.Presupunem $\pi(\theta)$de tip gama de parametrii α, β .Avem $\pi(\theta|x)$de tip gama de parametrii $\alpha + \sum_{i=1}^{n} x_i$, $\beta + n$, unde x_i este dat de $X_i = x_i$Acum deducem usor ca pentru $U(\theta, a) = \rho \cdot (\theta - a)^2$, estimatia Bayes a lui θ este

$$d(x) = \left(\alpha + \sum_{i=1}^{n} x_i\right) \cdot (\beta + n)^{-1}$$

Astfel ca riscul Bayes este dat de $r(\pi) = \alpha \cdot [\beta(\beta - n)]^{-1}$.

Daca consideram $U(\theta, a) = \theta^{-1} \cdot (\theta - a)^2$, atunci obtinem usor estimatia Bayes

$$d(x) = \left(\alpha - 1 + \sum_{i=1}^{n} x_i\right) \cdot (\beta + n)^{-1}$$

pentru orice valori ale lui $X = (X_1, \dots X_n)$ astfel ca $\alpha + \sum_{i=1}^{n} x_i > 0$ .Riscul Bayes corespunzator devine egal cu $(\beta + n)^{-1}$.

Urmatorul rezultat stabileste forma estimatiei Bayes pentru o alta functie pierdere cunoscuta, anume $U(\theta, a) = \rho \cdot |\theta - a|$ cu $\rho > 0$.

Teorema 2. *Fie* $U(\theta, a) = \rho \cdot |\theta - a|$. *Atunci orice mediana pentru* $\pi(\theta|x)$ *este o estimatie pentru* θ.

Demonstratie. Fie m o mediana pentru $\pi(\theta|x)$, iar $a > m$ o alta actiune din A. Observam ca

$$U(\theta, m) - U(\theta, a) = \begin{cases} m - a, & daca\ \theta \leq m \\ 2\theta - (m + a), & daca\ m < \theta < a \\ a - m\ , & daca\ \theta \geq a \end{cases}$$

deduce usor

$$U(\theta, m) - U(\theta, a) \leq \begin{cases} m - a\ , & daca\ \theta \leq m \\ a - m\ , & daca\ \theta > m \end{cases}$$

Deoarece m este mediana pentru $\pi(\theta|x)$ avem ca : $P(\theta \leq m|x) \geq 2^{-1}$, $P(\theta > m|x) \leq 2^{-1}$, si prin urmare deducem

$$\begin{aligned} E^{\pi(\theta|x)}[U(\theta, m) - U(\theta, a)] &\leq (m - a)P(\theta \leq m|x) + (a - m)P(\theta > m|x) \\ &\leq (m - a)2^{-1} + (a - m)2^{-1} = 0 \end{aligned}$$

ceea ce stabileste ca m da o pierdere medie aposteriori mai mica sau egala decat cea data de a . Urmand acelasi procedeu deducem un rezultat similar si pentru cazul $a < m$. Astfel teorema este demonstrata.

Observatie. De altfel se observa usor ca d este o estimatie Bayes in cazul $U(\theta, a) = \rho \cdot |\theta - a|$ daca si numai daca d este o mediana pentru $\pi(\theta|x)$. Implicit presupunem ca $E^{\pi(\theta|x)}[|\theta|] < \infty$.

***Exemplul* 3:**

Fie $X = (X_1, \dots X_n)\ cu\ X_i \sim N(\theta, \sigma)$ cu θ necunoscut iar $\sigma > 0$ dat. Fie $\pi(\theta)$ de tip $N(\mu, \tau)\tau$ iar $U(\theta, a)$ dat de teorema de mai sus. Atunci $\pi(\theta|x)$ fiind de tip $N\big((\tau\mu + \sigma \cdot \sum_{i=1}^{n} x_i)(\tau + n\sigma)^{-1}, \tau + n\sigma\big)$ deducem ca estimatia Bayes pentru θ este unica mediana a distributiei normale, mediana ce coindice cu media, adica

$$d(x) = \left(\tau\mu + \sigma \cdot \sum_{i=1}^{n} x_i\right) \cdot (\tau + n\sigma)^{-1}$$

Riscul Bayes corespunzator se determina din relatia $r(\pi) = \rho \cdot E^X\big(E^{\pi(\theta|x)}|\theta - d(X)|\big)$, deoarece distributia aposteriori pentru $\theta - d(X)$ cand $X = x$, trebuie sa fie o distributie normal cu media zero si precizia $\tau + n\sigma$, indifferent de valorile pentru $x = (x_1, \dots, x_n)$. Acum, daca $Y \sim N(0, p)$, se poate arata ca $E^{\pi(\theta|x)}|Y| = [2(\pi p)^{-1}]^{1/2}$ .Prin urmare, pentru orice valori pentru X, trebuie sa avem

$$r(\pi) = \rho \cdot [2 \cdot \pi^{-1} \cdot (\tau + n\sigma)^{-1}]^{1/2}$$

O extensie a teoremei de mai sus este,

Teorema 3. *Fie* $U(\theta, a) = \rho_0 \cdot (\theta - a)$ *daca* $\theta \geq a$ *, si* $= \rho_1 \cdot (a - \theta)$ *daca* $\theta < a$ *cu* $\rho_0 > 0$ *,* $\rho_1 > 0$. *Atunci estimatia Bayes pentru* θ *este data de cuantila de ordin* $\rho_0 \cdot (\rho_0 + \rho_1)^{-1}$ *pentru* $\pi(\theta|x)$.

Demonstratia urmeaza calea prezentata in Teorema 1.

