tanulmányok 42/1975

MTA Számitástechnikai és Automatizálási Kutató Intézet Budapest

MAGYAR TUDOMÁNYOS AKADÉMIA
SZÁMITÁSTECHNIKAI ÉS AUTOMATIZÁLÁSI KUTATÓ INTÉZETE

FEJEZETEK A MATEMATIKAI STATISZTIKÁBÓL SZÁMITÓGÉPES ALKALMAZÁSOKKAL

Arató Mátyás

Tanulmányok 42/1975

A kiadásért felelős

Dr. Gertler János

ISBN 963 311 010 6

Készült az Országos Müszaki Könyvtár és Dokumentációs
Központ házi sokszorositójában
F.v.: Janoch Gyula
Eng. sz.: 54549

Tartalomjegyzék

Az előkészületben lévő II.Fejezet tartalomjegyzéke

II.FEJEZET TÖBBDIMENZIÓS ANALIZIS

B e v e z e t é s

A könyv azoknak az előadásoknak az alapján iródott, amelyeket az 1971-es évtől kezdődően a budapesti Eötvös Lóránd Tudományegyetemen tartottam ötödéves matematikus hallgatóknak. Az előadásban feltételezem a valószinüségelmélet, a sztohasztikus folyamatok elméletének és a matematikai statisztika elemeinek ismeretét. Igy a könyv nem bevezető jellegü, és több helyen elég sokat tételez fel az olvasóról.

Az első fejezetben a szekvenciális eljárásokat vizsgálom mint hipotézisvizsgálati módszert. Bizonyitásra kerül a likelihood hányadoson alapuló eljárás optimális tulajdonsága. Alapvető egyenlőtlenségek bizonyitása mellett a Bayes féle módszert is ismertetem, amely szekvenciális eljárások esetén is alapvető döntési eljárás.

A szekvenciális eljárások megértésénél és a további feladatok kijelölésénél fontos szerepet játszik a Brown-mozgás folyamata (Wiener folyamat) statisztikai problémáinak vizsgálata. Igy a klasszikus szekvenciális feladaton kivül tárgyalom a "riasztás" vagy meghibásodás ("razladka") és bizonyos megállitási szabályok ("stopping rules") azon eredményeit, amelyek statisztikai értelmezése és haszna nyilvánvaló.

A meghibásodás feladata és megoldása lényeges lépést jelentett a hatvanas évek elején a statisztikai kutatásokban. A feladat jelentőségének megértésében hálával emlékezem azokra a beszélgetésekre, amelyeket A. Kolmogorovval és A. Sirjajevvel folytattam, akiknek a kutatásai uttörő jelentőségüek voltak.

A számitógépek operációs rendszereiben használatos statisztikus eljárások közül kiemeltem a lapolási eljárásoknál használatos szekvenciális eljárásokat. A könyvben ismertetésre kerülő hozzáállás és megoldás az irodalomban nem ismert. Bizonyos Markov lánc szabályozási és irányitási feladat bemutatását, valamint a Kálmán-Bucy szürő elvének ismertetését feltétlenül szükségesnek láttam.

Sem az előadásokban, sem a könyvben nem törekedhettem teljességre. Az érdeklődő olvasó az irodalom alapján részletesen tájékozódhat.

A könyv második fejezete a többdimenziós analizissel foglalkozik.* Itt is törekedtem megismertetni a hallgatóimat a legujabb eredmények-

* Előkészületben

kel. A többdimenziós normális eloszlás statisztikai problémáinak vizsgálata a huszas évek óta jelentős helyet foglal el a statisztikai vizsgálatokban. A számitógépek megjelenésével ez az irány egyre jobban virágzik. A ma már klasszikus eredményeket ismertetve vizsgálhatjuk a kanonikus korreláció, főkomponens analizis, diszkriminancia analizis, faktor analizis és cluster analizis eredményeit és módszereit. Ezeknek a módszereknek a hatékonyságát több példán keresztül is be kivánom mutatni.

A könyvben szép számmal szerepelnek példák és feladatok is. Megértésük ill. megoldásuk nagyban segithetik az olvasót az ismertetett anyag önálló elsajátitásában.

Ahol lehetséges kitérek a számitógépes realizálásokra és számitógépes feladatokra, hiszen csak igy képzelhető el a könyvben vizsgált módszerek általános alkalmazása. A többdimenziós analizis részben programok ismertetése is szerepel. Ezek felhasználása nagyban segitheti az olvasót a modern statisztikai módszerek gyors hasznositásában.

Az egyetemi előadások során igyekeztem érzékeltetni hallgatóimmal a statisztikus magatartását a mai viszonyok között. Ezt nehezebb könyvben bemutatni, hiszen annak hangsulyozása, hogy milyen hatással van a modern számitástechnika a valószinüségelméletre és a matematikai statisztikára, sokszor csak megjegyzéseken mulik. Az uj technika megjelenése nemcsak azt jelenti, hogy a jelenlegi számitógépek memóriája és sebessége szinte fantasztikus, hanem azt is, hogy milyen jelentőségü pl. az input-output müveletek és a rendezések megszervezése, az interaktiv időosztásos számitógépes terminálrendszerek kialakitása, továbbá a programozási nyelvek haszna. Ez utóbbival kapcsolatban a nálunk elsősorban használt FORTRAN, COBOL és SIMULA nyelvekre történik utalás. A feladatokat CDC-3300-ra készült FORTRAN-nyelvü programokkal illusztrálom.

Az egyetemi előadások célja elsősorban a matematikai statisztikai módszerek megismertetése volt. Szükség van közben a matematikai statisztikai (és részben a valószinüségelméleti) modellalkotás kimunkálására is. A jelenlegi szinten a modellalkotás lehetőségei jóval nagyobbak, mint a klasszikus matematikai statisztikában. A hallgatóknak érezni kell, hogy a számitógépek felhasználásával jóval nagyobb szabadságuk van a modell megválasztásában is.

A könyvben elsősorban a módszerek megismertetését tüztem ki célnak; ez azzal a veszéllyel járhat – és ez a veszély egyáltalán nem lebecsülendő –, hogy a matematikai statisztika módszereiben jártas hallgatóink valódi problémák megoldásában járatlanok maradnak.

Az a hozzáállás, amely lebecsüli a csiszolatlan formában felvetett kérdésekkel kapcsolatban a modellalkotás, az adatszolgáltatás (gyüjtés) megszervezésének jelentőségét, nem vezethet sikeres megoldásához még a legszebb, de a valósághoz nem illeszkedő formális matematikai apparátus felhasználása utján sem.

Az analitikus és számitógépes megoldások helyes arányának és lehetőségeinek megtalálása nem kis gyakorlatot kiván. Hazai viszonylatban ezen a téren még nincsenek komoly tapasztalataink, az egyetemi előadások és jelen könyv is csak első kisérletnek tekinthetők.

Javasolt irodalom

1. Blackwell, D., Girshick, M.: Teorija igr i sztatisztìcseszkih resenij. Il.(1958), Moszkva.
2. Ghosh, B. K.: Sequential tests of statistical hypotheses. Adison-Wesley (1970), New York.
3. DeGroot, M. H.: Optimal statistical decisions. McGraw Hill, (1970), New York (oroszul is).
4. Kendall, M. G., Stuart, A.: The advanced theory of statistics. I-III, Hafner (1966), New York.
5. Lehman, E.: Testing statistical hypotheses. (1959), New York, (oroszul is).
6. Lipcer, R., Shirjaev, A. N.: Sztatisztika szlucsajnüh processzov. Nauka, Moszkva (1974).
7. Morrison, D. F.: Multivariate statistical methods. McGraw-Hill, (1967), New York.
8. Prékopa A.: Valószinüségelmélet. Müszaki Kiadó (1962),Budapest.
9. Rao, R.: Linear statistical inference. Wiley, (1965), New York, (oroszul is).
10. Rényi A.: Valószinüségszámitás. Tankönyvkiadó, (1965), Budapest.
11. Sirjajev, A. N.: Sztatiszticseszkij poszledovatyelnüj analiz. Nauka, (1969), Moszkva.
12. Wald, A.: Sequential analysis. - (Poszledovatyelnüj analiz.) Fizmatgiz, (1960), Moszkva.
13. Wetherill, G. B.: Sequential methods in statistics. McMillan, (1966), London.

I. Fejezet

Szekvenciális eljárások és feladatok

1. A statisztikai hipotézisvizsgálatról.

Ha a megfigyelési eredmények alapján arról kell döntenünk, hogy valamilyen előre meghatározott paraméter érték (vagy értékek) elfogadható(k)-e hipotézis vizsgálatról beszélünk. Legyen H_0 egy egyszerü hipotézis. A H_0 feltétel mellett a kritikus (H_0-át elvető) tartomány W valószinüsége (a H_0 : P_0 jelölést használjuk a 0 paraméterü eloszlásra vonatkozó hipotézis jelölésére):

$$P_0(W) = \alpha \qquad (= \alpha_0) \qquad \text{elsőfaju hiba valószinüsége.}$$

(W a kritikus tartomány).

Az α értéket szignifikancia szintnek is szokás nevezni. A másodfaju hiba az ellenhipotézis elvetésével kapcsolatos. Ha S jelöli a teljes mintateret és a h paraméterhez tartozó hipotézist valamint eloszlást H_h ill. P_h-val jelöljük a

$$P_h(S-W) = \beta(h), \qquad (P_1(S-W) = \alpha_1)$$

jelölésekkel a következő definiciókat vezetjük be.*

A $\beta(h)$ függvényt erőfüggvénynek nevezzük. A másodfaju hiba valószinüségét összetett ellenhipotézis esetén a következő módon definiáljuk

$$\beta = \sup_h \beta(h)$$

(hasonlóan az elsőfaju hiba valószinüsége összetett hipotézis esetén $\alpha = \sup_{h \in H_0} P_h(W)$).

A klasszikus hipotézis vizsgálat alapvető eredményeit Neyman és Pearson érték el a harmincas években. Összefoglaló munka Lehmann (1959) könyve.

---

* A későbbiekben az alaptér jelölésére vagy S vagy X szolgál. A valószinüségi változókat vagy görög $\xi, \eta, \zeta, \dots$ vagy latin nagybetükkel jelöljük X, Y,... . Remélhetőleg ez nem vezet félreértésre.

1. Lemma. (A Neyman-Pearson lemma.)

Legyen $p_0(\underline{x})$ a H_0, $p_1(\underline{x})$ a hipotézis melletti sürüség. A w kritikus tartomány optimális, ha

$$\int_W p_0(\underline{x})\, d\underline{x} = \alpha \qquad (\alpha \text{ adott}),$$

$$\int_W p_1(\underline{x})\, d\underline{x} = \text{maximum},$$

teljesül. Az optimalizálási feladat megoldását a következő w tartomány adja:

$$W = \{\underline{x} : p_1(\underline{x}) \geqslant k\, p_0(\underline{x})\}.$$

2. Lemma. (Általánositás több függvény esetére)

Legyenek f_0, f_1, ... integrálhatók a μ mértékre mérve az S téren. w legyen olyan tartomány, melyre

$$(\ast) \quad \int_W f_i\, d\mu = C_i \qquad (i=1,2,\ldots,)$$

ahol a C_i számok adottak ($i=1,2,\ldots$). Létezzenek olyan $k_1, k_2, \ldots$ állandók, hogy a W_0 tartományban

$$f_0 \geqslant k_1 f_1 + k_2 f_2 + \cdots$$

és $\overline{W}_0$-ban

$$f_0 \leqq k_1 f_1 + k_2 f_2 + \cdots$$

Továbbá ($\ast$) is teljesül W_0-ra. Akkor

$$\int_{W_0} f_0\, d\mu \geqslant \int_W f_0\, d\mu .$$

Az általános Neyman-Pearson lemma bizonyitása. A C_i értékekből levonva a $W_0 \cap W$ halmazon az integrál értékét adódik, hogy

$$\int_{W - W\cap W_0} f_i\, d\mu = \int_{W_0 - W\cap W_0} f_i\, d\mu \qquad (i=1,2,\ldots)$$

Ennek alapján

$$\int_{W_0} f_0\, d\mu - \int_W f_0\, d\mu = \int_{W_0 - W\cap W_0} f_0\, d\mu - \int_{W - W\cap W_0} f_0\, d\mu \geqslant \int_{W_0 - W_0\cap W} \sum_1^\infty k_i f_i\, d\mu - \int_{W - W_0\cap W} \sum_1^\infty k_i f_i\, d\mu = 0$$

amiből az állitás következik.

3. Lemma. (Randomizált eset.) Legyenek az f_i függvények integrálhatók, φ pontfüggvény, melyre $0 \leqq \varphi \leqq 1$ és

(*) $\int f_i \varphi \, d\mu = c_i$ (i=1,2,...)

a c_i számok adottak. Legyen φ^* olyan, hogy (*) teljesül és

$$\varphi^* = \begin{cases} 0 & \text{ha } f_0 < k_1 f_1 + \dots \\ 1 & \text{ha } f_0 > k_1 f_1 + \dots \end{cases}$$

és φ^* tetszőleges, ha $f_0 = k_1 f_1 + \dots$. Akkor

$$\int f_0 \varphi^* d\mu \geq \int f_0 \varphi \, d\mu$$

Bizonyitás: Jelöljük a következő halmazokat S_i-vel (i=1,2,3),

$$S_1 = \{ \underline{x} : f_0 < \sum k_i f_i \}, \quad \varphi^*(x) - \varphi(x) \leq 0 \quad \text{ha} \quad x \in S_1,$$

$$S_2 = \{ x : f_0 > \sum k_i f_i \}, \quad \varphi^*(x) - \varphi(x) \geq 0 \quad \text{ha} \quad x \in S_2,$$

$$S_3 = \{ x : f_0 = \sum k_i f_i \}$$

Mindhárom halmazon $(\varphi^* - \varphi) f_0 \geq (\varphi^* - \varphi) \sum_1^\infty k_i f_i$. Igy

$$\int_{S_1 \cup S_2 \cup S_3} f_0 (\varphi^* - \varphi) d\mu \geq \int \sum_1^\infty k_i f_i (\varphi^* - \varphi) d\mu = 0$$

a (*) feltétel miatt. Ezzel a lemma bizonyitása kész.

2. A likelihood hányados alapján vizsgált szekvenciális eljárások

2.1. Alapvető összefüggések

A gyakorlatban valójában minden kisérletet igyekszünk szekvenciálisan végezni (mindaddig amig elegendő eredményünk nincs).

A Neyman-Pearson lemma optimalitást biztosit az n-dimenziós megfigyelési téren a következő szabályok alapján (feltéve, hogy léteznek a folytonos sürüségfüggvények):

$p_{1n}(\underline{x})/p_{0n}(\underline{x}) \geq C$ esetén elfogadni H_0-at,

$p_{1n}(\underline{x})/p_{0n}(\underline{x}) < C$ esetén elfogadni H_1-et.

Ezt a szabályt próbáljuk független megfigyelések esetén szekvenciális döntési szabállyá átalakitani. Előbb azonban lássunk néhány példát. Ritka vércsoporthoz tartozó egyének százalékos előfordulására vagyunk kiváncsiak. Ezt ugy oldjuk meg, hogy mindaddig vizsgálunk egyedeket, amig 50 (vagy más meghatározott számu) a vércsoportba tartozó egyedet nem találunk. A megvizsgált egyedek száma ebben az esetben nyilvánvalóan véletlen lesz.

Hasonlóan szekvenciális eljárást alkalmazunk olyan esetekben számológépes feldolgozásnál, amikor ritkán előforduló valószinüségi változók eloszlására kivánunk felvilágositást kapni (pl.SIMULA nyelven irt programok periféria igényét vizsgáljuk).

Az első kisérlet szekvenciális eljárásra Dodge és Romig nevéhez füződik (Sampling Inspection Tables, Wiley, 1944).

Megvizsgálva n_1 terméket azt tapasztaljuk, hogy r_1 a hibásak száma. A következő szabály alapján döntünk az egész tétel elfogadásáról:

$r_1 < c_1$ esetén elfogadni az egész tételt,

$r_1 > c_2$ " elvetni " " " ,

$c_1 < r_1 < c_2$ " folytatni a mintavételt n_2 ujabb megfigyeléssel és a következő döntési szabállyal:

A tétel, ha r_2 jelöli az ujabb hibás elemek számát,

$r_1 + r_2 < c$ esetén elfogadásra kerül,

$r_1 + r_2 > c$ " elvetésre " .

A modern szekvenciális eljárások kidolgozásának kezdete Wald Ábrahám nevéhez fűződik (Sequential analysis, Wiley, 1947.). Az általa javasolt likelihood hányadoson alapuló eljárást ismertetjük.

Legyenek megadva az A_0, A_1 konstansok $(0 < A_0 < 1 < A_1 < \infty)$ és az $X_1, X_2, \ldots, X_n$ megfigyelést folytassuk, amig teljesül az

$$(*) \qquad A_0 < \frac{p_{1n}(\underline{X})}{p_{0n}(\underline{X})} = R_n < A_1 \ , \quad \underline{X} = (X_1, \ldots, X_n), \quad (n=1,2,\ldots).$$

Azaz ha a (*) feltétel teljesül n=1,2,..., N-re az N+1-edik megfigyelést is elvégezzük. Ha $R_N < A_0$ elfogadjuk a H_0 hipotézist, vagy ha $R_N > A_1$ elfogadjuk a H_1 hipotézist.

Az eljárást az α_0, elsőfaju hiba valószinüsége (elsőfaju hiba: elvetni H_0-at, amikor igaz), és α_1 a másodfaju hiba valószinüsége (másodfaju hiba: elvetni H_1-et, amikor igaz) függvényében vizsgáljuk. Ha N az eljárás befejezéséhez szükséges megfigyelésszámot jelenti, akkor nyilvánvaló, hogy ez valószinüségi változó.*

Az ilymódon leirt szekvenciális eljárás optimális tulajdonságu, amennyiben érvényes a következő

1. Tétel. Tekintsük azokat a kritériumokat, amelyekre teljesülnek

P_1 (H_0 elfogadása) $\leq \alpha_1$, P_0 (H_0 elvetése) $\leq \alpha_0$

és $E_0(N)$, $E_1(N) < \infty$ feltételek. A likelihood hányados próbán alapuló szekvenciális eljárás egyszerre minimalizálja $E_0(N)$ és $E_1(N)$-et. Azaz bármilyen más eljáráson alapuló $\tilde{N}$ megfigyelésszámra $E_0(\tilde{N}) \geq E_0(N)$, $E_1(\tilde{N}) \geq E_1(N)$. Ezen alapvető tétel bizonyitása előtt vizsgáljuk meg a szekvenciális eljárásokat.

2.2. A szekvenciális eljárások néhány tulajdonsága

Az alábbiakban a likelihood hányadospróbán alapuló szekvenciális eljárással foglalkozunk.

* Az (α_0, α_1) párt a szekvenciális eljárás erejének nevezzük.

(I) Ha 1 valószinüséggel befejeződik az eljárás, azaz

$$\sum_{n=0}^{\infty} P_i(N=n) = 1, \qquad (i=1,2)$$

és (α_0, α_1) a hiba valószinüségek az (A_0, A_1) határok esetén, akkor

$$A_0 \geq \frac{\alpha_1}{1-\alpha_0}, \qquad A_1 \leq \frac{1-\alpha_1}{\alpha_0}$$

Bizonyitás. Legyen

$$W_n = \left\{ \underline{x} : A_0 < R_k < A_1 \quad \text{ha } k_1 = 1, \dots, n-1 \quad \text{és } A_1 \leq R_n \right\},$$

akkor

$$\alpha_0 = P_0(H_1) = \sum_{n=1}^{\infty} \int_{W_n} p_{0n}(\underline{x})\, d\underline{x} \leq \sum_{n=1}^{\infty} \frac{1}{A_1} \int_{W_n} p_{1n}(\underline{x})\, d\underline{x} = \frac{1}{A_1}(1-\alpha_1).$$

Ha

$$U_n = \left\{ x : N=n,\ A_0 \geq R_n \right\},\quad 1-\alpha_0 = \sum_{n=1}^{\infty} \int_{U_n} p_{0n} \geq \frac{\alpha_1}{A_0}.$$

Közben felhasználtuk, hogy

$$\sum_{n=1}^{\infty} P_i(N=n) = \sum_{n=1}^{\infty} \int_{W_n \cup U_n} p_{in} dx = 1 \qquad , (i=1,0).$$

(II) Ha adott α_0, α_1 esetén a határokat

$$A_0' = \frac{\alpha_1}{1-\alpha_0}, \qquad A_1' = \frac{1-\alpha_1}{\alpha_0}$$

alakunak választjuk, akkor (feltéve, hogy az eljárás 1 valószinüséggel befejeződik) az ennek megfelelő eljárás (α_0', α_1') erejü lesz, amelyre (I) alapján teljesül

$$\frac{\alpha_1'}{1-\alpha_0'} \leq A_0' = \frac{\alpha_1}{1-\alpha_0}, \qquad \frac{1-\alpha_1}{\alpha_0} = A_1' \leq \frac{1-\alpha_1'}{\alpha_0'},$$

ahonnan

$$\alpha_0' \leq (1-\alpha_1')\frac{\alpha_0}{1-\alpha_1} \leq \frac{\alpha_0}{1-\alpha_1}, \qquad \alpha_1' \leq (1-\alpha_0')\frac{\alpha_1}{1-\alpha_0} = \frac{\alpha_1}{1-\alpha_0}.$$

[Ha α_0, α_1 a 0,01 - 0,1 intervallumba esnek, nincs lényeges eltérés a két erő között, α_i' nem lehet sokkal nagyobb α_i-nál, ugyanis

$$\alpha_1'(1-\alpha_0) \leq \alpha_1(1-\alpha_0') \qquad \alpha_0'(1-\alpha_1) \leq (1-\alpha_1')\alpha_0$$

$$\alpha_1' - \alpha_1'\alpha_0 + \alpha_0' - \alpha_0'\alpha_1 \leq \alpha_1 - \alpha_1\alpha_0' + \alpha_0 - \alpha_1'\alpha_0$$

egyenlőtlenségből adódik, hogy $\alpha_1' + \alpha_0' \leq \alpha_1 + \alpha_0$.]

Különbség abból adódik, hogy α_0', α_1' lényegesen kisebbek lehetnek α_0, α_1 -nál (ez pedig a megfigyelésszám növekedésével jár). Példaként a binomiális eloszlás esetét nézzük meg. n megfigyelés esetén a

p_0 valószinüségü és $p_1(> p_0)$ valószinüségü esemény hipotéziseket akarjuk megkülönböztetni. A likelihood hányados

$$\frac{p_{1n}}{p_{0n}} = \left(\frac{p_1 q_0}{p_0 q_1}\right)^{\sum_1^n x_i} \left(\frac{q_1}{q_0}\right)^n \quad \text{, ahol } q_i = 1 - p_i.$$

Ha $\log(p_1 p_0^{-1})/\log(q_0 q_1^{-1})$ racionális pontos képletek ismertek[1] az erőre és átlagos megfigyelésszámra. Igy szemléltetésül[2] pl.

$p_0 = 0{,}05$, $p_1 = 0{,}17$ esetén az erőt

$\alpha_0 = 0{,}05$, $\alpha_1 = 0{,}1$ értékeknek választva $\alpha_0' = 0{,}031$, $\alpha_1' = 0{,}099$

adódik és az átlagos megfigyelésszám $E_0'(N) = 31{,}4$, $E_1'(N) = 30{,}0$ lesz. Fix megfigyelésszám és (α_0, α_1) erő (hiba) esetén 57 megfigyelésre lenne szükség.

(III) Legyenek $X_1, X_2, \dots, X_n, \dots$ független, azonos eloszlásu valószinüségi változók, legyen $Z_i = \log(p_1(X_i)/p_0(X_i))$. Tetszőleges olyan H hipotézisre, amelyre $P_H(|Z(x)| > 0) > 0$ teljesül, igazak a következő állitások:

a) $P_H(N < \infty) = 1$,

b) $E_H(e^{tN}) < \infty$, (ahol $-\infty < t < t_0$, $t_0 > 0$).

Tehát a likelihood hányados próbán alapuló szekvenciális eljárás nemcsak 1 valószinüséggel befejeződik, hanem az N megfigyelésszám összes momentumai is léteznek.

Bizonyitás. Vezessük be a következő jelöléseket

$$S_i = Z_1 + \dots + Z_i, \quad b = \log A_0, \quad a = \log A_1.$$

Legyen adott m és $k(< m)$ esetén $r = \left[\frac{m}{k}\right]$ (egész rész) és tekintsük az S_k, $S_{2k} - S_k$, ..., $S_{rk} - S_{(r-1)k}$ valószinüségi változókat. $N > m$ esetén $S_i \in (b,a)$, $i=1,2,\dots,m$ és $i=k,2k,\dots,rk$-ra is. Igy $|T_i| = |S_{ik} - S_{(i-1)k}| < (|b| + |a|) = c$, $i=1,2,\dots,r$. Mivel a T_i változók függetlenek kapjuk, hogy

1) Girschick (1946) Annals of Math. Stat. 282-298.

2) Robinson: A note on exact sequential analysis. Univ.Calif.Publs.1 (1948) 241-246.

$$P\{N > m\} \leq P\{|T_i| < c, \quad i = 1, 2, \ldots, r\} = (P\{|T_i| < c\})^r .$$

Ahonnan azonnal adódik a) teljesülése. Közben felhasználtuk, hogy k megválasztható ugy, hogy $P\{|T_1| > c\} > 0$ teljesül és igy $P\{|T_1| < c\} = \delta < 1$. Feltevésünk szerint van olyan h, hogy vagy $P(z > h) > 0$, vagy $P(z < -h) > 0$. Legyen $h > \frac{c}{k}$, egyszerü meggondolásokból adódik, hogy

$$P\{|T_1| > c\} = P(|z_1 + \ldots + z_k| > c) \geq P(z_i > \tfrac{c}{k}, \; i = 1, \ldots, k) + P(z_i < -\tfrac{c}{k}, i = 1, \ldots, k) \geq$$

$$= [P(z > h)]^k + [P(z < -h)]^k > 0,$$

amit bizonyitani akartunk.

Az állitás második (b)) része a következő egyenlőtlenség sorozatból adódik:

$$E(e^{tN}) = \sum_m P\{N = m\} e^{tm} \leq \sum_m e^{tm} \delta^r = \sum_m e^{tm} \delta^{m/k} \delta^{(r - m/k)} \leq$$

$$\leq \delta^{-1/k} \sum_m e^{tm} \delta^{m/k} = \delta^{-1/k} \sum_m (e^t, \delta^{1/k})^m ,$$

ahol

$$r = \left[\frac{m-1}{k}\right], \qquad \delta = P(|T_n| < c)$$

és

$$e^t \delta^{1/k} < 1, \quad \left(\text{vagy} \quad t < -\frac{\log \delta}{k} = t_0\right).$$

A sor konvergenciája innen már következik.

(IV) Ha N pozitiv egész értékü valószinüségi változó, igaz az

$$E(N) = \sum_{m=1}^{\infty} P(N \geq m)$$

azonosság. Definició szerint $E(N) = P(N = 1) + 2P(N = 2) + \ldots = P(N = 1) +$

$$+ P(N = 2) + \ldots + (N = 2) + \ldots = P(N \geq 1) + P(N \geq 2) + \ldots$$

(V) Legyenek $X_1, X_2, \ldots$ függetlenek, azonos eloszlásuak és $z = z(x)$ az előbbi változó. N jelölje a döntéshez szükséges lépésszámot a szekvenciális eljárásban (a likelihood hányados alapján), mely nem függ a jövőtől. $E_H(|Z(X)|) < \infty$, $E_H(N) < \infty$ teljesülése esetén igaz az

$$E_H(S_N) = E_H(Z) \cdot E_H(N).$$

azonosság (Wald nevéről szokás Wald-féle azonosságnak nevezni). Mint az előbbiekben $S_N = z(X_1) + \dots + z(X_N)$.

Bizonyitás. Legyen

$$y_i = \begin{cases} 1 & \text{ha döntés (i-1)-ig nincs} \\ 0 & \text{ha van döntés (i-1)-ig} \end{cases}$$

$$P(y_i = 1) = P(N \geq i)$$

$$y_i = y_i(X_1, \dots, X_{i-1})$$

Legyen $Z_1 y_1 + \dots + Z_N y_N + Z_{N+1}\, y_{N+1} + \dots \; (= S_N)$, akkor

$$E(S_N) = E(\sum_i Z_i y_i) \sum_i E(Z_i y_i) = \sum_i E(Z_i) E(y_i) = E(Z) \sum_i E(y_i) =$$
$$E(Z) \sum_i P(N \geq i) = E(Z)\, E(N),$$

hogy megtehető az integrálás és összegezés felcserélése az következik a $\sum_i E(|Z_i\, y_i|) = \sum E(|Z|)\, P(N \geq i) = E(|Z|)\, E(N) < \infty$ egyenlőtlenségből.

A teljes valószinüség (várható érték) tétel alapján is belátható az állitás; ugyanis

$$E(S_N) = \sum_i E(S_N \mid N = i)\, P(N = i) = \sum_i i\, E(Z) . P(N = i)$$

2.3. A szekvenciális módszer hatásossága

(I) A szekvenciális eljárás 1 valószinüséggel befejeződik H_0 és H_1 hipotézis teljesülése esetén is.

Bizonyitás. Ha $z(x) \not\equiv 0$, akkor $E_0(z(X)) < 0$. Ugyanis az információ-elméletből ismert, hogy $E_0(Z) = \int p_0 \log p_1/p_0 = -\int p_0 \log \frac{p_0}{p_1}$ és az állitás a Jensen egyenlőtlenségből adódik. Innen $P_0(Z < 0) > 0$ és $P_0(|Z| > 0) > 0$, azaz teljesül az előző pont (III) állitásában szereplő feltétel. Ezzel állitásunk bizonyitása kész.

(II) Legyenek $E_0(|Z|) < \infty$, $E_1(|Z|) < \infty$, akkor az átlagos megfigyelésszámra a következő közelitések igazak:

$$E_0(N) \sim \frac{b(1-\alpha_0) + a\alpha_0}{E_0(z)}$$

1) A Jensen egyenlőtlenség szerint $(f, g \geq 0)$: $\int f \log \frac{1}{g}\, d\mu \geq 0$ ha $f > 0$ és $\int (f-g) \geq 0$, egyenlőség csak $f=g$ esetén áll fenn.

$$E_1(N) \sim \frac{b\alpha_1 + a(1-\alpha_1)}{E_1(z)}$$

(α_0, α_1) a próba ereje,
($b = \log A_0$, $a = \log A_1$)

Bizonyitás. Az $E(S_N | S_N \leq b) \sim b$, $E(S_N | S_N \geq a) \sim a$ összefüggéseket használva (elhanyagolva a szintek átugrásánál fellépő eltéréseket) adódik

$$E_0(S_N) = E_0(S_N | S_N \leq b) P_0(S_N \leq b) + E_0(S_N | S_N \geq a) P_0(S_N \geq a) \sim (1-\alpha_0) b + \alpha_0 \cdot a$$

Továbbá

$$E_0(N) E_0(Z) = E_0(S_N)$$

ahonnan adódik az állitás.

(III) Tetszőleges 1 valószinüséggel befejeződő, (α_0, α_1) erővel rendelkező szekvenciális eljárásra

$$E_0(N) \geq \frac{(1-\alpha_0)\log\frac{\alpha_1}{1-\alpha_0} + \alpha_0 \log\frac{1-\alpha_1}{\alpha_0}}{E_0(Z)}$$

Mielőtt a bizonyitásra térnénk vizsgáljuk meg mi adódik a likelihood próbán alapuló eljárásra.

Az előbbi (II) közelitést valamint az előző fejezet (I) közelitését felhasználva a likelihood próba esetén

$$E_0(N) \sim \frac{b(1-\alpha_0) + a\alpha_0}{E_0(Z)} \sim \frac{\left(\log\frac{\alpha_1}{1-\alpha_0}\right)(1-\alpha_0) + \alpha_0 \log\frac{1-\alpha_1}{\alpha_0}}{E_0(Z)}$$

közelités adódik. Igy ez a közelités is mutatja, de nem bizonyitja, hogy a likelihood hányados próbán alapuló eljárás a legjobb a véges várható értékü megfigyelésszám esetében.

Bizonyitás. Tetszőleges eljárásra

$$E_0(S_N) = E_0(S_N / H_0 \text{ elfogadva}) P_0(H_0) + E_0(S_N / H_0 \text{ elvetve}) P_0(H_0 \text{ elvetve}).$$

A Jensen egyenlőtlenség alapján (mely szerint ha $f(x)$ konkáv $E(f(\xi)) \leq f(E(\xi))$)

$$(*) E_0(S_N | H_0 \text{ elfogadva}) \leq \log E_0(e^{S_N} / H_0 \text{ elfogadva}) =$$

$$= \log\left(\frac{1}{1-\alpha_0} \sum_{m=1}^{\infty} \int_{S-W_m} e^{S_m} p_0(x_1) \dots p_0(x_m)\, d\underline{x}\right) =$$

$$= \log\left(\frac{1}{1-\alpha_0}\sum_{m=1}^{\infty}\int_{S-W_m} p_1(x_1)\dots p_1(x_m)\,d\underline{x}\right) = \log\frac{\alpha_1}{1-\alpha_0},$$

hasonlóan bizonyitható, hogy

$$(**)\quad E_0(S_N \mid H_0 \text{ elvetve}) < \log\frac{1-\alpha_1}{\alpha_0}.$$

Igy a $(*)$ és $(**)$ egyenlőtlenségekből kapjuk, hogy

$$E_0(S_N) \le (1-\alpha_0)\log\frac{\alpha_1}{1-\alpha_0} + \alpha_0\log\frac{1-\alpha_1}{\alpha_0}.$$

Mivel $E_0(z) < 0$,

adódik a bizonyitás.

Példa. Legyen $H_0 : \mathcal{N}(\theta_0, \sigma^2), H_1 : \mathcal{N}(\theta_1, \sigma^2)$ és a paraméterekről tegyük fel, hogy

$$\theta_1 > \theta_0, \quad \sigma^2 \quad \text{ismert.}$$

A likelihood hányados logaritmusa m megfigyelés esetén

$$\log R_m = \frac{\theta_1 - \theta_0}{2\sigma^2}\left[2\sum_1^m x_i - m(\theta_1 + \theta_0)\right]$$

A szekvenciális eljárás a következőkben áll:

1. A kisérlet folytatása, ha

$$\frac{b\,\sigma^2}{\theta_1 - \theta_0} < \sum_{i=1}^m X_i - \frac{m(\theta_1+\theta_0)}{2} < \frac{a\,\sigma^2}{\theta_1-\theta_0},$$

2. H_0 elfogadása, ha $\displaystyle\sum_{i=1}^m X_i - \frac{m(\theta_1+\theta_0)}{2} \le \frac{b\,\sigma^2}{\theta_1-\theta_0}$

3. H_1 elfogadása, ha $\displaystyle\sum_{i=1}^m X_i - \frac{m(\theta_1+\theta_0)}{2} \ge \frac{a\,\sigma^2}{\theta_1-\theta_0}$

ahol az a és b értékekre a következő közelitést használjuk:

$$a \sim \log\frac{1-\alpha_1}{\alpha_0}, \qquad b \sim \log\frac{\alpha_1}{1-\alpha_0}.$$

Az átlagos megfigyelésszám meghatározására az

$$E_0(N) \sim \frac{b(1-\alpha_0) + \alpha_0 a}{E_0(Z)}$$

közelitést használjuk. Mivel

$$Z = -\frac{(X-\theta_1)^2-(X-\theta_0)^2}{2\sigma^2}, \qquad E_0(Z) = \frac{(\theta_1-\theta_0)^2}{2\sigma^2}$$

$$E_0(N) \sim -\frac{b(1-\alpha_0)+\alpha_0 a}{(\theta_1-\theta_0)^2/2\sigma^2}$$

Hasonlóan kapjuk az

$$E_1(N) \sim -\frac{b\alpha_1+a(1-\alpha_1)}{(\theta_1-\theta_0)^2/2\sigma^2}$$

összefüggést.

Fix elemszámu minta, és (α_0, α_1) erő esetén a kritikus tartományt

$$\sqrt{N}(\bar{x}-\theta_0) \geq d_{\alpha_0}\cdot\sigma$$

adja meg, ahol N a megfigyelésszám, d_{α_0} az $1-\alpha_0$ kvantilise $\mathcal{N}(0,1)$-nek.

A másodfaju hiba valószinüsége

$$\alpha_1 = P_{\theta_1}(\sqrt{N}(\bar{x}-\theta_0) < d_{\alpha_0}\sigma) = P_{\theta_1}\{\sqrt{N}(\bar{x}-\theta_1) < d_{\alpha_0}\sigma - \sqrt{N}(\theta_1-\theta_0)\}$$

lesz. Ha d_{α_1} az $1-\alpha_1$ kvantilise $\mathcal{N}(0,1)$-nek

$$d_{\alpha_1}\cdot\sigma = -d_{\alpha_0}\cdot\sigma + \sqrt{N}(\theta_1-\theta_0),$$

$$N = \frac{\sigma^2(d_{\alpha_1}+d_{\alpha_0})^2}{(\theta_1-\theta_0)^2}.$$

A fix megfigyelésszámu N, és a szekvenciális eljáráshoz tartozó $E_0(N)$ hányadosa

(*) $$-\frac{2[b(1-\alpha_0)+a\alpha_0]}{(d_{\alpha_0}+d_{\alpha_1})^2}$$

független $(\theta_1-\theta_0)$-tól ; $E_1(N)$-el összehasonlitva a fix megfigyelésszámot

(* *) $$\frac{2[b\alpha_1+a(1-\alpha_1)]}{(d_{\alpha_0}+d_{\alpha_1})^2}$$

adódik. Ha pl. $\alpha_0=\alpha_1=0.05$, $d_{\alpha_0}=d_{\alpha_1}=1,64$ és jó közelitéssel $a = 2,94$; $b = -2,94$.

A (*) és (* *) hányadosokra 0.49 adódik.

A valóságban a hányadosok függnek $\theta_1-\theta_0$-tól. Chernoff vizsgálta[1)] kis α_0, α_1 -ekre a megfigyelésszámok arányát, s nagy megfigyelésszámra 4:1 arányt talált.

---

1) Chernoff, Annals Math. Stat. (1959) 3o, 755-77o.

2.4. A szekvenciális eljárás alapvető azonossága

1. Lemma. Legyen Z olyan valószinüségi változó, hogy

(a) $P(Z>0)>0$ és $P(Z<0)>0$,

(b) $\varphi(t)=E(e^{tZ})$ minden valós t-re létezik,

(c) $E(Z)\neq 0$.

Akkor van olyan $\tau\neq 0$, hogy $\varphi(\tau)=1$. Ha $E(Z)>0$, $\tau<0$ és $E(Z)<0$ esetén $\tau>0$.

Bizonyitás. A feltételekből következik olyan c létezése, hogy

$$P(Z>c)=\delta>0.$$ Ezért

$$\varphi(t)=E(e^{tZ})=\int e^{tz}\,dF\geq\int_{z>c} e^{tz}\,dF>e^{tc}P(Z>c),\quad \text{ha}\quad t>0,$$

és ha $t\to\infty$ adódik, hogy $\varphi(t)\to\infty$.
Ugyanugy, ha

$t\to-\infty$, $\varphi(t)\to\infty$. [Ugyanis $E(e^{tZ})=\int e^{tz}dF\geq\int_{z<-c} e^{tz}dF\geq e^{-tc}\cdot P(Z<-c)$]

Mivel $\varphi(0)=1$ és $\varphi'(t)=E(Ze^{tZ})$ ez pedig $t=0$-ban $E(Z)$, a $\varphi'(0)$ értéke. Továbbá $\varphi''(t)=E(Z^2e^{tZ})>0$ minden t-re, igy egy minimuma lehet a $\varphi(t)$ függvénynek és a t=0 értéken kivül egyetlen τ helyen lehet $\varphi(\tau)=1$. Ha $E(Z)>0$ $\tau<0$ kell legyen, és megforditva.