In continuare dam cateva exemple in care punem in evident cateva extensii pentru notiunea de decizie Bayes. Considerarea acestor extensii este justificata de urmatorul exemplu :

***Exemplul* 4**:

Fie $\Theta = A = (-\infty, +\infty), U(\theta, a) = (\theta - a)^2$ iar $X \sim N(\theta, 1)$.Vom arata ca estimatia de maxima verosimilitate a lui θ , $d(x) = x$,(care se obtine ca solutie a ecuatiei $\frac{\partial f(x|\theta)}{\partial \theta} = 0$), nu este decizie Bayes relativ la nici-o distributie apriori π .Pe de alta parte, stim ca estimatia de maxima verosimilitate este absolute corecta si deci: $E_\theta d(X) = 0$ pentru orice $\theta \in \Theta$ .Astfel ca $r(\pi, d) = E^\pi E_\theta\big(\theta - d(X)\big)^2 = E^\pi E_\theta \theta^2 - 2E^\pi E_\theta \theta d(X) + E^\pi E_\theta d^2(X) = E^\pi \theta^2 - 2E^\pi\big(\theta E_\theta d(X)\big) + E^\pi E_\theta d^2(X) = E^\pi \theta^2 - 2E^\pi \theta^2 + E^\pi E_\theta d^2(X) = E^\pi E_\theta d^2(X) - E^\pi \theta^2$. Insa $E_\theta d(X) = E^\pi\big(\theta E_\theta d(X)\big) = E^\pi \theta^2$, si de asemenea $E_\theta d(X) = E\big(d(X) \cdot E^{\pi(\theta|x)}\theta\big) = E\big(d^2(X)\big) = E^\pi E_\theta d^2(X)$

,daca am presupune ca $d(x)$ este decizie Bayes relativ la distributia π(Aici am notat cu E media in raport cu distributia vectorului (X,θ)). Din cele de mai sus rezulta $r(\pi,d)=0$ Totusi, observam usor ca pentru $r(\pi,d)$ mai avem si relatia $r(\pi,d)=E^{\pi}E_{\theta}\big(\theta-d(X)\big)^2=E^{\pi}E_{\theta}(\theta-X)^2=E^{\pi}1=1$ deoarece $X\sim N(\theta,1)$ .Astfel obtinem o contradictie si deci estimatorul de maxima verosimilitate *nu este* decizie Bayes. In exemplul urmator aratam ca $d(x)=x$ este limita de functii de decizie Bayes.

Exemplul **5**:

In conditiile din exemplul 4, fie π_n distributia apriori de tip $N(0,n^2)$ .Deoarece densitatea marginala a lui X este

$$f(x)=\frac{1}{\big(2\pi(1-n^2)\big)^{1/2}}\exp\left[-\frac{x^2}{2(1+n^2)}\right]$$

Deducem usor ca

$$\pi_n(\theta|x)=\left(\frac{1+n^2}{2\pi\cdot n^2}\right)^{\frac{1}{2}}\exp\left[-\frac{(1+n^2)}{2n^2}\left(\theta-\frac{xn^2}{1+n^2}\right)^2\right]$$

Si deci $\pi_n(\theta|x)$ este de tip $N\left(\frac{xn^2}{1+n^2},\frac{n^2}{1+n^2}\right)$ .In concluzie decizia Bayes relativ la $\pi_n$este $d_n(x)=E^{\pi_n(\theta|x)}[\theta]=\frac{xn^2}{1+n^2}$ careia ii corespunde riscul Bayes :

$$\begin{aligned}r(\pi_n,d_n)&=(1+n^2)^{-2}E^{\pi_n}E_{\theta}[\theta+n^2(\theta-X)]^2=(1+n^2)^{-2}\{E^{\pi_n}\theta^2+n^4\}\\&=(1+n^2)^{-2}(n^2+n^4)=n^2(1+n^2)^{-1}\end{aligned}$$

Observam ca $\lim_{n\to\infty}d_n(x)=d(x)=x$,pentru orice x, adica d este limita de decizii Bayes.

Urmatoarele doua exemple stabilesc ca estimatorul de maxima verosimilitate de mai sus prezinta alte proprietati de tip Bayes.

Exemplul **6**:

Considerand in exemplul de mai sus o distributie apriori improprie, $\pi(\theta)=\theta\,,\theta\in\Theta$,deducem ca distributia aposteriori a lui θ este $N(x,1)$.Obtinem ca expresia

$E^{\pi}(U(\theta, d)f(x|\theta)$ este minima pentru $d(x) = x$ .Aceasta proprietate desemneaza faptul ca d este o decizie Bayes generalizat.

***Exemplul* 7:**

Aici aratam ca $d(x) = x$ are proprietatea ca pentru orice $\varepsilon > 0$ sufficient de mic, exista o distributie apriori $\pi(\varepsilon)$astfel ca d sa fie decizie ε –Bayes relativ la $\pi(\varepsilon)$.Luand in exemplul 5 distributia apriori $\pi(\varepsilon)$ de tip $N(0, (1-\varepsilon) \cdot \varepsilon^{-1})$ obtinem ca decizia Bayes relativ la $\pi(\varepsilon)$ este $d_{\varepsilon}(x) = x(1-\varepsilon)$iar $\inf r(\pi(\varepsilon), d) = r(\pi(\varepsilon), d_{\varepsilon}) = 1 - \varepsilon$. Insa, din exemplul 4, rezulta, intre altele, ca $r(\pi(\varepsilon), d) = 1$ .Astfel ca $r(\pi(\varepsilon), d) = \inf r(\pi(\varepsilon), d) + \varepsilon$ si prin urmare $d(x) = x$ are proprietatea de mai sus. Aceasta proprietate desemneaza faptul ca d este o functie de decizie Bayes extinsa.

1.6. CAZUL A MULTIME FINITA

Un caz particular important il reprezinta problemele de decizie in care spatial actiunilor este o multime finite. Aceste problem de decizie sunt cunoscute sub numele de problem de decizie multipla. Si aceasta clasa de problem de decizie contine cateva subclase binecunoscute. In continuare ne ocupam de cateva din aceste probleme.