2. Lemma. Legyen X_1, X_2, ..., azonos eloszlásu független valószinüségi változók sorozata p_H sürüséggel. Tekintsünk egy tetszőleges szekvenciális eljárást adott megállási szabállyal. Legyen n a szükséges megfigyelések száma. Ha $z_i = z(x_i)$ tetszőleges mérhető függvény, akkor a $P_H(n<\infty)=1$ összefüggésből adódik, hogy

$$E_H\left(e^{tS_n}[\varphi(t)]^{-n}\right)=P_{H_t}(n<\infty),$$

ahol a H_t hipotézis esetén a sürüség alakja

$$p_{H_t}(x)=\frac{e^{tz(x)}}{\varphi(t)}\cdot p_H(x),$$

$$S_n = Z_1 + \dots + Z_n, \quad \varphi(t) = E_H(e^{tZ}).$$

Bizonyitás. Legyen $W_m + U_m$ az m- dimenziós tér azon tartománya, melyben az eljárás pontosan m lépésben befejeződik. Akkor

$$E_H(e^{tS_n}[\varphi(t)]^{-n}) = \sum_{m=1}^{\infty} \int_{W_m+U_m} [\varphi(t)]^{-m} e^{tS_m} p_H(x_1)\dots p_H(x_m)dx_1\dots dx_m =$$

$$= \sum_{1}^{\infty} \int_{W_m} p_{H_t}(x_1)\dots p_{H_t}(x_m)\, dx_1\dots dx_m = P_{H_t} \qquad (h < \infty)$$

3. Lemma. (Alapazonosság) Legyen a likelihood hányados alapján működő szekvenciális eljáráshoz szükséges megfigyelésszám (valószinüségi változó) N és $z = \log \frac{p_1(x)}{p_0(x)}$. Akkor

$$E_H\{e^{tS_N}[\varphi(t)]^{-N}\} = 1, \quad \text{ha} \quad P_H\{|Z| > 0\} > 0.$$

Bizonyitás. $P_H(|Z| > 0) > 0 = P_{H_t}(|Z| > 0)$ és a 2. lemmából adódik a tétel, mivel $P_{H_t}(N < \infty) = 1$.

A terv operativ karakterisztikája (jelleggörbe, vagy OC görbe elnevezés is használatos).

A feladat abban áll, hogy a szekvenciális terv alapján elfogadunk egy szállitmányt vagy sem. Jelölje $\pi(H)$ az elfogadás valószinüségét a H hipotézis esetén.

Egyszerü alternativa esetén

H_0 elfogadása valószinüsége $1 - \alpha_0$ (ha H_0 igaz),

H_0 " " α_1 (ha H_1 igaz).

Ha szekvenciális likelihood próbával történik a szállitmány elfogadása a $\pi(H)$ függvényt, mint H függvényét az elfogadási terv operativ karakterisztikájának nevezzük. H_0, H_1 hipotézis és α_0, α_1 első-másodfaju hiba feltételezésével az előbbi pont eredményei alapján jó közelitések adhatók (a szállitmány elfogadása H_0 elfogadását jelenti).

Az 1. lemma szerint van olyan τ, hogy a karakterisztikus függvény $\varphi(\tau) = 1$. Innen és az alapazonosságból kapjuk, hogy

$$1 = E_H(e^{\tau S_N}) = P_H\{S_N \geq a\} E_H\{e^{\tau S_N} | S_N \geq a\} +$$

$$+ P_H\{S_N \leq b\} E_H\{e^{\tau S_N} | S_N \leq b\} \approx [1 - \pi(H)] e^{\tau a} + \pi(H) e^{\tau b}$$

ahonnan

$$\pi(H) \sim \frac{1-e^{\tau a}}{e^{\tau b}-e^{\tau a}}$$

Az átlagos megfigyelésszám (Average Sample Number, ASN) meghatározható az előbbi közelités valamint a 2.3. pont (II) közelitése alapján. Ugyanis

$$E_H(S_N) \sim b\pi(H) + [1-\pi(H)]\cdot a$$

és

$$E_H(N)\cdot E_H(Z) = E_H(S_N)$$

ahonnan

$$E_H(N) \sim \frac{b\pi(H) + a(1-\pi(H))}{E_H(z)}$$

Példa. Normális eloszlás feltételezése esetén (a 2.3. pont példája jelöléseivel)

$$Z = \frac{1}{2\sigma^2}\left[(X-\theta_1)^2-(X-\theta_0)^2\right]$$

$$\varphi(t) = E_H(e^{tZ}) = e^{\frac{q}{2\sigma^2}}, \quad q = t(\theta_1-\theta_0)[2\theta-\theta_1-\theta_0+t(\theta_1-\theta_0)]$$

Az a τ érték, melyre $\varphi(\tau)=1$ a következő

$$\tau(\theta) = \frac{\theta_1+\theta_0-2\theta}{\theta_1-\theta_0}.$$

Az operativ karakterisztika

$$\pi(\theta) = \frac{1-e^{a\tau(\theta)}}{e^{b\tau(\theta)}-e^{a\tau(\theta)}}$$

lesz. Mig az átlagos megfigyelésszám

$$E_\theta(N) = \frac{1-e^{b\tau(\theta)}}{E_\theta(z)}, \text{ ahol } E_\theta(Z) = \frac{(\theta-\theta_0)^2-(\theta-\theta_1)^2}{2\sigma^2}.$$

2.5. Időben folytonos folyamat vizsgálata és kapcsolata a diszkrét folyamattal.

Az előbbi pontok tételei általánositása Wiener (Brown-mozgás) folyamatra a következő.

1. Lemma. Ha w_t jelöli a standard Wiener folyamatot $Ew_t = 0$, $Ew_t^2 = t$ paraméterekkel, $w_0 = 0$ és $\tau = \tau(w)$ egy Markov pont, amelyre $E(\tau)<\infty$, akkor

$$Ew_\tau = 0, \quad Ew_\tau^2 = E\tau$$

(a bizonyitás megtalálható például A. Sirjajev:"Statisztikai szekvenciális analizis" 183.o., vagy I. Gihman - V. Szkorohod: "Bevezetés a sztohasztikus folyamatok elméletébe" c. könyv 493.o.).

Ez a lemma Wald lemmájának általánositása (2.2. pont V. állitás), előbb az $S_n \sim w_t$ majd $S_n \sim w_t^2$ helyettesitéssel

$$(Ew_\tau = Ew_1 \cdot E\tau = 0 \quad \text{és} \quad Ew_\tau^2 = Ew_1^2 \, E\tau = E\tau).$$

2. Lemma. Ha w_t a standard Wiener folyamat és $\tau = \tau(\omega)$ egy Markov pont, melyre teljesül a $P(\tau \le K) = 1$, ahol $K < \infty$, feltétel, akkor minden $-\infty < \lambda < \infty$ számra teljesül a következő összefüggés

$$E \exp\left(\lambda w_\tau - \frac{\lambda^2}{2}\tau\right) = 1.$$

(a bizonyitás megtalálható Sirjajev emlitett könyve 184.o.-on, kell hozzá Doob: "Stochastic Process" c. könyv VII. fejezet 11.7. tétel is.).

Ez a lemma a szekvenciális analizis alapvető azonosságának általánositása folyamatra, korlátos Markov pont esetén (v.ö. 2.4.pont 3. lemma). Felhasználva, hogy $E(e^{\lambda w_s}) = \varphi(\lambda) = e^{\frac{1}{2}\lambda^2 s}$.

Ez utóbbi lemma általánositása a Novikov féle tétel.

Tétel. Ha $f(t(\omega))$ olyan függvény (véletlen) amelyre értelmezhető a sztohasztikus integrál és teljesül a következő feltétel

$$E \exp\left\{\tfrac{1}{2}\int_0^T f^2(t,\omega)\,dt\right\} < \infty$$

akkor

$$E \exp\left\{\int_0^T f(t,\omega)\,dw(t) - \frac{1}{2}\int_0^T f^2(t,\omega)\,dt\right\} = 1.$$

(lásd A. Novikov "Teorija verojatnosztyej"(1972) 761-765.)

Az, hogy a 2. lemma a tétel következménye, belátható az $f(t,\omega) = \lambda\chi\,(t < \tau)$ választással.

Ezeknek a tételeknek a bizonyitására a sztochasztikus folyamatok statisztikai vizsgálatánál visszatérünk.

Feladatok

1. Vizsgáljunk független megfigyelési sorozatban egy p valószinüségü eseményt. m-szeri előfordulásig folytassuk a kisérletet. A kisérletek N száma valószinüségi változó,

$$P(N=n) = \binom{n-1}{m-1} p^m q^{n-m} \qquad n = m,\ m+1, \dots$$

eloszlással. Bizonyitsuk be, hogy 1-valószinüséggel befejeződik az eljárás. Mutassuk meg, hogy az eloszlás generátor függvénye

$$\left(\frac{pt}{1-qt} \right)^m$$

és a kumuláns generátor függvénye

$$\psi(t) = \log\left(\frac{pe^t}{1-qe^t}\right)^m = m \log\left(\frac{p}{e^{-t}-q}\right)$$

alaku, ahonnan sorfejtéssel

$$\varkappa_1 = \frac{m}{p},$$

$$\varkappa_2 = \frac{mq}{p^2}$$

adódik.

Azaz N várható értéke $\frac{m}{p}$. Ez nem jelenti, hogy $\frac{m}{N}$ torzitatlan becslése p-nek. $\frac{m-1}{N-1} = \xi$ lesz torzitatlan becslés.

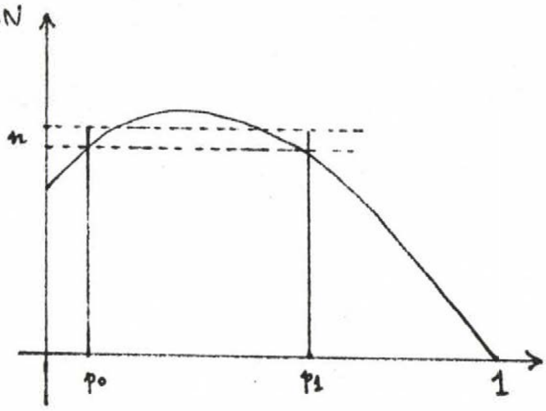
$$E\left(\frac{m-1}{N-1}\right) = \sum_{n=m}^{\infty} \left(\frac{m-1}{n-1}\right)\binom{n-1}{m-1} p^m (1-p)^{n-m} =$$

$$= p \sum_{m}^{\infty} \binom{n-2}{m-2} p^{m-1} q^{n-m} = p.$$

A szórásnégyzetet bonyolult meghatározni, de

$$E \frac{(m-1)(m-2)}{(N-1)(N-2)} = p^2$$

2. Legyen a selejtarány p_0 H_0 hipotézis esetén, mig H_1 esetén p_1. Nevezzük (n,c) tervnek a következőt: ha n darabot megvizsgálva a selejtek száma nagyobb c-nél, a tételt ($N \gg n$ darabszámu) elvetjük. Rajzoljuk meg az OC-görbét (Operating Characteristic) és adott (α_0, α_1) erő esetén n és c értékét.
Például $p_0 = 0{,}02$, $p_1 = 0{,}08$ és $\alpha_0 = 0{,}05$, $\alpha_1 = 0{,}1$ esetén $n = 100$, $c = 4$ adódik.
Alakitsuk át kétfokozatu tervvé (a Dodge-Romig eljárás alapján) a fenti feladatot (n_1, n_2, c_{12}, c_{1o}, c_{2o}) paraméterekkel. Táblázat található H. Dodge - H. Romig: Sampling inspection tables (Wiley, 1949, New York) c. könyvben. Ez utóbbi esetben határozzuk meg az ASN (átlagos megfigyelésszám) görbét. Mutassuk meg, hogy annak alakja az ábrán látható módon változik.



Például a $p_0=0,03$, $p_1=0,107$ és $\alpha_0= 0,05$, $\alpha_1=0,1$ értékekre egy megoldás $n_1=75$, $n_2=150$, $c_{10}=4$, $c_{11}=9$, $c_{20}=8$.

Határozzuk meg a fenti (p_0, p_1) és (α_0, α_1) értékekre szekvenciális eljárás esetén A_0, A_1 és E_0N, E_1N értékeket.

3. A tönkremenés problémája. Legyen A-nak (a) egységnyi pénze és B-nek (b) egységnyi pénze. A nyerési valószinüsége legyen p, B-é pedig q = 1-p. A játék mindaddig folytatódik, mig valamelyik eléri az a+b összeget.

Legyen u_x az A tönkremenetele valószinüsége, ha x összege van. Nyilván

$$u_x = pu_{x+1} + qu_{x-1}$$

az

$$u_0 = 1, \quad u_{a+b} = 0$$

feltételekkel. Ennek megoldása

$$u_x = c_1 t_1^{*} + c_2 t_2^{*}$$

ahol t_1, t_2 a

$$pt^2 - t + \varkappa = 0$$

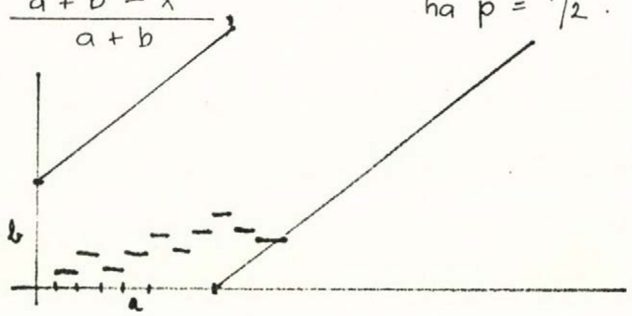
egyenlet gyökei ($t_1=1$, $t_2=q/p$). Feltéve, hogy $p \neq q$

$$u_x = \frac{\left(\frac{q}{p}\right)^{a+b} - \left(\frac{q}{p}\right)^{x}}{\left(\frac{q}{p}\right)^{a+b} - 1}, \qquad p \neq 1/2$$

és

$$u_x = \frac{a+b-x}{a+b}, \qquad \text{ha } p = 1/2 .$$

A játék ábrázolása



4. Mutassuk meg, hogy a $\frac{p_{1n}}{p_{0n}}(\xi) = \xi_n$ sorozat martingált alkot a P_0 mértékre nézve.

5. Poisson eloszlás esetén vizsgáljuk a likelihood hányadoson alapuló szekvenciális eljárást.

6. Egyszerü alternativa esetén az esemény valószinüsége a H_0 hipotézis esetén legyen p_0 (H_1 esetén $p_1 > p_0$). A likelihood hányados próbán alapuló szekvenciális eljárásról mutassuk meg, hogy m kisérlet esetén (m_1 darab sikeres kisérlet mellett) a következő:

   a) H_0 elfogadása, ha $m_1 \leq c.m + d_1$

   b) H_1 elfogadása, ha $m_1 \geq c.m + d_2$

   c) ujabb kisérlet egyébként.

   Adjuk meg az O.C. függvényt és az átlagos mintaszámot.

7. Legyen X exponenciális eloszlásu $\lambda e^{-\lambda x}$ sürüséggel. A λ_0 és λ_1 hipotézisek eldöntése közötti szekvenciális próba a

$$k_1 + (\lambda_1 - \lambda_0)\sum_{j=1}^{n} x_j \leq n \log(\lambda_1/\lambda_0) \leq k_2 + (\lambda_1 - \lambda_0)\sum_{j=1}^{n} x_i$$

egyenlőtlenség párban áll, ahol k_1 és k_2 konstansok.

8. Mutassuk meg, hogy $b = -\infty$ (vagy $a = +\infty$) esetén $EN = \infty$.

9. Mutassuk meg, hogy a momentum generáló függvény létezése esetén (a $t=0$ egy környezetében) az összes momentumok végesek.

3. Statisztikai döntések elmélete

3.1. Döntési eljárások

Minden x megfigyeléshez $\delta(x) = d$ értéket rendelünk. $x \in X$, $d \in D$. Ha az X valószinüségi változó valódi eloszlása P_θ a d döntés vesztesége legyen $L(\theta, d)$. Ismétlések hosszu sora folyamán a δ alkalmazásával adódó átlagos veszteség

$$E_\theta(L(\theta,\delta)) = R(\theta,\delta)$$

(ezt a függvényt rizikófüggvénynek is szokás nevezni).
A feladat olyan δ döntési eljárás keresése, melyre a rizikófüggvény minimális. A három elem P_θ, $\theta \in \Theta$, $d \in D$ és L alapvető szerepet játszik a statisztikai döntések elméletében. Természetesen a veszteségfüggvény megadása a feladatok milyenségétől függ

<u>1. Példa.</u> Legyen a $\xi_1, \ldots, \xi_n$ sorozat adott eloszlásu (pl. binomiális, Poisson vagy normális p, λ, (m, σ) paraméterekkel, ahol az utóbbi paraméterpár egyike a θ). Legyen $\gamma = \gamma(\theta)$ egy valós függvény.

(I) Megvizsgálandó γ nagyobb-e γ_0 -nál és a döntés álljon a következőkben

$$d_0 : \gamma > \gamma_0, \qquad d_1 : \gamma \leqq \gamma_0$$

(pl. zaj nagyobb-e, növekedés gyorsabb-e, orvosság hatása elég nagy-e egy megadott szintnél). A veszteségfüggvény a felhasználási területtől függ. Egységes azonban a szokás abban, hogy helyes döntés esetén a veszteség 0, helytelen esetén $L(\gamma, d_0)$ és $L(\gamma, d_1)$ a $|\gamma - \gamma_0|$ növekvő függvényei.

(II) Ha γ becslése a feladat - a probléma bonyolultabb. Ekkor d valós szám (γ becslése) és rendszerint $L(\gamma, d) = v(\gamma)w(|d-\gamma|)$, ahol w egy szigoruan növekvő függvény.

(III) Lehet, hogy

$$d_0 : \gamma < \gamma_0, \qquad d_1 : \gamma > \gamma_1, \qquad d_2 : \gamma_0 \leqq \gamma \leqq \gamma_1$$

Az (I) leirás a hipotézisvizsgálatnak, (II) a pontbecslésnek, mig (III) több lehetőségü döntési eljárások egy speciális esetének felel meg.

Sok esetben nem egyetlen megoldás van a problémára (azaz minden θ-hoz nem egyetlen olyan d létezik, melyre $L(\theta, d) = 0$). Ilyenkor a következtetés a fontos s nem a döntés. Ezt is példával illusztráljuk.

2. Példa. Legyen ξ egy $\mathcal{N}(m, \sigma)$ változó, ahol m-re $[\underline{L}(x), \overline{L}(x)]$ konfiderancia intervallum szerkesztendő. Az intervallum hossza legfeljebb $k\sigma$, ahol k konstans. A veszteség 0, ha helyes a döntés (lefedi az intervallum m-et), egyébként a veszteség az m-től való távolság lehet. Döntési eljárással megfogalmazva több megoldása is van a feladatnak.

A veszteségfüggvény megválasztása a legnehezebb. A legegyszerűbb esetben (amikor a veszteség pénzzel mérhető) is nehéz. Az (I) példában

$$L(\theta, d_0) = a, \quad \text{ha} \quad \gamma(\theta) \leq \gamma_0 \quad (d_0 : \gamma > \gamma_0),$$

$$L(\theta, d_1) = b, \quad \text{ha} \quad \gamma(\theta) > \gamma_0 \quad (d_1 : \gamma \leq \gamma_0).$$

Ekkor a rizikófüggvény

$$R(\theta, \delta) = \begin{cases} a \cdot P_\theta\{\delta(x) = d_0\} & , \text{ha} \quad \gamma \leq \gamma_0 \\ b \cdot P_\theta\{\delta(x) = d_1\} & , \text{ha} \quad \gamma > \gamma_0 \end{cases}$$

Általában nem megy ilyen könnyü egyszerüsitéssel a veszteség megadása. Egyszerü függvényeket keresnek, ahol a veszteséget

$$L(\theta, d) = v(\theta)(d - \gamma)^2$$

szokás megadni (ez sokszor jó közelités). Van ugy, hogy egyetlen L_1 komponenssel a veszteség nem jellemezhető. Konfidencia intervallum megadása esetén L_1 az intervallum hosszát, L_2 a veszteséget jelentse, ha nem fedi le az intervallum a valódi paraméter értéket.