Problemele de decizie in care multimea *A* contine doua elemente sunt cunoscute sub numele de probleme de verificare a ipotezelor statistice. In acest ca o functie de decizie nealeatoare poate fi identificata cu o multime masurabila din spatial de selectie.., cu interpretarea ca, daca $A = \{a_0, a_1\}$, se alege actiunea a_0 in situatia in care variabila aleatoare observata ia valori in acea multime masurabila, si se alege actiunea a_1 in caz contrar. Astfel, o decizie aleatoare revine la o distributie de probabilitate pe clasa tuturor multimilor masurabile din $\mathfrak{X}$ .O utilizare a acestor decizii nu este simpla. Totusi utilizand modul comportamental, lucrarile se simplifica.

Functia de decizie comportamentala este o functie $\delta: \mathfrak{X} \to [0,1]$ cu interpretarea : daca $X = x$, atunci se alege actiunea a_0 cu probabilitatea $\delta(x)$, si actiunea a_1 cu probabilitatea $1 - \delta(x)$ .Cu aceasta abordare, o decizie nealeatoare d se identifica cu decizia comportamentala δ definita prin :

$$\delta(x) = \begin{cases} 1 & ,daca\ x \in d^{-1}(a_0) \\ 0 & ,daca\ x \in d^{-1}(a_1) \end{cases}$$

.In cazul verificarii ipotezelor statistice o decizie comportamentala mai poarta numele de *test*.

Daca δ este o decizie atunci, pentru $\theta \in \Theta$, functia risc este definite de $R(\theta,\delta) = U(\theta,a_0) + \left(U(\theta,a_1) - U(\theta,a_0)\right) \cdot E_\theta\delta(X)$.Termenul $E_\theta\delta(X)$ poarta numele de functie putere a deciziei δ sa a testului δ .Remarcam ca functia risc depinde de δ doar prin intermediul functie putere, ceea ce ne determina sa presupunem ca $E_\theta\delta(X)$ exista pentru orice θ.

Daca $\Theta = \Theta_0 \cup \Theta_1$, $\Theta_0 \cap \Theta_1 = \square$ iar $U(\theta,a_i) = \chi_{\Theta_i}(\theta), i = 0,1$, atunci vorbim de “ipoteza nula” H_0 cand θ se afla in Θ_0 si de “ipoteza alternativa” H_1 cand θ se afla in Θ_1 .Una din aceste ipoteze disjuncte este adevarata, iar decidentul trebuie sa ghiceasca care din ele, pierderea fiind zero daca raspunsul este correct si unu in caz contrar. Prin urmare a_0 poate fi considerate actiunea ce consta in a accepta H_0 iar a_1 actiunea de a accepta H_1 sau de a respinge H_0.Cand Θ_i contine mai mult de un element spunem ca ipoteza H_i este compusa iar cand Θ_i contine un singur element spune ca H_i este o ipoteza simpla.

In acest caz determinarea unei decizii Bayes in raport cu o distributie apriori π revine, in cadrul formei extensive a abordarii bayesiene, la determinarea acelei actiuni pentru care $E^{\pi(\theta|x)}[U(\theta,a_i)]$, $i = 0,1$, este minima. Din forma functiei U obtinem, pentru $i = 0,1$, $E^{\pi(\theta|x)}[U(\theta,a_i)] = P(\Theta_i|x)$ adica decizia Bayes este data de ipoteza cu cea mai mare probabilitate aposteriori.

Cand U este data de : $U(\theta,a_i) = 0$ pentru $\theta \in \Theta_i$ si $= c_i > 0$ pentru $\theta \in \Theta_j$, $j \neq i$, $i,j = 0,1$, atunci pierderile medii aposteriori pentru a_0 si a_1 sunt $c_0 \cdot P(\Theta_1|x)$ si respectiv $c_1 \cdot P(\Theta_0|x)$ .Decizia Bayes revine la actiunea petru care pierderea medie aposteriori este mai mica. De exemplu, se adopta actiunea a_1 cand$c_1 \cdot P(\Theta_0|x) < c_0 \cdot P(\Theta_1|x)$. Cum $\Theta = \Theta_0 \cup \Theta_1$ avem $P(\Theta_0|x) = 1 - P(\Theta_1|x)$; adoptam actiunea a_1 cand $P(\Theta_1|x) > c_1(c_0 + c_1)^{-1}$ .Astfel, in terminologia clasica a testarii ipotezelor, regiunea de respingere a testului Bayesian este $x \in \mathfrak{X} | P(\Theta_1|x) >$

$c_1(c_0 + c_1)^{-1}\}$ care este de aceeasi natura cu regiunea de respingere a testului de verosimilitate.

De exemplu, fie $X \sim N(\theta, 1)$, $\Theta = (-\infty, +\infty)$, $\Theta_0 = \{\theta \in \Theta, \theta \geq \theta_0\}$, $\Theta_1 = \{\theta \in \Theta, \theta < \theta_0\}$, cu θ dat, iar π o functie de desitate normala cu media zero si dispersia unu. Pentru $\pi(\theta|x)$ gasim o densitate normala de medie $x/2$ si dispersie $1/2$. Atunci testul Bayes respinge ipoteza H_0 daca

$$c_1(c_0 + c_1)^{-1} < \frac{1}{\sqrt{\pi}} \int_{\infty}^{0} e^{-\left(\theta - \frac{x}{2}\right)^2} d\theta = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\sqrt{2}(\theta_0 - \frac{x}{2})} e^{-\frac{t^2}{2}} dt$$