Ki lehet kötni pl., hogy teljesülnie kell a

$$EL_1(\theta, \delta(x)) \leq \alpha$$

feltételnek, és a másik komponenst minimalizálni kell.

Randomizálás (A kisérlet megválasztása.)

A döntési eljárás fenti meghatározása elég szük. Minden x értéknek egyetlen döntés felel meg. Helyesebb valamilyen eloszlás szerint választani a döntésekből (ahol az eloszlás x-től függ).

A rizikófüggvények osztálya konvex lesz (ha a randomizáltakat is nézzük) és ez több előnnyel jár.

Az eddigiekben feltettük, hogy a kisérlet meg volt tervezve. Valójában ez nincs igy, hiszen gyakran előre nem tudjuk, hogy hány megfigyelésre lesz szükség.

3. Példa. 1. (m_1, σ) és (m_2, σ) megkülönböztetése a feladat. A megoldás σ-tól függ. Néhány elemből σ-t meg kell becsülni és megadni a megfigyelésszámot.

2. Eldöntendő egyszerü alternativára vonatkozó kisérletsorozatnál, hogy $p > \frac{1}{2}$ vagy $p \leq \frac{1}{2}$. A kisérletszám attól függ, mi a p.

3. Nemcsak az eloszlások lehetnek különbözőek (információs szempontból) hanem a megfigyelések is.

$[\theta - \frac{1}{2}, \theta + \frac{1}{2}]$-ben egyenletes változók esetén

θ jól becsülhető, ha $\max \xi_i - \min \xi_j \sim 1$. (Ha kicsi, egyetlen megfigyelésnél nem jelent többet.)

Optimális eljárások.

Már emlitettük, hogy $R(\theta,\delta) = E_\theta[L(\theta, \delta(x)]$ minimalizálása a feladat, azaz olyan δ keresése, melyre a rizikó minimális. A megoldás függ θ-tól. (Pl. ha θ_0-ra $\delta(x) = d_0$ olyan, hogy $L(\theta_0, d_0) = 0$, akkor $R(\theta_0, d_0) = 0$, viszont másutt nagy lesz a rizikó értéke).

Van-e egyenletesen legkisebb rizikó ?

Bizonyos megszoritások esetén - igen (pl. a torzitatlanság, invariánsság esetében).

Másik ut a legjobb eljárások esetére az eljárások rendezésében áll.

Bayes és minimax eljárások.

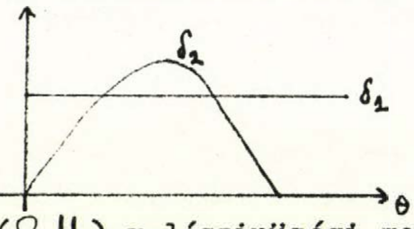
Ha θ valószinüségi változó $\varrho(\theta)$ sürüséggel, akkor a teljes átlagos veszteség δ alkalmazása esetén

$$r(\varrho,\delta) = \int R(\theta,\delta)\,\varrho(\theta)\,d\theta = \int E_\theta L(\theta, \delta(x))\,\varrho(\theta)\,d\theta .$$

Optimális eljárást az a δ szolgáltat, melyre $r(\varrho,\delta)$ minimális. Ez az u.n. Bayes féle megoldása a feladatnak, mely a $\varrho(\theta)$ apriori eloszlásnak felel meg. $\min_\delta r(\varrho,\delta)$ a Bayes féle rizikó.

Ismerni kell $\varrho(\theta)$-át ! Ez azonban nem mindig teljesül. Ha nem ismert a $\varrho(\theta)$ választható a következő eljárás: meg kell nézni adott δ

esetén $R(\theta,\delta)$ maximumát és két döntés esetén az a jobbik eljárás, melyre a maximum kisebb. Ez az u.n. minimax eljárás.



Legyen az $(\Omega, \mathcal{U})$ valószinüségi mezőn megadva a $P_\pi (= \pi P_0 + (1-\pi) P_1)$ mértékek összessége $(0 \leqq \pi \leqq 1)$. $\theta = \theta(\omega)$ valószinüségi változó, mely az 1 és 0 értékeket veszi fel. $P_\pi(\theta(\omega) = 0) = \pi$, $P_\pi(\theta(\omega) = 1) = 1 - \pi$.

A $\xi_1, \xi_2, \dots$ valószinüségi változók vagy P_0 vagy P_1 eloszlásuak. $\mathcal{F}_n = \sigma\{(\omega,\pi) : \pi, \xi_0, \xi_1, \dots, \xi_n\}$ ahol $\pi \in [0,1]$, $\xi_0 = 0$, $n \leqq N$,
$\mathfrak{M} = \{\tau\}$,
a Markov pontok halmaza $F = \{\mathcal{F}_n\}$ -re nézve; $D_\tau = \{d\}$ az $\mathcal{F}_\tau$ mérhető 0,1 értéket felvevő $d = d(\omega,\pi)$ függvények osztálya.

Ha $\tau \in \mathfrak{M}$, $d \in D_\tau$ a $\delta = (\tau, d)$ pár <u>döntési szabály</u>. $\Delta = \{\delta\}$ a döntési szabályok osztálya.

3.2. A szekvenciális eljárás optimum tulajdonságairól

Jelölje H_0 és H_1 a két egyszerü hipotézist

H_0: p_0, és a veszteség H_0 elvetése esetén (amikor igaz), legyen w_0;

H_1: p_1, a veszteség H_1 elvetése esetén (amikor igaz), legyen w_1.

A költség legyen egy megfigyelésre c.

A δ döntési eljárás esetén

α_0 az elsőfaju hiba valószinüsége $(P_0(\delta(w,\pi) = 1) = \alpha_0)$.

α_1 a másodfaju hiba valószinüsége $(P_1(\delta(w,\pi) = 0) = \alpha_1)$.

A szekvenciális eljárás rizikója (N megfigyelésszám) p_i valódi sürüség esetén

$$\alpha_i w_i + cE_i(N) \qquad (i = 0,1)$$

A hipotézisek indexeit valószinüségi változónak tekintjük, amely 0,1-et π, $(1-\pi)$ valószinüséggel veszi fel. Jelölje ezt a változót θ, $P(\theta = 0) = \pi$. A Bayes féle rizikó (teljes átlagos rizikó) a következő

$$(*) \qquad r(\pi,\delta) = \pi[\alpha_0 w_0 + c E_0(N)] + (1-\pi)[\alpha_1 w_1 + c E_1(N)].$$

Keressük a Bayes féle megoldását a feladatnak, azt a δ^* eljárást, melyre (*) minimális. Ezt a feladatot nevezzük a szekvenciális eljáráshoz tartozó kiegészitő feladatnak.

Legyen δ_0 a megfigyelés nélküli eljárás H_1 elfogadásával. Nyilván $r(\pi,\delta_0) = \alpha_0 w_0 \pi = w_0 \pi$.

Legyen δ_1 a megfigyelés nélküli eljárás H_0 elfogadásával. Nyilván $r(\pi,\delta_0) = w_1(1-\pi)$.

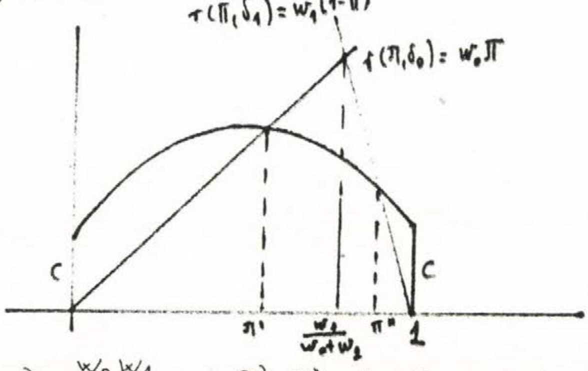
Legyen $\rho(\pi) = \inf r(\pi,\delta)$, $\delta \in \mathcal{C}$,

$\mathcal{C}$ jelöli a legalább 1 megfigyelést igénylő eljárások osztályát. Tetszőleges π_0, π_1 és $0<\lambda<1$ értékekre

$$\rho(\lambda\pi_0 + (1-\lambda)\pi_1) = \inf_{\delta\in\mathcal{C}} r[(\lambda\pi_0 + (1-\lambda)\pi_1),\delta] = \inf_{\delta\in\mathcal{C}}[\lambda r(\pi_0,\delta) + (1-\lambda) r(\pi_1,\delta)] \geq$$

$$\geq \lambda\rho(\pi_0) + (1-\lambda)\rho(\pi_1),$$

azaz a $\rho(\pi)$ függvény konkáv és mivel alulról korlátos $(\rho(\pi) \geq c)$, egyben folytonos.



Ha $\rho\left(\frac{w_1}{w_0+w_1}\right) < \frac{w_0 w_1}{w_0+w_1}$, a π', π'' értékek definiciója legyen a következő:

$$\rho(\pi',\delta_0) = \rho(\pi'), \qquad \rho(\pi'',\delta_1) = \rho(\pi'').$$

Ellenkező esetben legyen

$$\pi' = \pi'' = \frac{w_1}{w_0+w_1}.$$

1. Tétel. Legyen π' és π'' a fenti. Ha $0<\pi'<\pi''<1$, akkor minden $\pi'<\pi<\pi''$ értékre a Bayes féle rizikó minimalizálásának feltétele az

$$A_0 = \frac{\pi}{1-\pi}\cdot\frac{1-\pi''}{\pi''}, \qquad A_1 = \frac{\pi}{1-\pi}\,\frac{1-\pi'}{\pi'}$$

határokkal meghatározott szekvenciális likelihood hányadospróba használata.

Bizonyitás

1. Kell-e végezni megfigyelést ?

   A δ_0 eljárás (*)-ot akkor és csak akkor minimalizálja, ha $\pi < \pi'$ és H_0-át elvetjük.
   A δ_1 eljárás (*)-ot akkor és csak akkor minimalizálja, ha $\pi > \pi''$ és H_1-et elvetjük (H_0-át elfogadjuk).

2. A bizonyitás további része indukcióval történik.
   Legyen $\pi' < \pi < \pi''$, és n megfigyelésünket jelölje $X_1, X_2, \ldots, X_n$.
   Az, hogy n.c költségünk van, nincs hatással a feladatra, semmilyen jövőbeni esemény nem teszi kisebbé ezt a költséget.
   Vagy abbahagyjuk a kisérletet w_0 ill. w_1 veszteséggel, vagy végzünk egy ujabb X_{n+1} kisérletet. Az aposteriori valószinüségek alapján a döntés: ha $P\{\theta=\theta | X_1, \ldots, X_n\} < \pi'$, vagy $P\{\theta=0 / X_1, \ldots, X_n\} > \pi''$ nincs további kisérlet, mig $\pi' < P\{\theta=0 / X_1, \ldots X_n\} < \pi''$ esetén elvégezzük az n+1 kisérletet is.
   A Bayes féle tétel szerint a

$$P(\theta=0 / X_1, \ldots X_n) = \frac{\pi p_{0n}}{\pi p_{0n} + (1-\pi) p_{1n}} = \pi(X_1, \ldots, X_n)$$

   aposteriori valószinüségre a

$$\pi' < \pi(X_1, \ldots, X_n) < \pi''$$

   feltételekből adódik, hogy

$$\pi'(\pi p_{0n} + (1-\pi) \cdot p_{1n}) > \pi p_{0n},$$

$$\pi(\pi'-1) p_{0n} + \pi'(1-\pi) p_{1n} > 0,$$

$$\pi'(1-\pi) p_{1n} > \pi(1-\pi') p_{0n},$$

$$\frac{p_{1n}}{p_{0n}} < \frac{\pi}{1-\pi} \qquad \frac{1-\pi''}{\pi''},$$

teljesülése esetén elvetjük a H_0 hipotézist, mig

$$\frac{p_{1n}}{p_{0n}} < \frac{\pi}{1-\pi} \qquad \frac{1-\pi''}{\pi''}$$

esetén elfogadjuk a H_0 hipotézist.

Bayes tétele: ha B_i teljes esemény rendszer

$$P(B_k|A) = \frac{P(A|B_k)\,P(B_k)}{\sum_i P(A|B_i)\,P(B_i)}.$$

Sürüségfüggvényekre a Bayes tétel a

$$g(y|x) = \frac{f(x|y)\,g(y)}{\int f(x|t)\,g(t)\,dt}$$

alaku, vagy irhatjuk

$$P(a \le \eta < b \mid \xi = x) = \frac{\int_a^b f(x|y)\,dG(y)}{\int_{-\infty}^{\infty} f(x|t)\,dG(t)}$$

alakban is.

3. $\pi = \pi'$ (ill. $\pi = \pi''$) esetén nem egyértelmü a feladat megoldása (van δ_0 és $\delta \in \mathcal{C}$ is). Ha már van egy x_1 megfigyelés, a legjobb eljárás 2. szerinti. Nemcsak az első, hanem a további lépésekben is lényegtelen melyik utat választjuk.

2. Tétel. Tetszőleges $0 < \pi_0' < \pi_0'' < 1$ értékekhez találhatók olyan $0 < w < 1$ és $c > 0$ számok, hogy a kiegészitő feladat * Bayes megoldása – a $w_0 = 1 - w_1$, $w_1 = w$, c értékekkel és $\pi\ (\pi_0' < \pi < \pi_0'')$ apriori eloszlással – a szekvenciális likelihood hányadospróbával adódik

$$A_0 = \frac{\pi}{1-\pi}\;\frac{1-\pi_0''}{\pi_0''}, \qquad A_1 = \frac{\pi}{1-\pi}\;\frac{1-\pi_0'}{\pi_0'}$$

határokkal.

Bizonyitás. 1. Az 1. tétel szerint π' és π'' a w, c függvényei. Olyan w, c értékek kellenek, melyekre

$$\pi'(w,c) = \pi_0',$$
$$\pi''(w,c) = \pi_0''.$$

Fix w-re jelölje $\pi'(c) = \pi'(c,w)$, $\pi''(c) = \pi''(c,w)$. Ha c_0 a legkisebb c, melyre $\pi'(c_0) = \pi''(c_0)$, akkor $c < c_0$ esetén $\pi'(c)$ és $\pi''(c)$ az

$$(1-w)\pi' = \varrho(\pi', c), \quad (1-\pi'')w = \varrho(\pi'', c)$$

egyenletből kerülnek meghatározásra.

A $\varrho(\pi', c)$ függvény, mint c függvénye (π fix)

I. folytonos, ami a konkávság következménye,

II. szigoruan növekvő, mert $\delta \in \mathcal{C}$ esetén a veszteség nő c-vel és a minimális veszteség $\varrho(\pi', c)$ $\mathcal{C}$-beli eljárásra éretik el.

---

* lásd a 24. oldalt

III. ha $c \to 0$ $\varrho(\pi', c) \to 0$, $\varrho(\pi'' c) \to 0$, ami abból következik, hogy fix megfigyelésszám esetén (ha n elég nagy) a hibavalószinüség tetszőleges kicsi lehet.

Ezért $0 < c < c_0$-ban π' folytonos, monoton növekvő függvény, $\pi'(c) \to 0$, ha $c \to 0$. π'' folytonos, monoton csökkenő függvény, $\pi''(c) \to 1$, ha $c \to 0$.
Ha $c \to c_0$, $\pi''(c) - \pi'(c) \to w$, $\pi''(c) \to w$, $\pi'(c) \to w$, ahol w a $\pi'(1-w) = (1-\pi'')w$ egyenlet megoldása.
Fix w esetén

$$\lambda(c) = \frac{\pi'(c)}{1-\pi'(c)} \cdot \frac{1-\pi''(c)}{\pi''(c)}$$

függvény folytonos, szigoruan növekvő $0 < c < c_0 (= c_0(w))$ és $0 \leq \lambda(c) \leq 1$.

2. Legyen

$$\lambda(w,c) = \frac{\pi'(w,c)}{1-\pi'(w,c)} \cdot \frac{1-\pi''(w,c)}{\pi''(w,c)}, \quad \gamma(w,c) = \frac{\pi''(w,c)}{1-\pi''(w,c)}.$$

Ezekkel egyszerübb dolgozni, mint π', π''-vel. Megmutatjuk léteznek olyan w és c értékek, hogy

$$\lambda(w,c) = \frac{\pi_0'}{1-\pi_0'} \cdot \frac{1-\pi_0''}{\pi_0''} = \lambda_0, \quad \gamma(w,c) = \frac{\pi_0''}{1-\pi_0''} = \gamma_0.$$

Az 1. részben láttuk, hogy tetszőleges w esetén létezik c=c(w), melyre $\lambda(w,c) = \lambda_0$. Alább bizonyitjuk, hogy $\gamma(w) = \gamma[w, c(w)]$ kölcsönösen egyértelmü leképezése $0 < w < 1$ -nek $0 < \gamma < \infty$ -re, és van egyetlen olyan w, melyre $\gamma(w) = \gamma_0$.

3. Az 1. tétel szerint a kiegészitő feladatban w, c=c(w) és $\pi = \pi'[w, c(w)]$ konstansok esetén létezik δ' Bayes megoldás, mégpedig a szekvenciális likelihood hányados próba

$$A_0' = \frac{\pi'[w, c(w)]}{1-\pi'[w,c(w)]} \cdot \frac{1-\pi''[w, c(w)]}{\pi''[w, c(w)]} = \lambda[w, c(w)] = \lambda_0, \quad A_1' = \frac{\pi'}{1-\pi'} \cdot \frac{1-\pi'}{\pi'} = 1$$

határokkal.
Legyen δ'' a megfelelő megoldás w, c=c(w), $\pi = \pi''[w, c(w)]$ értékek esetén, azaz a szekvenciális likelihood hányados próba

$$A_0'' = 1, \quad A_1'' = \frac{\pi''[w, c(w)]}{1-\pi''[w,c(w)]} \cdot \frac{1-\pi'[w, c(w)]}{\pi'[w,c(w)]} = \frac{1}{\lambda_0}$$

határokkal.
Ekkor $\alpha_0', \alpha_1', E_0'(N), E_1'(N)$ a δ' Bayes féle döntési eljárás, mig $\alpha_0'', \alpha_1'', E_0''(N), E_1''(N)$ a δ'' Bayes féle döntési eljárás esetén a w,c-től csak λ_0 -on keresztül függenek,

s nem függnek γ-tól. Ha $\lambda_0$fix, a fenti értékek is fix számok. A Bayes féle veszteségek (rizikók) $\pi = \pi'(w, c(w))$ és $\pi = \pi''[w, c(w)]$ esetén

$$\rho(\pi') = r(\pi', \delta') \quad \text{és} \quad \rho(\pi'') = r(\pi'', \delta).$$

A π' és π'' definiciójából (valamint δ_0, δ_1 definiciójából) adódik, hogy $r(\pi', \delta_0) = r(\pi', \delta')$ és $r(\pi'', \delta'') = r(\pi'', \delta_1)$, vagy kiirva ezeket az összefüggéseket

$$\pi'(1-w) = \pi'[\alpha_0'(1-w) + cE_0'(N)] + (1-\pi')[\alpha_1' w + cE_1'(N)]$$

és

$$w(1-\pi'') = \pi''[\alpha_0''(1-w) + cE_0''(N)] + (1-\pi'')[\alpha_1'' w + cE_1''(N)].$$

A $\pi'/(1-\pi')$ helyett $\lambda_0\gamma$ és $\pi''/(1-\pi'')$ helyett γ-értéket irva és c-t kiküszöbölve γ és w közötti összefüggés

$$\{\lambda_0\gamma(1-\alpha_0') - w[\lambda_0\gamma(1-\alpha_0') + \alpha_1']\}\{\gamma E_0''(N) + E_1''(N)\} =$$
$$= -\gamma\alpha_0'' + w[(1-\alpha_1'') + \gamma\alpha_0'']\}\{\lambda_0\gamma E_0'(N) + E_1'(N)\}$$

w-ben lineáris és $\gamma>0$ esetén van $0<w<1$ megoldása. γ-ban négyzetes kifejezés a konstans tag és a négyzetes tag ellenkező előjelüek $0<w<1$ esetén, igy van egy, és csakis egy pozitiv megoldása, mely megadja a w és γ közötti kapcsolatot.

A szekvenciális módszer alaptételének bizonyitása

Tekintsünk egy tetszőleges szekvenciális likelihood hányados eljárást $A_0 < 1 < A_1$ határokkal és $0 < \pi < 1$ valószinüséggel. Legyen

$$\pi' = \frac{\pi}{A_1(1-\pi) + \pi}, \qquad \pi'' = \frac{\pi}{A_0(1-\pi) + \pi}$$

akkor

$$A_0 = \frac{\pi}{1-\pi} \; \frac{1-\pi''}{\pi''}$$

és $A_1 = \pi/(1-\pi) \cdot (1-\pi')/\pi'$. A 2. tétel alapján találhatók olyan w, c értékek, hogy a próba a Bayes féle kiegészitő feladatra, ahol π valószinüséggel p_0 a sürüségfüggvény, a veszteség $w_0=1-w$, $w_1=w$; egy megfigyelés c értékü. Legyen α_0, α_1 a hibavalószinüség $(E_0(N), E_1(N))$ az átlagos megfigyelésszám. Legyen δ^* tetszőleges eljárás $\alpha_i^* \leq \alpha_i$ (i=1,2) hibavalószinüséggel, és $E_i^*(N) < \infty$ megfigyelésszámmal.

Mivel a szekvenciális likelihood eljárás minimalizálja a

Bayes féle rizikót

$$\pi[(1-w)\alpha_0+cE_0(N)]+(1-\pi)[w\alpha_1+cE_1(N)]\leq$$
$$\leq\pi[(1-w)\alpha_0^*+cE_0^*(N)]+(1-\pi)[w\alpha_1^*+cE_1^*(N)],$$

és igy $\pi E_0(N)+(1-\pi)E_1(N)\leq\pi E_0^*(N)+(1-\pi)E_1^*(N)$ minden $0<\pi<1$ -re, ahonnan adódik a bizonyitandó $E_0(N)\leq E_0^*(N)$, $E_1(N)\leq E_1^*(N)$ egyenlőtlenség pár.

3.3. Szekvenciális becslésekről

A szekvenciális becslés alapgondolata: olyan mintát keresünk, mely előre megadott pontosságu becslést ad, vagy maximalizálja a pontosságot adott mintaár esetén.

Intervallum becslés esetén az intervallum hosszával adható meg a pontosság (adott megbizhatósági szinten). Pontbecslés esetén előre megadott veszteségfüggvény lehet a pontosság*.

Stein adta meg normális eloszlásu változó várható értéke megadott hosszuságu konfidencia intervallum becslését két lépésben, amikor a szórás ismeretlen. Vizsgáljuk meg a Stein féle eljárást.

Ha a megfigyelésszám n fix a μ várható értékre $(1-\alpha)$ szintü konfidencia intervallum a t-eloszlás alapján

$$\bar{X}\pm\frac{s\cdot t_{\alpha/2}}{\sqrt{n}},\quad\left(\bar{X}=\frac{1}{n}\sum X_i,\; s^2=\frac{1}{n-1}\sum(X_i-\bar{X})^2\right)$$

lesz, ahol $t_{\alpha/2}$ az n-1 szabadságfoku eloszlás $1-\alpha/2$ kvantilise. A konfidencia intervallum hossza $\frac{2s\cdot t_{\alpha/2}}{\sqrt{n}}$ maga is valószinüségi változó (mely nagy értékeket is felvehet).
(Dantzig (1940) bizonyitotta, hogy egy minta alapján ismeretlen σ esetén nem lehetséges meghatározott hosszuságu konfidencia intervallum szerkesztése μ-re.)
Legyen $X_1,\dots,X_m$ egy minta (m előre adott), $\bar{X}_m$, s^2 empirikus értékekkel. Jelölje n azt a legkisebb $(n>m)$ egészet, melyre

$$\frac{s^2}{n}\leq\frac{c^2}{b^2}$$

ahol b^2 egy m-1 szabadságfoku t-eloszlásu változó szórásnégyzete, c adott állandó. Ezután kiegészitő mintát veszünk, mely n-m nagyságu kell legyen. Az összes megfigyelés $\bar{X}_n=\frac{1}{n}(X_1+\dots+X_n)$ közép-

* A szekvenciális becslés problémakörével foglalkozott F.Anscombe (Proc.Cambridge Phil.Soc.(1952)48,600-607, J.Roy.Stat.Soc.B.15(1953) 1-29.),Ch. Stein (Ann.Math.Stat.(1945)16,243-258.),J. Linnik (Dokladii, 1968-72. számaiban)

értékére igazak a következő állitások.

1. Állitás: $\bar{X}_n$ - torzitatlan becslése μ-nek, szórása $\leq$ c.

A $\frac{\sqrt{n}(\bar{X}_n-\mu)}{\sigma}$ és $\frac{(m-1)}{\sigma^2}s^2$ változók függetlenek N(0,1) ill. χ^2_{m-1} eloszlásuak, ezért

$\frac{\sqrt{n}(\bar{X}-\mu)}{s}$ t eloszlásu (m-1 szabadságfoku) és

$$\frac{E\sqrt{n}(\bar{X}-\mu)}{\sigma}=0$$

Mivel n is valószinüségi változó, nem adódik azonnal, hogy $E(\bar{X}-\mu)=0$.

$$E(\bar{X}-\mu)=E\left(\frac{\sqrt{n}(\bar{X}-\mu)}{\sigma}\cdot\frac{\sigma}{\sqrt{n}}\right)=E\left(\frac{\sqrt{n}(\bar{X}-\mu)}{\sigma}\cdot f(s^2,\sigma)\right)=E\left(\frac{\sqrt{n}(\bar{X}-\mu)}{\sigma}\right)E(f)=0.$$

Továbbá

$$E(\bar{X}-\mu)^2=E\left(\frac{n(\bar{X}-\mu)^2}{s^2}\cdot\frac{s^2}{n}\right)\leq E\left(\frac{n(\bar{X}-\mu)^2}{s^2}\cdot\frac{c^2}{b^2}\right)=$$

$$=\frac{c^2}{b^2}E\left(\frac{n(\bar{X}-\mu)^2}{s^2}\right)=\frac{c^2}{b^2}\cdot b^2=c^2.$$

2. Állitás: Legyen a μ konfidencia intervallumának hossza 2ℓ, akkor $c=b\ell/t_{\alpha/2}$ választással $\bar{X}\pm\ell$ megbizhatósága legalább $1-\alpha$.

$$P\{|\bar{X}-\mu|<\ell\}=P\left\{\frac{|\bar{X}-\mu|}{s}\sqrt{n}<\frac{\sqrt{n}\,\ell}{s}\right\}\geq P\left\{\frac{|\sqrt{n}(\bar{X}-\mu)|}{s}<\frac{b\ell}{c}\right\}$$

ez utóbbi valószinüsége $1-\alpha$ (mivel $\frac{b\ell}{c}=t_{\alpha/2}$). Tehát ha az első minta elemszáma adott (tetszőleges), kétlépéses eljárással szerkeszthető megadott hosszuságu intervallum becslés.

3.4. A szekvenciális likelihood próba ereje.

Monoton a likelihood hányados, ha tetszőleges θ',θ-ra a P_θ, $P_{\theta'}$ eloszlások különbözőek, és létezik olyan T(x), hogy a sürüségfüggvények hányadosa $p_{\theta'}(x)/p_\theta(x)$ T(x)-nek nemcsökkenő függvénye.

Meg lehet mutatni, hogy a hipergeometrikus eloszlás monoton likelihood hányadosu. Jelölje D a selejtesek számát N-ből, n legyen a minta elemszáma és ξ jelölje a mintában a selejtesek számát. Akkor (D a paraméter)

$$P(\xi=x)=p_D(x)=\binom{D}{x}\binom{N-D}{n-x}\bigg/\binom{N}{n},\quad \frac{p_{D+1}(x)}{p_D(x)}=\frac{D+1}{N-D}\cdot\frac{N-D-n+x}{D+1-x}$$

ahonnan látható, hogy T(x) = x-ben monton likelihood hányadosu eloszlás seregről van szó.

Rögzitett megfigyelésszám esetén a

$$H_0: \theta \leq \theta_0,$$

$$H_1: \theta > \theta_0,$$

összetett hipotézisekre létezik egyenletesen legerősebb próba, ha a likelihood hányados monoton (lásd pl. Lehmann könyve 101.o.).

Ha P_θ egyparaméteres exponenciális eloszlássereg, azaz $P_\theta(x) = C(\theta)\exp\{Q(\theta)T(x)\}h(x)$, ahol $Q(\theta)$ monoton, a likelihood hányadosról könnyü megmutatni a monotonitást.

Ilyen exponenciális eloszlás pl. a binomiális eloszlás $\binom{n}{x}p^x(1-p)^{n-x}$ - valószinüségekkel, ahol $T(x) = x$, $\theta = p$, $Q(p) = \log[p/(1-p)]$.

Egyenletesen legerősebb próba szekvenciális eljárás esetén nem létezik. Ha θ_0 és θ_1 esetén a likelihood hányados próba szekvenciális eljárása α_0, α_1 hibákkal optimális, attól $\theta_2 \neq \theta_1$ esetén a szekvenciális eljárás nem lesz azonos a likelihood hányadoson alapuló szekvenciális eljárással θ_1 mellett, igy nem lesz a legjobb.

<u>Lemma</u>. Legyen $X_1, X_2, \ldots,$ független sorozat $p_\theta(x)$ sürüséggel. Legyen $p_\theta(x)$ likelihood hányadosa $T(x)$-ben monoton. Tetszőleges, a likelihood hányadoson alapuló szekvenciális eljárás θ_0 ellenőrzésére θ_1 $(\theta_0 < \theta_1)$ ellenében nemcsökkenő erőfüggvényü.

<u>Bizonyitás</u>. Legyen $z_i = \log[p_{\theta_1}(x_i)/p_{\theta_0}(x_i)] = h(T_i)$; ahol a feltevés szerint $h(T)$ nemcsökkenő, és $\theta < \theta'$. A bizonyitásban felhasználunk két, később bizonyitásra kerülő segédtételt (lásd 36.o.).

Ha $\theta < \theta'$ a 2. segédtétel szerint $F_{\theta'}(t) \leq F_\theta(t)$ minden t-re. Az 1. segédtétel szerint léteznek olyan V változó és f, f' függvények ($f(v) \leq f'(v)$), hogy $f(V)$, $f'(V)$ eloszlása F_θ, F'_θ. A szekvenciális eljárás folytatódik mindaddig, amig $\log A_0 < \sum_1^n h(t_i) < \log A_1$ feltétel teljesül. A hipotézis elvetve, ha a felső határon tulmegy az összeg. A kiszámitáshoz θ esetén T_i helyett $f(V_i)$, mig θ' esetén T_i helyett $f'(V_i)$ helyettesitendő.

Mivel $f(V_i) \leq f'(V_i)$, ha $\sum h(f(V_i)) > \log A_1$, egyben $\sum h(f'(V_i)) > \log A_1$ is teljesül. Innen következik, hogy P_θ (a hipotézis (θ) elvetése) $\leq P_{\theta'}$ (a hipotézis (θ') elvetése), azaz $\beta(\theta) \leq \beta(\theta')$, amit bizonyitani akartunk.

Az a kivánalom tehát, hogy

$$(*)\qquad \begin{array}{lll} \beta(\theta) \leq \alpha_0, & \text{ha } \theta \leq \theta_0 & (\theta_0 < \theta_1) \\ \beta(\theta) \geq \alpha_1, & \text{ha } \theta \geq \theta_1 & \end{array}$$

, teljesül.

A szekvenciális eljárás főtétele szerint a fenti (*) feltételnek eleget tevő eljárások közül a szekvenciális likelihood hányados próba $\theta = \theta_0$ és $\theta = \theta_1$ esetén minimalizálja a megfigyelésszámot (átlagban).

Az erőfüggvény és operativ karakterisztika, mely fogalmat korábban ismertünk meg, nyilván kapcsolatban vannak, mégpedig $\beta(\theta) = 1 - \pi_\theta(H)$.

$E_\theta(N)$ viselkedése érdekel bennünket, mint θ függvénye. A szekvenciális eljárás (likelihood hányadoson alapuló) optimális tulajdonsága miatt a megfigyelésszám $(E_\theta(N))\theta = \theta_0$ és $\theta = \theta_1$ esetén minimális. Rendszerint $E_\theta(N)$ a (θ_0, θ_1) intervallum belsejében éri el maximumát. Ez sokszor kisebb a fix n_0 megfigyelésszámnál. Azonban nem igaz, hogy minden esetben $E_\theta(N) \leq n_0$ a fix megfigyelésszám θ_0, θ_1 megkülönböztetésére (α_0, α_1) erő esetén).
Például binomiális esetben $p_0 = 0.4$ $p_1 = 0.6$ $\alpha_0 = \alpha_1 = 0.005$ esetén $n_0 = 160$ és $p = 1/2$ esetén $E_p(N) = 170$.

Megoldatlan probléma olyan kritérium keresése, melyre $\sup_\theta E_\theta(N)$ minimális a (*) feltétel mellett.

Két segédtétel.

1. Segédtétel. Legyenek $F_0(x)$, $F_1(x)$ eloszlásfüggvények. Az $F_1(x) \leq F_0(x)$ egyenlőtlenség akkor és csak akkor teljesül, ha léteznek nemcsökkenő f_0, f_1 függvények és V-valószinüségi változó, hogy

a) $f_0(v) \leq f_1(v)$ minden v-re,

b) $f_0(V) \leq f_1(V)$ eloszlásai F_0 és F_1.

2. Segédtétel. Legyen $p_\theta(x)$ monoton likelihood hányadosu x-ben (azaz $p_\theta(x)/p_\theta(x)$ monoton növekvő).

I. Ha Ψ nemcsökkenő, akkor $E_\theta \Psi$ θ-ben nemcsökkenő.
Ha $X_1, X_2, \dots, X_n$ függetlenek $p_\theta(x)$ sürüséggel, $\Psi' = \Psi'(x_1, \dots, x_n)$ és minden változóban nemcsökkenő, akkor $E_\theta \Psi'(x_1, \dots, x_n)$ nemcsökkenő.

II. Minden $\theta < \theta'$ esetén az eloszlásokra fennáll az $F_{\theta'}(x) \leq F_\theta(x)$ (minden x-re) egyenlőtlenség.

III. Ψ pontosan egyszer váltson előjelet. (Pontosan $\Psi(x) \leq 0$, ha $x < x_0$ és $\Psi(x) > 0$, ha $x \geq x_0$).
Akkor létezik olyan θ_0, melyre $E_\theta \Psi(x) \leq 0$, ha $\theta < \theta_0$ és $E_\theta \Psi(x) \geq 0$, ha $\theta > \theta_0$, és kizárt, hogy $E_\theta \Psi(x)$ pozitiv (vagy negativ) minden θ-ra.

1. Bizonyitása. Ha létezik a kivánt tulajdonságu f_0, f_1 és V, akkor $F_1(x)=P\{f_1(V)\leq x\}\leq P\{f_0(V)\leq x\}=F_0(x)$.
Megforditva legyen $F_1(x)\leq F_0(x)$ és $f_i(y)=\inf\{x: F_i(x-0)\leq y\leq F_i(x)\}$, $i=0,1$
(Feltételezzük, hogy $F(x)=P(\xi\leq x)$, azaz $F(x)$ jobbról folytonos)
Az f_i függvények nemcsökkenőek és $f(F(x))\leq x$, $F(f(y))\geq y$, minden x,y-ra.
Ezért $y\leq F(x_0)$ -ból következik, hogy $f(y)\leq f(F(x_0))\leq x_0$ és megforditva $f(y)\leq x_0$ -ból $F(f(y))\leq F(x_0)$ és igy $y\leq F(x_0)$ következik.
Ez az $f(y)\leq x_0$ és $y\leq F(x_0)$ egyenlőtlenségek ekvivalens voltát jelenti. Legyen V egyenletes eloszlásu változó [0,1]-ben, akkor

$$P\{f_i(V)\leq x\}=P\{V\leq F_i(x)\}=F_i(x).$$

Abból, hogy $F_1(x)\leq F_0(x)$, minden x-re, következik

$$f_0(y)\leq f_1(y),$$

amivel az 1. segédtétel bizonyitása kész.
Egyszerü példa a sztohasztikusan rendezett sokaságra az eltolás (elmozdulás) paraméterrel megadott eloszlás sereg:

$$F_\theta(x)=F(x-\theta).$$

Ha $\theta<\theta'$

$$F(x-\theta)=P\{\xi\leq x-\theta\}\geq P\{\xi\leq x-\theta'\}=F(x-\theta').$$

2. Bizonyitása. I. Legyen $\theta<\theta'$. $A=\{x: p_{\theta'}(x)<p_\theta(x)\}$

$B=\{x: p_{\theta'}(x)>p_\theta(x)\}$

Ha $a=\sup_A \Psi(x)$ és $b=\inf_B \Psi(x)$ igaz, hogy $b-a\geq 0$, továbbá

$$\int\Psi(x)[p_{\theta'}-p_\theta]d\mu\geq b\int_B(p_{\theta'}-p_\theta)d\mu+a\int_A(p_{\theta'}-p_\theta)d\mu=$$

$$=(b-a)\int_B(p_{\theta'}-p_\theta)d\mu\geq 0,$$

és ez az első állitás $n>1$-re indukcióval adódik a bizonyitás.

II. Legyen $\Psi(x)=\begin{cases}1 & x>x_0\\ 0 & \text{egyébként}\end{cases}$ akkor II. az I. következménye.

III. Először megmutatjuk, hogy tetszőleges $\theta'<\theta''$ esetén az $E_{\theta'}\Psi(x)>0$ egyenlőtlenségből adódik $E_{\theta''}\Psi(x)>0$.
Ha $p_{\theta''}(x_0)/p_{\theta'}(x_0)=\infty$ akkor $p_{\theta'}(x)=0$ $x>x_0$-ra (a monotonitás miatt) és $E_{\theta'}\Psi(x)\leq 0$. Legyen ezért
$p_{\theta''}(x_0)/p_{\theta'}(x_0)=C<\infty$, akkor $\Psi(x)\geq 0$ az $S=\{x: p_{\theta'}(x)=0$ és $p_{\theta''}(x)>0\}$ halmazon, igy

$$E_{\theta''}\Psi(x)\geq\int_{\bar S}\Psi\frac{p_{\theta''}}{p_{\theta'}}p_{\theta'}d\mu\geq\int_{x_0}^{\infty}C\Psi p_{\theta'}d\mu+\int_{-\infty}^{x_0}C\Psi p_{\theta'}d\mu=CE_{\theta'}\Psi\geq 0$$

(az (x_0,∞)-ben $\Psi \geq 0$, a$(-\infty, x_0)$-ban $\Psi \leq 0$).

Legyen $\theta_0 = \inf\{\theta : E_\theta \Psi(x) > 0\}$.

A II. rész szerinti monoton likelihood hányados esetén az eloszlás sztohasztikusan növő. A megforditás általában nem igaz, példa erre az $\frac{1}{\pi}\frac{1}{1+(x-\theta)^2}$ Cauchy eloszlás, mely nem monoton likelihood hányadosu, bár sztohasztikusan növő eloszlássereg.

Feladatok.