Daca $z(\alpha)$ este α −cuantila distributiei normale de medie zero si dispersie unu, atunci testul Bayes respinge H_0 daca $\sqrt{2}\left(\theta_0 - \frac{x}{2}\right) > z(c_1(c_0 + c_1)^{-1})$ sau, in mod echivalent, daca

$$x < 2\theta_0 - \sqrt{2}\, z(c_1(c_0 + c_1)^{-1})$$

Fie $\Theta_0 = \{\theta_0\}$, $\Theta_1 = \{\theta_1\}$ .Notam cu $\alpha_0(\delta) = E_{\theta_0}[\delta(X)]$ probabilitatea care consta in respingerea ipotezei H_0 cand ea este adevarata, numita si probabilitatea erorii de ordinal intai, iar cu $\alpha_1(\delta) = E_{\theta_1}[1 - \delta(X)]$ probabilitatea care consta in a admite ipoteza H_0 cand ea este falsa, numita si probabilitatea erorii de ordinul al doilea. In acest caz spunem ca un test δ este un cel mai bun test la nivelul α pentru a testa H_0 in raport cu H_1 daca$\alpha_0(\delta) = \alpha$ si daca pentru orice test δ' pentru care $\alpha_0(\delta') \leq \alpha$ avem $\alpha_1(\delta) \leq \alpha_1(\delta')$.Se observa ca daca δ este admisibil, atunci el este un cel mai bun test la nivel $\alpha = \alpha_0(\delta)$. Totusi, un cel mai bun test la nivelul $\alpha_0(\delta)$ nu este neaparat admisibil. Lema fundamentala data de Neyman si Pearson (1933) da o metoda generala pentru determinarea celor mai bune teste de acest gen.

Cand consideram problema testarii ipotezei compuse $H_0: \theta \in \Theta_0$ in raport cu alternative compusa $H_1: \theta \in \Theta_1$ notiunea de cel mai bun test la nivelul α se generealizeaza dupa cum urmeaza : δ este un cel mai bun test la nivelul α daca $\sup_{\theta \in \Theta_0} E_\theta[\delta(X)] = \alpha$ .Un test δ_0 este uniform cel mai puternic la nivelul α .Daca are nivelul α si pentru orice alt test δ' la nivelul cel mult α avem $E_\theta[\delta_0(X)] \geq E_\theta[\delta'(X)]$ pentru orice $\theta \in \Theta_1$.

In general nu exista un test uniform cel mai puternic . Totusi, raspunsul este afirmativ pentru clasa densitatilor de probabilitate cu raportul de verosimilitate monoton unde $\Theta \subseteq \mathbb{R}^1$.Aceasta inseamna ca, pentru $\theta_1 < \theta_2$, raportul de verosimilitate $\frac{f(x|\theta_2)}{f(x|\theta_1)}$ este o functie nedescrescatoare de x pe multimea pe care are sens (adica acest raport are sens in x daca cel putin una din cantitatile care-l definesc este pozitiva).

O astfel de clasa este clasa densitatilor exponentiale ce depind de un parametru, date de : $f(x|\theta) = c(\theta)\, H(x) \exp[Q(\theta) \cdot T(x)]$, unde Q si T sunt nedescrescatoare. Aceasta clasa include densitatile normal, binomiala, Poisson, gama si beta.

In continuare ne ocupam de cazul in care multimea actiunilor este finita dar contine mai mult de doua elemente. Aceste probleme poarta numele de probleme de decizie multipla. Fie $A = \{a_1, \dots, a_k\}$ iar δ o functie de decizie comportamentala data de $\delta(x) = \left(\delta(x), \dots, \delta^k(x)\right)$ unde $\delta^i(x)$ este probabilitatea de a adopta actiunea a_i in ipoteza ca $X = x$. Riscul asociat deciziei δ in ipoteza ca θ este starea adevarata se reduce la :

$$R(\theta, \delta) = \sum_{i=1}^{k} \int_{\mathfrak{X}} \delta(x) U(\theta, a_i) f(x|\theta)\, d\mu$$

iar riscul Bayes relative la o distributie apriori π este

$$r(\pi, \delta) = \sum_{i=1}^{k} \delta^i(x) [\int_{\Theta} U(\theta, a_i) f(x|\theta) d\pi]\, d\mu$$

Pentru a stabili forma deciziilor Bayes relativ la π notam pentru $x \in \mathfrak{X}, a_i \in A$:

$$\Delta_i(x, \pi) = \int_{\Theta} [U(\theta, a) - U(\theta, a_1)]\, f(x|\theta) d\pi(\theta)$$

si definim multimile :

$$\mathfrak{X}_i' = \{x \in \mathfrak{X} \,|\Delta_i(x, \pi) > \min_{1 \le j \le k} \Delta_j(x, \pi)\}$$

$i = 1,2, \dots, k$. Fie δ_π functia de decizie data de : $\delta_\pi(x) = \{\delta_\pi^1(x), \dots, \delta_\pi^k(x)\}$ unde $\delta_\pi^i(x) = 0$ pentru $x \in \mathfrak{X}_i'$.

Acum avem :

Teorema 1. *δ_π este o decizie Bayes relativ la π .*

Demonstratie. Fie $\delta_0 \in \mathcal{I}$ .Aratam ca $r(\pi,\delta_0) \geq r(\pi,\delta_\pi)$ .Avem

$$r(\pi,\delta_0) - r(\pi,\delta_\pi) = \sum_{j=1}^{k} \int_{\mathfrak{X}} \delta_0^j(x) \cdot \Delta_j(\pi,x) d\mu - \sum_{i=1}^{k} \int_{\mathfrak{X}} \delta_\pi^i(x) \cdot \Delta_i(\pi,x) d\mu$$

$$= \sum_{j=1}^{k} \sum_{i=1}^{k} \int_{\mathfrak{X}} \delta_0^j(x) \cdot \delta_\pi^i(x) \cdot \left[\Delta_j(\pi,x) - \Delta_i(\pi,x)\right] d\mu$$

Ultima expresie de aici nu poate fi negativa deoarece atunci cand $\Delta_j(\pi,x) < \Delta_i(\pi,x)$ avem $\delta_\pi^i(x) = 0$ Astfel teorema este demonstrata.

Se observa ca daca $\Delta_i(\pi,x) < \Delta_j(\pi,x)$ pentru $j \neq i$, atunci $\delta_\pi^i(x) = 1$.

Daca $\Theta = \{\theta_1, \dots, \theta_k\}$, atunci a poate fi interpretata ca actiunea de a considera ca X are densitatea $f(x|\theta_i)$ .Asemenea probleme de decizie poarta numele de probleme de clasificare si putem considera $U(\theta_i, a_j) = 1$ daca $j \neq i$ si zero in rest. Astfel, riscul asociat deciziei δ, in ipoteza ca starea adevarata este θ_i , are forma $R(\theta_i,\delta) = 1 - \int_{\mathfrak{X}} \delta^i(x) f(x|\theta_i) d\mu$ .Deoarece $\delta^i(x)$ este probabilitatea de a clasifica faptul ca X are densitatea $f(x|\theta_i)$ rezulta ca $\int_{\mathfrak{X}} \delta^i(x) f(x|\theta_i) d\mu$ reprezinta probabilitatea unei clasificari corecte ca $X = x$ sa provina dintr-o populatie cu parametrul θ_i .In concluzie, $R(\theta_i,\delta)$ poate fi interpretata ca probabilitatea unei clasificari incorecte daca se utilizeaza decizia δ iar θ_i, este starea adevarata. Pentru o distributie apriori $\pi = (\pi_1, \dots, \pi_k)$ avem $\Delta_i(\pi,x) = 1 - \pi f(x|\theta_i)$ si $\delta_\pi^i(x) = 0$ daca $\pi f(x|\theta_i) < \max_{1 \leq j \leq k} \pi_j f(x|\theta_j)$.,pentru $1 \leq i \leq k$.

***Exemplul* 1**:

Fie $\mathfrak{X} = (0,\infty), \theta_1 > \theta_2 > \cdots > \theta_k > 0 , f(x|\theta_i) = \theta_i e^{-x\theta_i}$, iar π este distributia uniforma pe Θ .

Pentru $i < j$ avem $f(x|\theta_i) > f(x|\theta_j)$ daca si numai daca $x < \varphi(\theta_i,\theta_j)$, unde $\varphi(\theta_i,\theta_j) = (\theta_i - \theta_j)^{-1} \cdot \ln(\theta_i/\theta_j)$ .Cand θ_i este fixat, functia $\varphi(\theta_i,\cdot)$ este descrescatoare in θ pentru $\theta < \theta_i$; pentru θ_j , fixat, functia $\varphi(\cdot,\theta_j)$ este descrescatoare in θ pentru $\theta < \theta_j$, deoarece $\ln z \leq z - 1$ atunci cand $z > 0$.

Fie $x_i = \varphi(\theta_i,\theta_{i+1})$ cu $i = 1,2, \dots, k-1$, iar $x_0 = 0$, $x_k = +\infty$.

Utilizand Teorema 1, functia de decizie δ_π in raport cu $\pi = \left(\frac{1}{k}, \dots, \frac{1}{k}\right)$ trebuie determinate astfel ca in ipoteza $X = x$ sa se adopte actiunea a_i pentru care indicele i face maxima cantitatea $f(x|\theta_i)$ .Dar $f(x|\theta_i)$ este descrescatoare in x pe $\mathfrak{X}$ .Prin urmare trebuie sa avem :

$$\delta_\pi^i(x) = \begin{cases} 1 & x_{i-1} \leq x < x_i \\ 0 & in\,rest \end{cases}$$

Sa observam ca daca $\pi = (\pi_1, \dots, \pi_k)$ unde $\pi_i = \left[\theta_i \cdot \sum_{j=1}^{k} \theta_j^{-1}\right]^{-1}$ atunci$e^{-x\theta_i} < e^{-x\theta_j}$ pentru $i < j$. Deci pentru orice $i \neq k$,$\delta_\pi^i(x) = 0$ si $\delta_\pi^k(x) = 1$ cand $x \in \mathfrak{X}$ In concluzie trebuie sa renuntam la informatia furnizata de variabila aleatoare X.

2. APLICATII

2.1. TESTE GRILA

Pentru un examen la care se prezinta $N \geq 1$ candidati se da un test.Testul cuprinde $n > 1$ rubrici,fiecare cu $a > 1$ alternative de raspuns.Candidatul trebuie sa aleaga un raspuns, in mod independent,omisiunile nefiind premise.Se presupune ca un candidat cunoaste raspunsul correct la θ rubrici $(\theta = 0,1, \dots, n)$ si da un raspuns gresit drept raspuns correct la ω rubrici $(\omega = 0,1, \dots, n - \theta)$ .Candidatul are deci informatie gresita asupra ω rubrici.Atat θ cat si ω sunt necunoscute celui care inregistreaza rezultatul testului ele fiind specific fiecarui candidat in parte. Numarul total de raspunsuri exacte poate fi marit de candidat prin ghicire la intamplare prin a alternative pentru fiecare din cele $n - \theta - \omega$ rubrici la care el nu cumoaste raspunsul.

Numarul total de raspunsuri corecte R ,se considera drept o variabila aleatoare exprimata prin suma raspunsurilor corecte θ si o componenta repartizata binominal de parametrii $n - \theta - \omega$ si $1/a$.

Modelul traditional ,numit si “cunoastere sau ghicire la intamplare” ,nu-l ia in consideratie pe ω, acesta fiind presupus egal cu 0. Astfel ca numarul total de raspunsuri exacte ,R, este considerat acum ,o variabile aleatoare egala cu suma dintre θ si o componenta repartizata binominal ,dar de parametri $n - \theta$ si $1/a$.