1. Mutassuk meg, hogy a 3.§.-ban definiált Bayes féle minimális veszteség függvény (rizikó), $\min_{\delta} r(\varrho,\delta) = r^*(\varrho)$ a $\varrho(\alpha)$ függvények konkáv függvénye. Azaz, ha $\varrho_1(\alpha)$ és $\varrho_2(\alpha)$ tetszőleges eloszlásai (sürüségfüggvényei) α-nak és $0 \leq \lambda \leq 1$ tetszőleges, akkor $\lambda\varrho_1 + (1-\lambda)\varrho_2$ ismét eloszlás (sürüségfüggvény) és
$$r^*[\lambda\varrho_1 + (1-\lambda)\varrho_2] \geq \lambda r^*(\varrho_1) + (1-\lambda) r^*(\varrho_2).$$

2. Legyen a paraméter tér két elemü $\theta = \{0,1\}$ és a döntések tere legyen egy intervallum $D = [0,1]$, a veszteségfüggvény $R(\theta,d) = EL(\theta,d) = |\theta-d|^\alpha$, ahol α adott egész szám. Legyen továbbá az apriori eloszlás $P(\theta=0) = 3/4$, $P(\theta=1) = 1/4$. Mutassuk meg, hogy $\alpha = 1$ esetén az egyetlen Bayes féle megoldás (döntés) $d^* = 0$ és ekkor $r^*(\varrho) = 1/4$. (Vegyük észre, hogy $D=(0,1]$ esetén $r^*(\varrho) = 1/4$, de nincs Bayes féle megoldás.) Ha $\alpha>1$ az egyetlen Bayes féle megoldás $d^* = (1 + 3^{1/\alpha-1})^{-1}$. [Utmutatás: $r(\varrho,d) = 3/4\, d^\alpha + \frac{1}{4}(1-d)^\alpha$.]

3. Randomizált döntési eljárás esetén, ha x megfigyelési eredmény esetén az i-edik döntés (di) valószinüsége $\lambda_i(x)$ $[i=1,2,\dots m]$, ahol $\lambda_i(x) \geq 0$ $\sum_{i=1}^{m} \lambda_i(x) = 1$,
$$L(\theta_i, \delta(x)) = \sum_{j=1}^{m} \lambda_j(x) L(\theta_i, d_j)$$
és az
$$L = (L_1,\dots,L_k)^* = (L(\theta_1,\delta(x)),\dots, L(\theta_k,\delta(x)))^*$$
vektor jellemzi a veszteséget. A $\delta_2(x)$ döntési eljárás jobb, mint $\delta_1(x)$, ha $L_{2i} \leq L_{1i}$ minden $i=1,\dots,k$ értékre és legalább egy i értékre egyenlőtlenség áll fenn. A rizikófüggvényt (vektorfüggvényt)

(x) $R = (R_1,\dots,R_k)^* = (E_{\theta_1} L(\theta_1,\delta(x)),\dots, E_{\theta_k} L(\theta_k,\delta(x)))^*$

adja meg. $\delta(x)$ megengedhető, ha nincs nála jobb eljárás (azaz olyan $\tilde{\delta}$ melyre legalább egy i-re $\tilde{R}_i < R_i$). Ha a kezdeti eloszlással, azaz $P(\theta_\eta = \theta_i) = \varrho_i$, akkor a Bayes féle veszteség

$$(**) \quad r(\varrho,\delta) = \sum_{i=1}^{k} \varrho_i R_i$$

vagy folytonos eloszlás esetén

$$r(\varrho,\delta) = \int L_\theta\, \varrho(\theta) d\theta .$$

A (*) előállitásból látható, hogy az összes lehetséges R pontot tartalmazó G halmaz az $R(d_1),\dots, R(d_m)$ egyszerü döntéseknek megfelelő m pont konvex burka. Az $r(\varrho,\delta)$ rizikó a $\underline{\varrho}$ és $\underline{R}$ vektorok skalár sorozata. δ^* akkor és csak akkor Bayes féle megoldás, ha a $\underline{\varrho}.\underline{R}(\delta)$ sorozatot minimalizálja. Mutassuk meg, hogy ha δ^*megengedhető döntés, akkor $R(\delta^*)$ akkor és csak akkor pontja G Bayes féle hatásának, ha létezik olyan $\underline{\varrho}$ eloszlás, melyre δ^* Bayes féle döntés. (Az $\underline{x} = (x_1,\dots x_n) \in G$ pont akkor tartozik G Bayes féle határához, ha nincs olyan $\underline{y} = (y_1,\dots,y_n) \in G$ melyre $y_i < x_i$, i=1,2,...k teljesül.) [Utmutatás: részletes leirás található M. DeGroot könyve (lásd 7.§.) 8. fejezetében.]

4. Tegyük fel, hogy

$$P\{\xi=1\mid\theta=1\} = 3/4, \qquad P\{\xi=0\mid\theta=1\} = 1/4,$$
$$P\{\xi=1\mid\theta=0\} = 1/3, \qquad P\{\xi=0\mid\theta=0\} = 2/3,$$
$$P\{\theta=1\} = \pi \qquad P\{\theta=0\} = 1-\pi, \qquad (0 \le \pi \le 1).$$

Az R veszteséget a következő táblázat adja meg:

	d_1	d_0
$\theta=1$	0	5
$\theta=0$	10	0

A ξ megfigyelés értéke 0 vagy 1 lehet. A Bayes féle tétel alapján

$$P(\theta=1|\xi=1) = \pi(1) = \frac{3/4\,\pi}{3/4\,\pi + \frac{1}{3}(1-\pi)}, \quad P(\theta=1|\xi=0) = \frac{1/4\,\pi}{1/4\,\pi + 3/4(1-\pi)} = \pi(0)$$

ξ megfigyelés után d_1 és d_0 döntések közül kell választani.

A d_1 döntés esetén a rizikó $10[1-\pi(\xi)]$ míg d_0 esetén $5\pi(\xi)$, és Bayes-i a döntés, ha: $\pi(\xi)>2/3$ esetén d_1, $\pi(\xi)<2/3$ esetén d_0 (ha $\pi(\xi)=2/3$ mindkét döntés Bayes-i).
Mutassuk meg, hogy a Bayes-i döntésfüggvény $\delta^*(\xi)$ a következő alaku

$$\delta^*(1)=\begin{cases} d_0 \text{ ha } & \pi(1)<2/3, \text{ (azaz } \pi<8/17), \\ d_1 \text{ ha } & \pi(1)>2/3, \text{ (azaz } \pi>8/17), \end{cases}$$

$$\delta^*(0)=\begin{cases} d_1 \text{ ha } & \pi(0)>2/3, \text{ (azaz } \pi>(\frac{16}{19})), \\ d_0 \text{ ha } & \pi(0)<2/3, \text{ (azaz } \pi<(\frac{16}{19})). \end{cases}$$

A Bayes féle döntés tehát a következőkben áll: $\pi<8/17$ esetén a ξ megfigyelés kimenetelétől függetlenül d_0, $8/17<\pi<\frac{16}{19}$ esetén $\delta^*(0)=d_0$, $\delta^*(1)=d_1$, végül $16/19<\pi$ esetén a ξ megfigyelés kimenetelétől függetlenül d_1.

Mutassuk meg, hogy a Bayes féle rizikó

$$r^*(\pi)=\begin{cases} 5\pi & \text{,ha } 0\le\pi\le 8/17 \\ \frac{5}{4}\pi+\frac{10}{3}(1-\pi) & \text{,ha } 8/17\le\pi\le\frac{16}{19} \\ 10(1-\pi) & \text{,ha } \frac{16}{19}\le\pi\le 1. \end{cases}$$

5. Tegyük fel, hogy az előbbi feladat feltételei teljesülnek, de egy megfigyelés értéke c. Milyen c értéket érdemes fizetni ? (Rajzoljuk meg az $r^*(\pi)$ és a megfigyelés nélküli $r_0^*(\pi)$ függvényeket.)

6. A 4. feladat feltételei mellett vizsgáljuk a Bayes féle döntést és a rizikó függvényt, ha a megfigyelések száma n. [Utmutatás:
$$\pi(x_1,\dots x_n)=\frac{\pi\cdot\frac{3^{\Sigma x_i}}{4^n}}{\pi\cdot\frac{3^{\Sigma x_i}}{4^n}+(1-\pi)\frac{2^{n-\Sigma x_i}}{3^n}}\ .]$$

7. Legyen a paraméter tér $\theta=\{0,1\}$ a döntési tér $D=\{0,1\}$ és az átlagos veszteséget adja meg a következő táblázat

	$d=0$	$d=1$
$\theta=0$	0	a_0
$\theta=1$	a_1	0

Legyen θ apriori eloszlása π ($P(\theta=0)=\pi$). A δ döntési eljárás esetén a hibás döntés valószinüsége legyen $\alpha(\delta)$, ha $\theta=0$ és $\beta(\delta)$, ha $\theta=1$. A rizikó

$$r(\pi,\delta)=a_0\pi\,\alpha(\delta)+a_1(1-\pi)\beta(\delta)=a\,\alpha(\delta)+b\,\beta(\delta).$$

Legyen ξ feltételes sürüségfüggvénye $\theta=i$ esetén $(i=0,1)$ $p_i(x)$.

A Neyman-Pearson lemma alapján bizonyitsuk be a következő állitást: tetszőleges a,b pozitiv értékekre a

$$\delta^*(x) = \begin{cases} d=0 & ,\text{ha} \quad af_0(x) > bf_1(x) \\ d=1 & ,\text{ha} \quad af_0(x) < bf_1(x) \\ d=0 \text{ vagy } d=1 & \text{az} \quad af_0(x) = bf_1(x) \text{ esetben,} \end{cases}$$

döntés a legjobb abban az értelemben, hogy tetszőleges δ-döntésre

$$a\alpha(\delta^*) + b\beta(\delta^*) \le a\alpha(\delta) + b\beta(\delta) .$$

8. Legyen a paraméter tér $\{0,1\}$ és a döntési tér is $\{0,1\} = D$. A feltételes sürüség legyen normális -1,9 (várható érték, szórásnégyzet), mig $\theta = 1$ esetén +1,9. Ha egy megfigyelés ára 1 egység és az átlagos veszteség

	$d=0$	$d=1$
$\theta=0$	0	1
$\theta=1$	1	0

A megfigyelések száma n. Mutassuk meg, hogy az optimális megfigyelésszám n=42, és a minimális teljes rizikó 57.4.

9. Bizonyitsuk be, hogy monoton likelihood hányadosu (T(x)-ben) ξ valószinüségi változó $p_\theta(x)$ sürüségfüggvénye esetén igazak a következő állitások:

a) A $H_0: \theta \le \theta_0$ és $H_1: \theta > \theta_0$ alternativ hipotézisek esetén létezik egyenletesen legerősebb próba, melyet a C és γ konstansokkal az alábbi módon lehet megadni (a kritikus tartomány indikátorfüggvénye helyett)

$$\varphi(x) = \begin{cases} 1 & \text{ha} \quad T(x) > C \\ \gamma & \text{ha} \quad T(x) = C \\ 0 & \text{ha} \quad T(x) < C \end{cases} , \quad E_{\theta_0}\varphi(x) = \alpha .$$

b) A $\beta(\theta) = E_\theta\varphi(\xi)$ erőfüggvény a $\beta(\theta) < 1$ pontokban szigoruan növekvő.

4. A Brown-mozgás folyamat szekvenciális feladatai.

1. Tekintsük a $w(t)$ standard Brown-mozgás folyamatot a $t \geq 0$ értékekre. Feltevésünk szerint $Ew(t) = 0$, $Ew^2(t) = t$ és $w(0) \equiv 0$ (1 valószinüséggel). Az $(\Omega, \mathcal{F})$ mértéktéren legyenek megadva a P_0 és P_1 mértékek, melyek a

$$\sigma w(t), \quad (t \geq 0, \sigma > 0),$$

illetve az

$$r \cdot t + \sigma w(t), \quad (r \neq 0),$$

folyamatnak felelnek meg. A továbbiakban feltesszük, hogy megfigyelhető az a $\xi(t)$ folyamat, amelynek differenciálja

$$d\xi(t) = r\theta dt + \sigma dw(t), \quad \xi(0) \equiv 0, \quad r \neq 0, \quad \sigma > 0$$

alaku, és a θ ismeretlen paraméter lehetséges értékei: $\theta = 1$ (H_1 hipotézis), $\theta = 0$(H_0 hipotézis). A feladat az ismeretlen θ paraméter meghatározásában áll a $\xi(t)$, $(t \geq 0)$, egyetlen realizációja alapján, mégpedig a lehető "legkevesebb" veszteséggel.

A fenti feladat a Bayes féle megfogalmazásban a következő alapfeltevésekből indul ki. Tekintsük a P_π valószinüségi mértékek következő seregét: $P_\pi = \pi P_1 + (1-\pi)P_0$, ahol $0 \leq \pi \leq 1$.
A θ mennyiséget valószinüségi változónak tekintjük, melyre $P_\pi(\theta = 1) = \pi$, $P_\pi(\theta = 0) = 1-\pi$.

A d döntés ($d = d(\omega, \pi)$) lehetséges értékei 0 és 1. $\tau = \tau(\omega, \pi)$ jelöljön Markov féle megállási pontot (a Markov megállási pontok osztálya legyen $\mathfrak{M} = \{\tau\}$). Ha $\mathcal{F}_t^\xi = \sigma\{(\omega,\pi) : \pi, \xi(s), 0 \leq s \leq t\}$, akkor jelölje $D_\tau = \{d\}$ az $\mathcal{F}_\tau$ - mérhető $d = d(\omega, \pi)$ döntések függvényosztályát.

A $\delta = (\tau, d)$, $\tau \in \mathfrak{M}$, $d \in D_\tau$ párt <u>döntési függvénynek</u>, vagy <u>döntési eljárásnak</u> (szabály) nevezzük. $\Delta = \{\delta\}$ a döntési eljárások osztálya.

A fenti feladatban τ jelenti a $\xi(t)$ folyamat megfigyelése megszakításának időpontját és $d = d(\omega, \pi)$ megmutatja (a π és $\xi(s)$, $0 \leq s \leq t$ alapján), hogy a H_0 vagy H_1 hipotézist kell-e elfogadni. Ha $d=1$ a H_1 hipotézist, ellenkező esetben a H_0 hipotézist kell <u>végső döntésként</u> elfogadni.

Jelölje

$$\alpha(\pi, \delta) = P_1\{\omega : d(\omega, \pi) = 0\}, \quad \beta(\pi, \delta) = P_0\{\omega : d(\omega, \pi) = 1\}$$

az első illetve másodfaju hibavalószinüséget a δ döntési eljárás esetén.

Az átlagos veszteséget a következőképpen mérjük:

$$\varrho(\pi,\delta)=\pi[cE_1\tau+aP_1(\omega:d(\omega,\pi)=0)]+(1-\pi)[cE_0\tau+bP_0(\omega:d(\omega,\pi)=1)],$$

ahol a,b,c adott nemnegativ értékek.

A $\delta^*=(\tau^*,d^*)\in\Delta$ döntési eljárás π-Bayes féle, ha

$$\varrho(\pi,\delta^*)=\inf_{\delta\in\Delta}\varrho(\pi,\delta).$$

δ^* döntési eljárás Bayes féle, ha δ^* minden π-re $(0\le\pi\le1)$ π-Bayes féle.

A két hipotézis megkülönböztetésére vonatkozó feladat - a Bayes féle megfogalmazásban - a π-Bayes féle és a Bayes féle döntési eljárás megkeresésében áll.

Az alábbiakban megmutatjuk, hogy a Brown-mozgásra vonatkozó hipotézis vizsgálati feladat bizonyos Markov folyamat optimális megállitására vezethető vissza. Legyen

$$\pi_t(\omega,\pi)=P_\pi\{\theta(\omega)=1\,|\,\mathcal{F}_t^\xi\},\ \sigma\{\omega:\xi_s(\omega),\ s\le t\}=\mathcal{F}_t^\xi\ .$$

Ismeretes továbbá, hogy

$$\varphi_t(\omega)=\frac{dP_1}{dP_0}(\mathcal{F}_t^\xi)(\omega)=\exp\left\{\frac{r}{\sigma^2}\left(\xi(t,\omega)-\frac{r}{2}t\right)\right\},\quad (P_\pi\ \text{m.m}).$$

Ha $\pi=1$, akkor $P_1\{\pi_t(\omega,1)=1\}=1$,

$$\pi_t(\omega)=P_\pi(\theta(\omega)=1\,|\,\mathcal{F}_t^\xi)=\frac{P_\pi(\theta(\omega)=1,\ \mathcal{F}_t^\xi)}{P_\pi(\mathcal{F}_t^\xi)}=\pi\frac{dP_1}{dP_\pi}=$$

$$=\pi\frac{\frac{dP_1}{dP_0}}{\pi\frac{dP_1}{dP_0}+(1-\pi)\cdot1}=\frac{\frac{\pi}{1-\pi}\varphi_t(\omega)}{1+\frac{\pi}{1-\pi}\varphi_t(\omega)}.$$

Vegyük észre, hogy φ_t definiciója alapján

$$\varphi_{t+s}(\omega)=\varphi_t(\omega)\exp\left\{\frac{r}{\sigma^2}\left[\theta(\omega)s+\sigma\cdot(w(t+s)-w(t))-\frac{r}{2}s\right]\right\}$$

és

$$P_\pi\{\theta(\omega)\cdot s+\sigma(w(t+s,\omega)-w(t,\omega))<x\,|\,\mathcal{F}_t^\xi\}=$$

$$=\pi_t(\omega,\pi)P_\pi\left\{w(t+s)-w(s)\le\frac{x-s}{\sigma}\right\}+(1-\pi_t)P_\pi\left\{w(t+s)-w(s)\le\frac{x}{\sigma}\right\}.$$

Innen látszik, hogy a $\{\pi_t(\omega,\pi),\ \mathcal{F}_t^\xi,\ P_\pi\}$ és $\{\varphi_t,\ \mathcal{F}_t^\xi,\ P_\pi\}$ rendszerek minden π-re Markov folyamatot alkotnak.

A hipotézisvizsgálatra vonatkozó Bayes féle feladat megoldható Brown-mozgás folyamat esetén. Ezt mutatja az alábbi tétel, melynek bizonyitására nem térünk ki (megtalálható pl. A.N.Sirjajev: "Statiszticseszkij poszledovatyelnüj analiz" c. könyvében, 172.o.).

Tétel. A $H_1: \theta = 1$ és $H_0: \theta = 0$ egyszerü hipotézisek megkülönböztetésére, a $\{\xi(t), t \geq 0\}$ megfigyelések alapján, létezik Bayes féle döntési eljárás, legyen ez $\delta^* = (\tau^*(\omega,\pi), d^*(\omega,\pi))$, amelyet a következő megállási szabállyal lehet megadni:

$$\tau^*(\omega,\pi) = \inf\{t \geq 0 : \pi_t(\omega,\pi) \bar{\in} (A^*, B^*)\},$$

$$d^*(\omega,\pi) = \begin{cases} 1 & \text{ha } \pi_{\tau^*}(\omega,\pi) \geq B^* \\ 0 & \text{ha } \pi_{\tau^*}(\omega,\pi) \leq A^* \end{cases}$$

Az A^*, B^* értékeket egyértelmüen meg lehet kapni a következő egyenletrendszerből

$$b + a = C\{\Psi(A^*) - \Psi(B^*)\},$$

$$b(1 - B^*) = aA^* + (B^* - A^*)(a - C\Psi(A^*)) + C(\Psi(B^*) - \Psi(A^*)),$$

ahol

$$C = c\left(\frac{r^2}{2\sigma^2}\right)^{-1},$$

$$\Psi(\pi) = (1 - 2\pi)\ln\frac{\pi}{1-\pi},$$

$$\Psi(\pi) = \Psi'(\pi) = \left(\frac{1-\pi}{\pi} - \frac{\pi}{1-\pi}\right) + 2\ln\frac{1-\pi}{\pi}.$$

A rizikó

$$\rho(\pi) = \begin{cases} g(\pi), \text{ ha } & \pi \bar{\in} (A^*, B^*) \\ g(A^*) + (\pi - A^*)(a - C\Psi(A^*)) + C(\Psi(\pi) - \Psi(A^*)), & \pi \in (A^*, B^*). \end{cases}$$

($g(\pi) = \min(a\pi, b(1-\pi))$ és a,b,c az átlagos veszteségfüggvény definiciójában szereplő állandók).

2. A hipotézisvizsgálat u.n. variációs feladatában a θ mennyiségről nem tételezzük fel, hogy valószinüségi változó.

Az $(\Omega, \mathcal{F})$ mérhető téren legyenek megadva a P_0 és P_1 mértékek, $\mathfrak{M}^\xi = \{\tau\}$ jelölje a Markov megállási pontok halmazát, $\tau = \tau(\omega) \leq t \in \mathcal{F}_t^\xi$. Jelölje továbbá $\mathcal{D}_\tau^\xi = \{d\}$ az $\mathcal{F}_\tau^\xi$ mérhető d döntések összességét (d = 0 vagy 1). Az $E_0\tau < \infty$ és $E_1\tau < \infty$ feltételnek eleget tevő $\delta = (\tau, d)$ döntési szabályok osztálya legyen $\Delta^\xi = \Delta^\xi(\alpha, \beta)$, ha az első és másodfaju hiba

$$\alpha(\delta) = P_1\{\omega : d(\omega) = 0\} \leq \alpha, \qquad \beta(\delta) = P_0\{\omega : d(\omega) = 1\} \leq \beta,$$

ahol az α, β nemnegativ számok, kielégitik a $0 < \alpha + \beta < 1$ feltételt.

A variációs feladat Wald féle megfogalmazása a következőképpen szól. Legyenek megadva az α, β nemnegativ számok, $\alpha + \beta < 1$. Olyan $\tilde{\delta}$ döntési szabályt keresünk a $\Delta^\xi(\alpha, \beta)$ osztályban, amelyre egyszerre teljesülnek a

$$E_0\tilde{\tau} \le E_0\tau, \quad E_1\tilde{\tau} \le E_1\tau$$

feltételek, minden $\delta=(\tau,d)\in\Delta^{\xi}(\alpha,\beta)$ döntési szabályra. A hipotézisvizsgálat fenti variációs feladatát a Brown-mozgás folyamat középértékére vonatkozóan meg lehet oldani. Ezt a megoldást mutatjuk be az alábbiakban.

Legyen az $\{\Omega,\mathcal{F},P\}$ valószinüségi téren megadva a $w(t)$ standard Brown-mozgás folyamat. Feltesszük, hogy csak a $\xi(t)$ folyamat figyelhető meg $t\ge 0$ esetén, ahol

$$d\xi(t)=r\theta dt+\sigma dw(t),\quad \xi(0)=0,\quad r\neq 0,\ \sigma>0.$$

A θ ismeretlen paraméter lehetséges értékei 0 (H_0 hipotézis), illetve 1 (H_1 hipotézis). $\mathcal{F}_t^{\xi}=\sigma\{\omega:\xi(s),\ s\le t\}$ és P_1 a $\theta=1$ érték esetén a $\xi(t)$ által generált mérték (i=0,1).
Igaz a következő állitás.

<u>Tétel.</u> Legyen $\alpha+\beta<1$. Ekkor létezik a $\Delta^{\xi}(\alpha,\beta)$ osztály egy olyan $\tilde{\delta}=(\tilde{\tau},\tilde{d})$ döntési eljárása hogy tetszőleges más (α,β) szintü döntési eljárásra

$$E_1\tilde{\tau}\le E_1\tau,\qquad E_0\tilde{\tau}\le E_0\tau.$$

Az optimális döntési eljárás a következő megállási szabály alapján müködik

$$\tilde{\tau}(\omega)=\inf\{t\ge 0;\ \lambda_t\bar{\in}(\tilde{A},\tilde{B})\},$$

$$\tilde{d}(\omega)=\begin{cases}1, & \lambda_{\tau}\ge\tilde{B},\\ 0, & \lambda_{\tau}\le\tilde{A},\end{cases}$$

ahol

$$\lambda_t(\omega)=\ln\varphi_t(\omega)=\frac{r}{\sigma^2}\left\{\xi_t(\omega)-\frac{r}{2}t\right\},\quad \lambda_0(\omega)\equiv 0,$$

$$\tilde{A}=\ln\frac{\alpha}{1-\beta},\quad \tilde{B}=\ln\frac{1-\alpha}{\beta},$$

$$E_0\tilde{\tau}(\omega)=\frac{\omega(\beta,\alpha)}{\rho},\quad E_1\tilde{\tau}(\omega)=\frac{\omega(\alpha,\beta)}{\rho},$$

és

$$\omega(x,y)=(1-x)\ln\frac{1-x}{y}+x\ln\frac{x}{1-y},\quad \rho=\frac{r^2}{2\sigma^2}.$$

A tétel bizonyitását több lépésben végezzük. A szükséges segédeszközök önmagukban is érdekesek, igy legtöbbjük bizonyitását is közöljük.
Legyen

$$\lambda_t^x=x+\frac{r}{\sigma^2}\left\{\xi_t(\omega)-\frac{r}{2}t\right\},\quad \lambda_t(\omega)=\lambda_t^0(\omega)$$

$$\tau_{A,B}^x(\omega)=\inf\{t\ge 0:\lambda_t^x(\omega)\bar{\in}(A,B)\},$$

$$\alpha(x)=P_1(\lambda^x_{\tau^x_{A,B}}=A),\quad \beta(x)=P_0(\lambda^x_{\tau^x_{A,B}}=B)$$

ahol

$$A\le x\le B$$

1. Lemma. Minden $A \le x \le B$-re

$$\alpha(x) = \frac{e^A(e^{B-x}-1)}{e^B-e^A}, \qquad \beta(x) = \frac{e^x - e^A}{e^B - e^A}.$$

Bizonyitás. Ismert (Dünkin* 13.16 tétel), hogy

$$\alpha''(x) + \alpha'(x) = 0, \quad A < x < B, \quad \alpha(A) = 1, \quad \alpha(B) = 0.$$

Hasonlóan

$$\beta''(B) = 1, \qquad A < x < B,$$

és

$$\beta(B) = 1, \qquad \beta(A) = 0.$$

2. Lemma. Legyen $m_i(x) = E_i\tau^x_{A,B}(\omega)$, $A \le x \le B$. Akkor

$$m_1(x) = \frac{1}{\rho}\left\{\frac{(e^B - e^{A+B-x})(B-A)}{e^B - e^A} + A - x\right\}$$

$$m_0(x) = \frac{1}{\rho}\left\{\frac{(e^B - e^x)(B-A)}{e^B - e^A} - B + x\right\}$$

Bizonyitás. Elég megjegyezni (l. Dünkin), hogy $m_i(x)$ kielégiti a

$$\frac{\tau^2}{2\sigma^2} m_i''(x) + (-1)^{1-i}\frac{\tau^2}{2\sigma^2} m_i'(x) = -1$$

egyenletet $m(A) = m(B) = 0$ feltétellel.

3. Lemma. Legyen $w(t)$ a standard Brown-mozgás folyamat, $\mathcal{F}_t$-re nézve mérhető. Minden $\tau = \tau(\omega)$ Markov pontra, amelyre $E\tau < \infty$ igaz a Wald féle azonosság:

$$Ew(\tau) = 0, \quad Ew^2(\tau) = E\tau.$$

Bizonyitás. Legyen $\chi_{(\tau > t)} = \begin{cases} 1 & \text{ha} \quad t < \tau, \\ 0 & \text{ha} \quad t \ge \tau. \end{cases}$

Mivel 1 valószinüséggel $\int\limits_0^\infty \chi^2_{(\tau \ge t)}\,dt = \tau(\omega) < \infty$,

értelmezett a $\int\limits_0^\infty \chi_{(\tau \ge t)}\,dw(t)$ integrál és $w(\tau,\omega) = \int\limits_0^\infty \chi(\tau \ge t)\,dw(t)$.

Mivel

$$\int\limits_0^\infty E\chi^2_{(\tau \ge t)}\,dt = \int\limits_0^\infty \overbrace{P(\tau \ge t)}^{[1-F(t)]}\,dt = E\tau < \infty,$$

a sztohasztikus integrálok tulajdonságai miatt

$$Ew(\tau) = E\int \chi(\tau \ge t)\,dw(t) = 0,$$

$$Ew^2(\tau) = E\left[\int \chi(\tau \ge t)\,dw(t)\right]^2 = \int\limits_0^\infty E\chi^2(\tau \ge t)\,dt = E\tau,$$

ami az állitás bizonyitása.

---

* Dünkin, E. B.: Markovszkije processzü (1963, Moszkva).

Megjegyzések.

1. Az $E\tau<\infty$ feltétel szükséges az állításhoz. Legyen ugyanis $\tau=\inf\{t\geq 0: w(t)=1\}$. Kimutatható az 1., 2. lemmák alapján, hogy $E\tau=\infty$, $P(\tau<\infty)=1$, ahonnan
$$1=Ew(\tau)^2\neq E\tau=\infty .$$

2. Legyen $\tau=\inf\{t\geq 0: |w(t)|=A\}$, ahol $A<\infty$. Akkor $E\tau=A^2$. Ugyanis, ha $\tau_N=\min(\tau,N)$ akkor a 3. lemma szerint $Ew(\tau_N)^2=E\tau_N$, ahonnan $E\tau_N\leq A^2$ és igy $E\tau=\lim_{N\to\infty}E\tau_N\leq A^2<\infty$.
A 3. lemma ujbóli alkalmazásával $Ew(\tau)^2=E\tau$. Mivel $P(\tau<\infty)=1$ ezért $E\tau=Ew^2(\tau)=A^2$.

3. Legyen $\tau=\inf\{t\geq 0: |w(t)|=a\sqrt{t+b}\}$, ahol $0<b<\infty$, $0\leq a<1$.
Akkor
$$E\tau=\frac{a^2b}{1-a^2}$$

Bizonyitás. Legyen $\tau_N=\min(\tau,N)$ (N egész), akkor ~~$E\tau_N=$~~ $=Ew^2(\tau_N)\leq a^2E(\tau_N+b)$, azaz $E\tau_N\leq\frac{a^2b}{1-a^2}$ és igy
$$E\tau=\lim_{N\to\infty}E\tau_N\leq\frac{a^2b}{1-a^2}<\infty .$$
Az iterált logaritmus tételből adódik, hogy $P(\tau<\infty)=1$. Ezért
$$E\tau=Ew^2(\tau)=\int_{(\tau<\infty)}w^2(\tau)\,dP=a^2\int_{(\tau<\infty)}[\tau+b]\,dP=a^2[E\tau+b],$$
ahonnan adódik az állítás.

4. Lemma. Legyen $w(t)$ a standard Brown-mozgás folyamat. Legyen τ -Markov pont $\mathcal{F}_t$-re nézve, mégpedig $P\{\tau\leq N\}=1$, $N<\infty$, akkor tetszőleges $-\infty<\lambda<\infty$ értékre
$$E\exp\left\{\lambda w(\tau)-\frac{\lambda^2}{2}\tau\right\}=1 .$$

Bizonyitás. Az
$$\eta_t=\exp\left\{\lambda w(t)-\frac{\lambda^2}{2}t\right\}$$
folyamat $\{\eta_t,\mathcal{F}_t,P\}$ martingál és $E\eta_t=1$. Ismert martingál tulajdonságokból adódik, hogy (Doob*: VII.fejezet, II.7.tétel)
$$E\eta_t\leq 3\sup_{t\leq N}E\eta_t=3<\infty .$$
Nyilvánvaló továbbá, hogy
$$\lim_{t\to\infty}\int_{(\tau>t)}\eta_t\,dP=0,\quad \text{mivel}\quad P\{\tau\leq N\}=1,\ N<\infty .$$
Alkalmazva az $Ex_\tau=Ex_0$ teljesülésének feltételét (Sirjajev idézett könyve, 42.o.), mely szerint elegendő, hogy

* Doob, J. L.: Stochastic processes (1953, Wiley).

$$E|x_\tau| < \infty,$$

$$\lim_{t\to\infty}\int_{\{\tau>t\}} x_t\, dP = 0 .$$

A lemma állítása nem más, mint a fundamentális azonosság a szekvenciális analizisben.

A tétel akkor is igaz marad - bizonyos esetekben -, ha τ nem korlátos megállási pont. Például

$$\tau = \inf\{t \ge 0 : |w(t)| = a\sqrt{t+b}\},\quad 0 < b < \infty,\quad 0 \le a \le 1 .$$

<u>Az alaptétel bizonyitása</u>

Először megmutatható, hogy tetszőleges $\delta = \{\tau(\omega), d(\omega)\}$ szabályra

$$E_0\tau \ge \frac{\omega(\beta,\alpha)}{\rho},\qquad E_1\tau \ge \frac{\omega(\alpha,\beta)}{\rho} .$$

$\theta = \theta(\omega)$ nem megfigyelhető. Megfigyelés alatt a $\xi(t) = r\theta(\omega)t+\sigma.w(t)$ áll.

A φ_t definiciója és a Wald féle azonosság alapján

$$E_1 \ln\varphi_\tau(\omega) = E_1 \ln\frac{dP_1}{dP_0} = E_1\left\{\frac{r}{\sigma^2}\xi(\tau) - \frac{r^2}{2\sigma^2}\tau\right\} = E_1\left\{\frac{r^2}{2\sigma^2}\tau + \frac{r}{\sigma^2}w(\tau)\right\} =$$

$$= \frac{r^2}{2\sigma^2}E_1\tau = \rho E_1\tau .$$

Másrészt

$$E_1 \ln\varphi_\tau(\omega) = -E_1\ln\frac{dP_0}{dP_1} = -\int_{\{\omega:d(\omega)=1\}}\ln\frac{dP_0}{dP_1}dP_1 - \int_{\{\omega:d(\omega)=0\}}\ln\frac{dP_0}{dP_1}dP_1 =$$

$$= -P_1\{d(\omega)=1\}\int_\Omega \ln\frac{dP_0}{dP_1}dP_1(\omega|d(\omega)=1) - P_1(d(\omega)=0)\int_\Omega \ln\frac{dP_0}{dP_1}dP_1(\omega|d(\omega)=0) \ge$$

$$\ge -P_1(d(\omega)=1)\ln\int\frac{dP_0}{dP_1}dP_1(\omega|d(\omega)=1) - P_1(d(\omega)=0)\ln\int\frac{dP_0}{dP_1}dP_1(\omega|d=0)$$

(a Jensen egyenlőtlenség szerint

ha $\eta \ge 0$, akkor $\ln E\eta \ge E\ln\eta$).

folytatva az egyenlőtlenséget

$$= -P_1(d=1)\ln\left[\frac{1}{P_1(d=1)}\int_{(\omega:d=1)}\frac{dP_0}{dP_1}dP_1\right] - P_1(d=0)\ln\left[\frac{1}{P_1(d=0)}\int_{\{\omega:d=0\}}\frac{dP_0}{dP_1}dP_1\right]$$

$$= -P_1(d=1)\ln\frac{P_0(d=1)}{P_1(d=1)} - P_1(d=0)\ln\frac{P_0(d=0)}{P_1(d=0)} \ge$$

$$\ge -(1-\alpha)\ln\frac{\beta}{1-\alpha} - \alpha\ln\frac{1-\beta}{\alpha} = (1-\alpha)\ln\frac{1-\alpha}{\beta} + \alpha\ln\frac{\alpha}{1-\beta} = \omega(\alpha,\beta)$$

Hasonlóan bizonyitható az $E_0\tau$ -ra vonatkozó egyenlőtlenség. Legyen $(\tilde\tau,\tilde d) = \tilde\delta$ a tételben meghatározott döntési szabály (azaz

$$\tilde d = \begin{cases} 1 & \lambda_{\tilde\tau} \ge \tilde B \\ 0 & \lambda_{\tilde\tau} \le \tilde A \end{cases})$$

$$P_1(\tilde d=0) = \alpha(0) = \frac{e^{\tilde A}(e^{\tilde B}-1)}{e^{\tilde B}-e^{\tilde A}} = \alpha \qquad \text{(az 1. lemma alapján)}$$

$$P_0(\tilde{d}=1)=\beta(0)=\frac{1-e^{\tilde{A}}}{e^{\tilde{B}}-e^{\tilde{A}}}=\beta$$

A 2.lemma alapján

$$E_1\tilde{\tau}=\frac{1}{\varrho}\left\{\frac{(e^{\tilde{B}}-e^{\tilde{A}+\tilde{B}})(\tilde{B}-\tilde{A})}{e^{\tilde{B}}-e^{\tilde{A}}}+\tilde{A}\right\}=\frac{1}{\varrho}\,\frac{\tilde{B}e^{\tilde{B}}(1-e^{\tilde{A}})+e^{\tilde{A}}\tilde{A}(e^{\tilde{B}}-1)}{e^{\tilde{B}}-e^{\tilde{A}}}=\frac{1}{\varrho}\,\omega(\alpha,\beta)$$

és

$$E_0\tau=\frac{1}{\varrho}\left\{\frac{(e^{\tilde{B}}-1)(\tilde{B}-\tilde{A})}{e^{\tilde{B}}-e^{\tilde{A}}}-\tilde{B}\right\}=\frac{1}{\varrho}\,\frac{\tilde{B}(e^{\tilde{A}}-1)+\tilde{A}(1-e^{\tilde{B}})}{e^{\tilde{B}}-e^{\tilde{A}}}=\frac{1}{\varrho}\,\omega(\beta,\alpha).$$

A most bebizonyitott tétel alapján a szekvenciális likelihood hányados próbán alapuló eljárást összehasonlithatjuk a klasszikus hipotézis vizsgálati eredményekkel, ahol ugyancsak feltesszük, hogy teljesülnek a

$$P_1(d_t=0)\le\alpha,\quad P_0(d_t=1)\le\beta.$$

feltételek. A Neyman-Pearson lemma alapján a döntési szabály

$$d_{t,(\alpha,\beta)}(\omega)=\begin{cases}1 & \lambda_t(\alpha,\beta)(\omega)\ge h(\alpha,\beta),\\ 0 & \lambda_t(\alpha,\beta)(\omega)<h(\alpha,\beta),\end{cases}$$

alaku, ahol

$$t(\alpha,\beta)=\frac{(c_\alpha+c_\beta)^2}{2\varrho},\qquad h(\alpha,\beta)=\frac{c_\beta^2-c_\alpha^2}{2}.$$

A c_γ együttható meghatározása a következőképpen történik:

$$\frac{1}{\sqrt{2\pi}}\int_{c_\gamma}^{\infty}e^{-x^2/2}\,dx=\gamma,\quad(0\le\gamma<1).$$

Ugyanis

$$P_0(d_t=1)=P_0(\lambda_t\ge h)=P_0\left\{\frac{r}{\sigma^2}\left[\xi_t-\frac{r}{2}t\right]\ge h\right\}=$$

$$=P\left\{\frac{r}{\sigma}w(t)\ge h+\frac{r^2}{2\sigma^2}t\right\}=P\left\{\frac{w_t}{\sqrt{t}}\ge\frac{h+\varrho t}{\frac{r}{\sigma}\sqrt{t}}\right\}=\phi\left(\frac{h+\varrho t}{\frac{r}{\sigma}\sqrt{t}}\right)=s$$

és

$$P_1(d_t=0)=1-\phi\left(\frac{h-\varrho t}{\frac{r}{\sigma}\sqrt{t}}\right)=\alpha.$$

Innen adódik, hogy

$$c_\beta=\frac{h+\varrho t}{\frac{r}{\sigma}\sqrt{t}},\qquad\frac{h-\varrho t}{\frac{r}{\sigma}\sqrt{t}}=-c_\alpha.$$

Igy $\alpha+\beta<1$ esetén

$$\frac{E_0\tilde{\tau}}{t(\alpha,\beta)}=2\,\frac{\omega(\beta,\alpha)}{(c_\alpha+c_\beta)^2}$$

$$\frac{E_1\tilde{\tau}}{t(\alpha,\beta)}=2\,\frac{\omega(\alpha,\beta)}{(c_\alpha+c_\beta)^2}$$

Ha $\alpha,\beta\le 0{,}03$

$$E_0\tilde{\tau}\le\frac{17}{30}t(\alpha,\beta),\qquad E_1\tilde{\tau}\le\frac{17}{30}t(\alpha,\beta)$$

Ha $\alpha=\beta$ Ajvazjan bizonyitotta a következő határérték összefüggéseket:

$$\lim_{\alpha\downarrow 0}\frac{E_0\tilde{\tau}}{t(\alpha,\alpha)}=\lim_{\alpha\downarrow 0}\frac{E_1\tilde{\tau}}{t(\alpha,\alpha)}=1/4 .$$