Ceea ce stie cel care inregistreaza rezultatul testului despre un candidat este numarul de raspunsuri corecte date .Ceea ce stie un candidat despre un test este cuplul (θ, ω).Este natural ca inregistrarea sa se faca dupa o functie de acest cuplu .Intrucat valorile adevarate ale lui θ si ω nu sunt cunoscute ,etse necesar ca acestea sa fie estimate.

Sa presupunem ca θ are o repartitie apriori data de functia de frecventa $\pi(\theta)$

$$\theta = 0,1,2,\dots,n\ ;\ \pi(\theta) \geq 0;\ \sum_{\theta=0}^{n} \pi(\theta) = 1$$

Alegem pentru ω o reparitie conditionata de θ ,si anume,uniforma pe intregii $\{0,1,2,\dots,n-\theta\}$. Atunci densitatea apriori comuna a lui θ si ω este

$$\pi(\theta)/(n-\theta+1)$$

$$\theta = 0,1,\dots,n\ ;\ \omega = 0,1,\dots,n-\theta\ .$$

Repartitia lui r din selectie,sau functia de verosimilitate,este data de

$$f(r|\theta,\omega) = P(R = r|\theta,\omega) = P(R-\theta = r-\theta|\theta,\omega)$$

$$= C_{n-\theta-\omega}^{r-\theta}\left(\frac{1}{a}\right)^{r-\theta}\left(\frac{a-1}{a}\right)^{n-\omega-r}$$

Date fiind repartitia apriori si functia de verosimilitate ,dupa teorema lui Bayes,densitatea aposteriori este

$$\pi(\theta,\omega|r) = \frac{C_{n-\theta-\omega}^{r-\theta}\,\frac{a^{\omega+\theta}}{(a-1)^{\omega}}\cdot\frac{\pi(\theta)}{n-\theta+1}}{\sum_{i=0}^{r}\frac{a^{i}\pi(i)}{n-i+1}\cdot\sum_{j=0}^{n-r}C_{n-i-j}^{r-i}\,\frac{a^{j}}{(a-1)^{j}}}$$

$$\theta = 0,1,\dots,r\ ;\ \omega = 0,1,\dots,n-r\ .$$

Densitatile aposteriori marginale ale lui θ si ω se gasesc sumand in relatia de mai sus dupa ω si ,respectiv θ.

Daca se considera modelul traditional, rezultatele sunt mai simple ca expresii matematice.Functia de verosimilitate este:

$$f(r|\theta) = C_{n-\theta}^{r-\theta}\left(\frac{1}{a}\right)^{r-\theta}\left(\frac{a-1}{a}\right)^{n-\theta}\ ,$$

Iar repartitia aposteriori este data de

$$\pi(\theta|r) = \frac{C_{n-\theta}^{r-\theta}\pi(\theta)a^{\theta}}{\sum_{i=0}^{r}C_{n-i}^{r-i}\pi(i)a^{i}}\ ;\ \theta = 0,1,\dots,r$$

O persoana care s-a ocupat mai mult timp cu darea si inregistrarea rezultatelor testelor,ar putea folosi experienta lor trecuta pentru a evalua repartitia neconditionata

$P(R = r)$, $r = 0,1,2,\dots,n$, mai usor decat sa evalueze direct pe $\pi(\theta)$.Din densitatea de repartitie comuna

$$f(r|\theta) = P(R = r|\theta)\pi(\theta)$$

Pentru modelul traditional si

$$f(r,\theta,\omega) = P(R = r|\theta,\omega)\pi(\theta,\omega)$$

Pentru modelul cu informatie gresita,se obtine densitatea marginala $P(R = r)$,sumand dupa parametrii respective.Astfel se obtin urmatoarele doua sisteme de ecuatii,care trebuiesc verificate de $\pi(\theta)$ in fiecare din cele doua cazuri.

$$P(R = r) = \sum_{\theta=0}^{r} C_{n-\theta}^{r-\theta}\left(\frac{1}{a}\right)^{r-\theta}\left(\frac{a-1}{a}\right)^{n-r}\pi(\theta)$$

Si

$$P(R = r) = \sum_{\theta=0}^{r}\left[\frac{1}{n-\theta+1}\sum_{\omega=0}^{n-r} C_{n-\theta-\omega}^{r-\theta}\left(\frac{1}{a}\right)^{r-\theta}\left(\frac{a-1}{a}\right)^{n-\omega-r}\right]\pi(\theta)$$

$$r = 0,1,2,\dots,n.$$

Deoarece matricea sistemului de mai sus este inferior triunghiulara, avand elementele de pe diagonala pozitive, rezulta ca acest system are solutie unica. Dar pentru modelul traditional, substitutia unui vector arbitrar de probabilitati pentru $P(R = r), r = 0,1,2,\dots,n$ poate duce in system la o solutie , $\pi(\theta), \theta = 0,1,\dots,n$ care nu este neaparat o probabilitate , avand $\sum_{\theta=0}^{n}\pi(\theta) = 1$,dar putand avea valori negative sau mai mari decat 1.

O alta metoda pentru evaluarea unei repartitii aprori pentru θ este aceea care foloseste

$$(1) \qquad P(R = r) = C_n^r\left(\frac{1}{a}\right)^r\left(\frac{a-1}{a}\right)^{n-r}.$$

In modelul traditioanal aceasta duce intotdeauna la $(0) = 1$.

Daca consideram $n = 2$ si $a = 2$,si egalam partile din dreapta ale celor doua ecuatii de mai sus se obtine sistemul de ecuatii:

$$\frac{1}{4} = \frac{7}{12}\,\pi(0)$$

$$\frac{1}{2} = \frac{1}{3}\pi(0) + \frac{3}{4}\pi(1)$$

$$\frac{1}{4} = \frac{1}{12}\pi(0) + \frac{1}{4}\pi(1) + \pi(2)$$

Sistemul are urmatoarea solutie: $\pi(0) = \frac{9}{21}$; $\pi(1) = \frac{10}{21}$; $\pi(2) = \frac{2}{21}$.

Valorile probabilitatii aposteriori pentru modelul de informative gresita si traditional si valorile medii bazate pe probabilitatile apriori din relatia (1) sunt expuse in tabelul urmator:

r	θ			
	0	1	2	$E(\theta\|r)$
0	1.00 (1.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
1	0.20 (0.47)	0.80 (0.53)	0.00 (0.00)	0.80 (0.53)
2	0.14 (0.24)	0.48 (0.54)	0.38 (0.22)	1.24 (0.98)

Pentru exemplificare am anexat si un astfel de test grila.

Subiecte de admitere la ASE Bucuresti
Facultatea: Finante-Asigurari

I.Fie

$M = \{m \in R \mid x^4 - (m-1)x^3 + mx^2 - (m-1)x + 1 = 0$ are numai doua radacini reale $x_1, x_2, x_1 \neq x_2\}$

Pentru m=13, fie $S = x_1^4 + x_2^4$. Atunci:

1) (5 p.) A) $M \subset (-\infty,1] \cup [2,\infty)$ B) $M \cap (-1,2) = \varnothing$ C) $M \supset (3,\infty)$ D) $M = [0,20]$ E) $M = \{13\}$
2) (3 p.) A) S = 431 B) S = 14159 C) S = 17304 D) S = 357 E) S = 14100

II. Pe multimea $A = \mathbf{Q} \times \mathbf{R}$ se definesc legile de compozitie "$\circ$" si "$*$" astfel:

$(a,x) \circ (b,y) = (a+b, x+y)$

$(a,x) * (b,y) = (ab, ax + by + xy),\ \forall (a,x),(b,y) \in A$

Daca B este multimea elementelor inversabile ale inelului $(A,\circ,*)$ si T=a+x, unde (a,x) este inversul elementului $\left(\frac{1}{3},5\right) \in A$, atunci:

3) (5 p.) A) $B \subset \mathbf{Q} \times \mathbf{Q}$ B) $\mathbf{Q} \times \mathbf{Q} \subset B$ C) $B = \{(a,x) | a \in \mathbf{Q} \setminus \{0\}, x \in \mathbf{R} \setminus \{-a\}\}$

D) $B = \{(a,x) | a \in \mathbf{Q} \setminus \{0\}, x \in \mathbf{Q} \setminus \{-a\}\}$ E) $B = \left\{\left(\frac{1}{a}, x\right) \middle| a \in \mathbf{Z} \setminus \{0\}, x \in \mathbf{Z}\right\}$

4) (3 p.) A) $T = \frac{16}{3}$ B) $T = 4$ C) $T = \frac{5}{3}$ D) $T = \frac{3}{5}$ E) $T = \frac{3}{16}$

III. Se considera functia $f_a : \mathbf{R} \to \mathbf{R},\ f_a(x) = \sqrt[3]{4(e^x - x - 1) - x^3 + (a-3)x^2}, a \in \mathbf{R}$

Daca $A = \{a \in \mathbf{R} | f_a \text{ este derivabila in } x = 0\}$ si $T = f_a'(0)$ pentru $a \in A$, atunci:

5) (5 p.) A) $A \subset \left(-3, -\frac{1}{2}\right)$ B) $A \subset \left(-\frac{1}{2}, \frac{3}{2}\right)$ C) $A \subset \left(\frac{5}{2}, 5\right)$ D) $A \subset \left(\frac{9}{2}, \frac{13}{2}\right)$

E) $A \subset (7,15)$

6) (3 p.) A) $T = -\sqrt{2}$ B) $T = \sqrt[3]{3}$ C) $T = \frac{1}{2}$ D) $T = -\frac{1}{\sqrt[3]{3}}$ E) $T = \sqrt[3]{e+1}$

IV. Se considera functia $f_b : D \to \mathbf{R}$ unde

$f_b(x) = \lim_{n\to\infty} \frac{bx^2 \cdot e^{n^{\ln x}} + 8x}{x \cdot e^{n^{\ln x}} + 2}$, $b \in \mathbf{R}$, D fiind domeniul maxim de definitie al functiei. Daca $B = \{b \in \mathbf{R} \mid f_b \text{ este derivabila pe D}\}$ si $I = \int_{\frac{1}{2}}^{\frac{7}{2}} f_b(x)dx$ pentru $b \in B$, atunci:

7) (5 p.) A) $B \subset \left(-1, -\frac{1}{2}\right)$ B) $B \subset \left(-\frac{1}{2}, \frac{1}{2}\right)$ C) $B \subset \left(\frac{3}{4}, \frac{5}{3}\right)$ D) $B \subset \left(\frac{7}{3}, \frac{13}{3}\right)$

E) $B \subset \left(\frac{16}{3}, \frac{17}{2}\right)$

8) (3 p.) A) I=13 B) I=17 C) I=24 D) I=21 E) I=32

V. Se considera matricea:

$$A_m = \begin{pmatrix} 3 & m & -1 & 2 \\ 2 & 3 & -4 & 4 \\ 4 & 2 & -2m & 1 \\ 1 & 4 & -3 & 3 \end{pmatrix}, m \in \mathbf{R}$$

Daca $M = \{m \in \mathbf{R} \mid rang\ A_m = 3\}$ si $T = \det A^*_{-3}$, atunci:

9) (5 p.) A) $M \subset \{-6, -2, 0, 1, 2, 3\}$ B) $M \subset (-3, 3)$ C) $M = \{0, 1, 4\}$

D) $M \subset \{-4, -2, 2, 4, 8\}$ E) $M \subset \{-5, -4, -1, 2, 5\}$

10) (3 p.) A) T = -512 B) T = -256 C) T = -2 D) T = 64 E) T = 112

2.2. REPARTIZAREA ELEVILOR DE LICEU

In majoritatea liceelor economice sau cu profil tehnologic,la sfarsitul clasei a zecea elevi sunt repartizati pe specializari in functie de media dupa primi doi ani de studiu.

Inainte de repartitie elevi sunt ajutati sa isi exprime optiunea pentru calificarea pe care vor sa o urmeze .Pentru aceasta se folosesc ca mod de orientare mediile de admitere din anul anterior.Acestea reprezinta si informatia apriori pe care o detinem inaintea efectuari repartitiei.

Fie $n \geq 1$ specializari,pe care le consideram populatii normale de dispersii egale cu unu si medii cunoscute $\mu_1, \dots \mu_n$ astfel ca $\mu_1 < \mu_2 < \cdots < \mu_n$. Se face cate o observatie (consideram cate un elev) pentru fiecare din aceste populatii iar apoi aceste observatii se pun intr-o ordine arbitrara, de exemplu $X_1, \dots, X_n$ .Acum problema consta in a decide in ce specializare se poate incadra fiecare elev X_i.

Sa consideram $\Theta = A = \{\sigma | \sigma \ permutare\ pe\ \{1, \dots, n\}\}$ si

$$U(\theta, a) = \begin{cases} 0 & \text{, daca } \theta = a \\ 1 & \text{, daca } \theta \neq a \end{cases}$$

Cum variabilele aleatoare $X_1, \dots, X_n$ sunt independente, iar in ipoteza $\theta = (\sigma_1, \dots, \sigma_n)$ variabila aleatoare X_i este repartizata normal cu media μ_{σ_i} si dispersia 1, densitatea de probabilitate a vectorului $X = (X_1, \dots, X_n)$ este:

$$f(x_1, \dots, x_n | \theta) = (2\pi)^{-n/2} \exp\{-\frac{1}{2} \sum_{i=1}^{n} (x - \mu_{\sigma_i})^2\}$$

Pe de alta parte, deoarece observatiile au fost puse in ordine aleatoare avem ca θ are o distributie uniforma pe Θ care are $k = n!$ elemente. Deci, fie $\pi = (\pi_1, \dots, \pi_k)$ distributia uniforma pe Θ .Conform Teoremei 1, functia de decizie Bayes relativ la π , ne conduce la o actiune $a = (\sigma_1, \ldots, \sigma_n)$ pentru care $f(x_1, \ldots, x_n|\theta_n)$ este maxima. Atunci trebuie sa determinam $(\sigma_1, \ldots, \sigma_n)$ pentru care

$$\sum_{i=1}^{n} \left(x - \mu_{\sigma_i}\right)^2$$

este minima. Insa aceasta expresie este minima pentru acea permutare $\tilde{\sigma}$ pentru care $x_{\widetilde{\sigma_1}} > x_{\widetilde{\sigma_2}} > \cdots > x_{\widetilde{\sigma_n}}$.

Pentru a arata acest lucru fie $\bar{\sigma}$ inversa permutarii $\tilde{\sigma}$.

Atunci $x_{\bar{\sigma}_1} < x_{\bar{\sigma}_2} < \cdots < x_{\bar{\sigma}_n}$. Daca σ este o permutare arbitrara cu permutarea inversa σ' avem:

$$\sum_{i=1}^{n} \left(x_i - \mu_{\sigma_i}\right)^2 - \sum_{i=1}^{n} \left(x_i - \mu_{\bar{\sigma}_i}\right)^2 = \sum_{i=1}^{n} \left(x_{\sigma'_i} - \mu_i\right)^2 - \sum_{i=1}^{n} (x_{\tilde{\sigma}} - \mu_i)^2 =$$

$$= 2\sum_{i=1}^{n} (x_{\tilde{\sigma}_i} - x_{\sigma'_i})\mu_i = 2\sum_{i=2}^{n} [\sum_{j=i}^{n} (x_{\sigma'_i} - x_{\sigma'_i})] \cdot (\mu_i - \mu_{i-1}) \geq 0$$

Deoarece

$$\sum_{i=1}^{n} x^2{}_{\widetilde{\sigma_i}} = \sum_{i=1}^{n} x^2{}_{\sigma'_i} \ , (\mu_i > \mu_{i-1}),$$

pentru orice i ,si

$$\sum_{j=i}^{n} x_{\widetilde{\sigma_i}} \geq \sum_{j=i}^{n} x_{\sigma'_i}$$

Pentru orice i si orice permutare σ' .

In continuare sunt prezentate datele de la un Colegiu Economic in care se aplica procedura prezentata mai sus pentru $n = 5$.

SPECIALIZARE	MEDIE
Tehnician în achiziţii şi contractări	**8,62**
Tehnician în activităţi economice	**8,95**
Tehnician in activitati de comert	**8,65**
Organizator banqueting	**7.58**
Tehnician în turism	**8,39**

BIBLIOGRAFIE

1. Preda,V. Teoria deciziilor statistice, Editura Academiei Romane,Bucuresti,1992.
2. Ciucu, G., Craiu, V.,Stefanescu, A.,Stefanescu, M.V. Statistica Matematica si Cercetari Operationale, Editura Didactica si Pedagogica , Bucuresti, 1982.
3. Craiu, V., Paunescu, V. , Elemente de statistica matematica cu aplicatii, Editura Mondo-Ec, 1998
4. Craiu, V., Preda, V., Probleme dedecizie multipla ,Editura Universitatii Bucuresti, 1980.
5. http://en.wikipedia.org/wiki/Bayesian_inference