Feladatok

1. Jelentse standard Brown-mozgás folyamat esetén τ_a az a szint első elérésének időpontját. Lássuk be a következő összefüggéseket

$$P\{\tau_a\le t\mid w(0)=x\}=P\{\tau_0\le t\mid w(0)=x-a\},$$
$$P\{\tau_y-\tau_x\le t\mid w(0)=0\}=P\{\tau_{y-x}\le t\mid w(0)=0\}\qquad (y\ge x\ge 0),$$
$$P\{\tau_a\le a^2t\mid w(0)=0\}=P\{\tau_1\le t\mid w(0)=0\}\quad,\qquad (a>0).$$

Igaz továbbá, hogy

$$P\{\tau_a<t\mid w(0)=0\}=\frac{a}{\sqrt{2\pi}}\int_0^t s^{-3/2}e^{-\frac{a^2}{2s}}\,ds\ ,\qquad (0\le t<\infty).$$

2. Mutassuk meg, hogy $a>0$ esetén

$$P\{\max_{0\le s\le t} w(s)\ge a\mid w(0)=0\}=\frac{2}{\sqrt{2\pi t}}\int_a^\infty e^{-\frac{y^2}{2t}}\,dy .$$

3. Igazak a következő összefüggések, melyeknek bizonyitása megtalálható Gihman-Szkorohod: "Bevezetés a sztohasztikus folyamatok elméletébe" c. könyvében,

$$P\{\max_{0\le s\le t} w(s)\ge a,\ w(t)\in[c,d]\}=$$
$$=\frac{1}{\sqrt{2\pi t}}\int_{\max(c,a)}^{\max(d,a)}e^{-\frac{x^2}{2t}}\,dx+\frac{1}{\sqrt{2\pi t}}\int_{\max(2a-d,a)}^{\max(2a-c,a)}e^{-\frac{x^2}{2t}}\,dx$$

és ha $a>0$ továbbá $[c,d]\subseteq[-a,a]$, akkor

$$P\{\max_{0\le s\le t}|\xi(s)|<a,\ \xi(t)\in[c,d]\}=$$
$$=\int_c^d\frac{1}{\sqrt{2\pi t}}\sum_{k=-\infty}^{\infty}(-1)^k e^{-\frac{(x-2ka)^2}{2t}}\,dx .$$

4. T.W.Anderson vizsgálta* a következő határokból való első kilépés feladatát. A (t,x) sikon legyenek megadva a

$$x_1(t)=\gamma_1+\delta_1 t,\qquad x_2(t)=\gamma_2+\delta_2 t$$

egyenesek, ahol $\gamma_1>0$, $\gamma_2<0$, mindaddig, amig $x_1(t)\ge x_2(t)$, $w(0)=0$. Legyen $\tau_1=\tau_1(\omega)$ a felső határon való első kilépés időpontja, $\tau_2=\tau_2(\omega)$ az alsó határon való első kilépés időpontja. Igaz a következő összefüggés

$$P\{\tau_1\le\min(t,\tau_2)\}=1-\Phi\left(\frac{\delta_1 t+\gamma_1}{\sqrt{t}}\right)+$$

* A modification of the sequential probability ratio test to reduce the sample size. Annals Math. Stat., 31, (1960) 165-198.

$$
\begin{aligned}
&+\sum_{k=1}^{\infty}\Big\{e^{-2[k\gamma_1-(k-1)\gamma_2][k\delta_1-(r-1)\delta_2]}\,\Phi\Big(\frac{\delta_1 t+2(k-1)\gamma_2-2(k-1)\gamma_1}{\sqrt{t}}\Big)-\\
&-e^{-2[k^2(\gamma_1\delta_2+\gamma_2\delta_2)-k(k-1)\gamma_1\delta_2-k(k+1)\gamma_2\delta_1]}\,\Phi\Big(\frac{\delta_1 t+2k\gamma_2-(2k-1)\gamma_1}{\sqrt{t}}\Big)-\\
&-e^{-2[(k-1)\gamma_1-k\gamma_2][(k-1)\delta_1-k\delta_2]}\Big[1-\Phi\Big(\frac{\delta_1 t-2k\gamma_2+(2k-1)\gamma_1}{\sqrt{t}}\Big)\Big]\\
&+e^{-2[k^2(\gamma_1\delta_1+\gamma_2\delta_2)-(k-1)k\gamma_2\delta_1-k(k+1)\gamma_1\delta_2}\Big[1-\Phi\Big(\frac{\delta_1 t+(2k+1)\gamma_1-2k\gamma_2}{\sqrt{t}}\Big)\Big]\Big\},
\end{aligned}
$$

ahol $\phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{u^2}{2}}\,du$ a standard normális eloszlás.

5. Legyen τ_a^* a $0 \le s \le t$ intervallumban a $w(s) = a$ egyenlet megoldásai közül a legnagyobb (az a szint utolsó elérése).
Igaz a következő állitás

$$P\{\tau_a^* = s \mid w(0)=a\} = \frac{2}{\pi}\arcsin\sqrt{\frac{s}{t}}, \qquad (0 \le s \le t).$$

Hasonló (de más bizonyitással) összefüggés igaz a $w(0) = 0$ esetén a $[0,t]$ intervallumban a 0 szint feletti tartózkodási időre. Ha τ jelöli azt az időtartamot, amig $w(u) \ge 0$ $(0 \le u \le t)$, akkor

$$P\{\tau \le s \mid w(0)=0\} = \frac{2}{\pi}\arcsin\sqrt{\frac{s}{t}}, \qquad (0 \le s \le t).$$

6. Állitsuk elő a Brown-mozgás folyamat közelitését a következő két módon:

a) Legyenek ξ_{mn}, $m=0,\dots,n-1$ azonos eloszlásu független valószinüségi változók

$$E\,\xi_{mn}=0,\quad E\,\xi_{mn}^2=\frac{1}{n} \qquad (n=1,2,\dots).$$

Tekintsük a

$$\xi_n(t)=\sum_{m=0}^{\ell-1}\xi_{mn}+(nt-k)\,\xi_{kn}$$

folytonos folyamatot, ahol $k=[nt]$ az nt egész része (ha $0 \le t \le \frac{1}{n}$ definició szerint $\xi_n(t) = nt\,\xi_{no}$). A $\xi_n(t)$ folyamat által generált P_n mérték a folytonos függvények terén gyengén tart a Wiener mértékhez.

b) Legyen

$$w(t)=\sum_{k=0}^{\infty}\xi_k\,\varphi_k(t)$$

ahol ξ_k független Gauss sorozat

$$E\,\xi_k=0,\quad E\,\xi_k^2=\frac{1}{[\frac{\pi}{2}(2k+1)]^2} \qquad (k=0,1,2,\dots),$$

$$\varphi_k(t)=\sin\Big[\frac{\pi}{2}(2k+1)t\Big].$$

A $w(t)$ folyamat Brown-mozgás.

Készitsünk mindkét előállitásra FORTRAN programot.

5. A meghibásodás gyors felismerése (a "riasztás" feladata)[1]

A hipotézisvizsgálat feladatában a $\xi_1, \xi_2, \ldots$ valószinüségi változók egy és ugyanazon (bár ismeretlen) valószinüségeloszlással rendelkeznek. A gyakorlatban gyakran találkozunk a következő tipusu feladattal. A megfigyelt valószinüségi változók valószinüségi jellemzői egy véletlen $\theta = \theta(\omega)$ időpillanatban megváltoznak (a "meghibásodás" időpontja). Az alábbiakban néhány feladat megfogalmazást mutatunk be, és vázoljuk azokat a lehetőségeket, melyek segitségével a megoldások is nyerhetők.

Előbb az időben diszkrét eset megfogalmazását tekintsük. Legyen $(\Omega, \mathcal{F})$ mérhető tér és P_π, $0 \le \pi \le 1$ valószinüségi mértékek egy serege ezen a téren. A valószinüségi téren legyenek értelmezve
1. a $\theta = \theta(\omega)$ nemnegativ egészértékü valószinüségi változó és
2. a $\xi_1, \xi_2, \ldots$ valószinüségi változó sorozat.

Feltesszük, hogy minden $0 \le \pi \le 1$ értékre

$$P_\pi(\theta = 0) = \pi, \qquad P(\theta = n \mid \theta > 0) = (1-p)^{n-1} p,$$

ahol $0 < p < 1$ ismert és független π-től.

Feltesszük továbbá, hogy ha $\theta = 0$ akkor a $\xi_1, \xi_2, \ldots$ változók függetlenek, azonos eloszlásuak, sürüségfüggvényük $p_1(x)$. Ha $\theta = i$, akkor a $\xi_1, \xi_2, \ldots, \xi_{i-1}, \xi_i, \ldots$ változók továbbra is függetlenek összességükben, mégpedig $\xi_1, \ldots, \xi_{i-1}$ azonos eloszlásuak $p_0(x)$ sürüségfüggvénnyel, $\xi_i, \xi_{i+1}, \ldots$ ugyancsak azonos eloszlásuak $p_1(x)$ sürüségfüggvénnyel. A θ geometria eloszlására vonatkozó feltételt az egyszerüség kedvéért vezettük be. A szóban megfogalmazott feltevések a következőt jelentik. Tetszőleges $n \ge 1$ esetén

$$P_\pi(x_1, \ldots, x_n) = P_\pi(\xi_1 \le x_1, \ldots, \xi_n \le x_n)$$

valószinüségek olyanok, hogy

$$P_\pi(x_1, \ldots, x_n) = \pi P_1(x_1, \ldots, x_n) + (1-\pi) \sum_{i=0}^{n-1} p(1-p)^i P_0(x_1, \ldots, x_i) P_1(x_{i+1}, \ldots, x_n) +$$
$$+ (1-\pi)(1-p)^n P_0(x_1, \ldots, x_n).$$

és

$$P_k(x_1, \ldots, x_n) = P_k(x_1) \cdots P_k(x_n) \qquad (k = 0, 1).$$

Az általánosság megszoritása nélkül feltehető, hogy léteznek a

[1] Az orosznyelvü irodalomban a feladatot "razladka", az angolnyelvü irodalomban "disorder" cimen szokták idézni.

$p_k(x)$ $(k = 0,1)$ sürüségfüggvények.

Általánosan megfogalmazzuk a riasztási feladatot. Legyen $\xi_0(\omega) \equiv 0$, $\mathcal{F}_n = \sigma\{(\omega,\pi):\pi, \xi_0(\omega), \xi_1(\omega), \dots, \xi_n(\omega)\}$, $n \geq 0$, $\mathfrak{m} = \{\tau\}$ a Markov megállási pontok halmaza $\{\mathcal{F}_n\}$ -re nézve.

Minden $\tau \geq 0$-nak megfelel egy "hibás riasztási" valószinüség $P_\pi(\tau < \theta)$ és "átlagos késés" $E_\pi(\tau - \theta | \tau \geq \theta)$, a $\tau \geq \theta$ feltétel mellett. Természetes kivánság olyan megállási pont keresése, amelyre mind a valószinüség, mind az átlagos késés kicsi lesz.

A " riasztási" feladat minőségellenőrzési vagy folyamatszabályozási és termelésellenőrzési jelentősége közvetlenül látható. Az alábbiakban egy, a számitógépek operációs rendszere szempontjából fontos feladatot kivánunk bemutatni.

Jelölje $\xi_1, \xi_2, \dots$ a T_1, $2T_1-T_1$,..... időintervallumban (ahol T_1 egy fix konstans, amely a multiprogramozásu számitógép kvantum idejénél nagyobb) fellépő I/O megszakitások számát (azaz a külső memóriákhoz, perifériákhoz való fordulások számát). A multiprogramozásu gépekben igen nagy jelentőségü a prioritások adása. A feladat annak megállapitása, hogy szükség van-e a prioritási rendszer megváltoztatására mivel a ξ_i változók eloszlása változott. A ξ_i értékek regisztrálása a gépben folyamatosan történik és az eloszlás változása ugyanazon a programon belül is történhet. Hasonló tipusu feladat a lapolási hibák számának eloszlásában történő változás, valamint a Denning által bevezetett munka mezők nagyságában történő változás gyors jelzése is. Ezeknek a feladatoknak a részletes leirására egy külön fejezetben kivánunk kitérni, egy más tipusu probléma megfogalmazása kapcsán (lásd 7.§.).

Variációs feladat. Legyen π fix, $0 \leq \pi < 1$, és $\mathfrak{m}(\alpha,\pi) \leq \mathfrak{m}$ azon megállási pontok összessége (Markov pontok) $\tau = \tau(\omega,\pi)$, melyekre

(1) $P_\pi(\tau < \theta) \leq \alpha$,

ahol α megadott konstans, $0 \leq \alpha \leq 1$. A $\tilde{\tau} = \tilde{\tau}(\omega,\pi) \in \mathfrak{m}(\alpha,\pi)$ megállást optimálisnak nevezzük, ha

(2) $E_\pi(\tilde{\tau} - \theta | \tilde{\tau} \geq \theta) \leq E_\pi(\tau - \theta | \tau \geq \theta)$

minden $\tau \in \mathfrak{m}(\alpha,\pi)$ megállásra. Tehát adott hibás "riasztási" valószinüség mellett az átlagos késést kivánjuk minimalizálni. Később feltételt adunk az optimális megállás létezésére.

Bayes féle feladat. Legyen

$$\rho(\pi,\tau) = P_\pi(\tau < \theta) + c\,E_\pi(\tau - \theta | \tau \geq \theta)\,P_\pi(\tau \geq \theta)$$

ahol $c > 0$, $\varrho(\pi) = \inf_{\tau \in \mathfrak{M}} \varrho(\pi,\tau)$.

A τ^* π-Bayes féle megoldás, ha adott π-re

(3) $\varrho(\pi,\tau^*) = \varrho(\pi)$.

Ha (3) minden $0 \le \pi \le 1$-re igaz, τ^* Bayes féle megoldás.
Igaz a következő állitás.

__1. Tétel.__ Legyen $c > 0$, $p > 0$ és

$$\pi_n(\omega,\pi) = P_\pi\{\theta(\omega) \le n \mid \mathcal{F}_n\}, \quad n \ge 1,$$

$(\pi_0(\omega,\pi) = \pi)$, a meghibásodás aposteriori valószinüsége.
A

$$\tau^*(\omega,\pi) = \inf\{n > 0 : \pi_n(\omega,\pi) \ge A^*\},$$

ahol A^* konstans, Bayes féle megoldás.

__2. Tétel.__ Legyen $0 \le \alpha \le 1$, $0 \le \pi \le 1$, $p > 0$ és $\mathfrak{M}(\alpha,\pi)$ azon τ Markov pontok összessége, melyre $P_\pi(\tau < \theta) \le \alpha$. Akkor, ha minden π-re $\alpha_A(\pi)$ folytonos A-ban,

(ahol $\tau_A(\omega,\pi) = \inf\{n \ge c : \pi_n(\omega,\pi) \ge A\}$

$$\alpha_A(\pi) = P_\pi\{\tau_A < \theta\} = E_\pi\{1 - \pi_{\tau_A}\}$$

akkor a

$$\tilde{\tau} = \tau^*_{c_\alpha} = \inf\{n \ge 0 : \pi_n \ge \tilde{A}\}$$

megállási szabály (ahol $\tilde{A} = A^*(c_\alpha)$ a c veszteség esetén a Bayes féle feladatban szereplő A^*, $c^*(A^*)$ az a minimális c, melyre $A^*(c) = A^*$) optimális abban az értelemben, hogy minden $\tau \in \mathfrak{M}(\alpha;\pi)$-re

$$E_\pi(\tilde{\tau} - \theta \mid \tilde{\tau} \ge \theta) \le E_\pi(\tau - \theta \mid \tau \ge \theta).$$

Adott α esetén megtalálni a pontos $\tilde{A} = A^*(c_\alpha)$ értéket nehéz, de $\tilde{A} \le 1 - \alpha$ igaz.

A fenti állitások megfogalmazása és a "riasztás" feladatának megoldása Kolmogorovtól és Sirjajevtől származik. A problémakör irodalma és a bizonyitások megtalálható Sirjajev idézett könyvében[1].
A "riasztás" feladatának olyan megoldása, amely az átlagos riasztási idő minimalizálása helyett adott (fix)T idő alatt 1 valószinüséggel irja elő a riasztást (a gyakorlatban ez a megfogalmazás sokszor szükségszerü) matematikailag nem megoldott. Az ilyenirányu erőfeszi-

[1] További irodalom: R. Lipcer-A. Sirjajev, Sztatisztika szlucsajnüh processzov (Moszkva, 1974.).

tések gyakorlatilag jól működő algoritmikus eljárások - és az ellenőrzés több komponensre kiterjedő megszervezésére vonatkoznak.

Az 1. Tétel bizonyítása. A Bayes formula alapján P_π majdnem mindenütt

$$(4)\quad \pi_{n+1}(\omega,\pi)=\frac{\pi_n p_1(\xi_{n+1})+(1-\pi_n)p\cdot p_1(\xi_{n+1})}{\pi_n p_1(\xi_{n+1})+(1-\pi_n)p\cdot p_1(\xi_{n+1})+(1-\pi_n)(1-p)p_0(\xi_{n+1})}$$

ahol

$$\pi_n=\pi_n(\omega,\pi),\quad \xi_{n+1}=\xi_{n+1}(\omega).$$

Megmutatható, hogy $\{\pi_n(\omega,\pi),\mathcal{F}_n,P_\pi\}$ egy Markov sorozat.

Minden $0\le\pi\le 1$-re, $\tau\in\mathfrak{M}$ -re

$$(5)\quad \rho(\pi,\tau)=P_\pi(\tau<\theta)+c\,E_\pi(\tau-\theta/\tau\ge\theta)\,P_\pi(\tau\ge\theta)=$$

$$=E_\pi(1-\pi_\tau)+c\,E_\pi\max(\tau-\theta,0)\quad.$$

Az utóbbi tagot próbáljuk másként meghatározni. $n\ge 0$ esetén

$$E_\pi[\max(n-\theta,0)|\mathcal{F}_n]=\sum_{k=0}^{n}(n-k)P_\pi(\theta=k|\mathcal{F}_n)=\sum_{k=0}^{n-1}P_\pi(\theta\le k\mid\mathcal{F}_n)=$$

$$=\sum_{k=0}^{n-1}[P_\pi(\theta\le k\mid\mathcal{F}_n)-P_\pi(\theta\le k\mid\mathcal{F}_k)]+\sum_{k=0}^{n-1}P_\pi(\theta\le k\mid\mathcal{F}_n)=$$

$$\sum_{k=0}^{n-1}[P_\pi(\theta\le k|\mathcal{F}_n)-P_\pi(\theta\le k|\mathcal{F}_k)]+\sum_{0}^{n-1}{}'\,\pi_k$$

Legyen

$$\Psi_n(\omega,\pi)=\sum_{k=0}^{n-1}[P_\pi(\theta\le k|\mathcal{F}_n)-P_\pi(\theta\le k|\mathcal{F}_k)]=$$

$$=\sum_{k=0}^{n-1}[P_\pi(\theta>k|\mathcal{F}_n)-P_\pi(\theta>k|\mathcal{F}_k)]$$

A $\{\Psi_n,\mathcal{F}_n,P_\pi\}$ – martingál ($n\ge 0$, $0\le\pi\le 1$). Ugyanis nyilvánvalóan $E|\Psi_n|<\infty$ és

$$\Psi_{n+1}=\sum_{k=0}^{n}[P_\pi(\theta\le k\mid\mathcal{F}_{n+1})-P_\pi(\theta\le k\mid\mathcal{F}_k)]$$

ahonnan

$$E(\Psi_{n+1}\mid\mathcal{F}_n)=\Psi_n$$

Mivel

$$|\Psi_n(\omega,\pi)|\le\sum_{k=0}^{\infty}P_\pi(\theta>k|\mathcal{F}_n)+\sum_{k=0}^{\infty}P_\pi(\theta>k|\mathcal{F}_k),$$

ahol ($p>0$ miatt)

$$E_\pi\sum_{k=0}^{\infty}P_\pi(\theta>k|\mathcal{F}_n)=\sum_{k=0}^{\infty}P_\pi(\theta>k)=E_\pi\theta<\infty$$

és

$$E_\pi\sum_{k=0}^{\infty}P_\pi(\theta>k|\mathcal{F}_k)=E_\pi\theta<\infty$$

a $\tau(\omega,\pi) \in \mathfrak{m}$ megállásra

(6) $$\lim_{n\to\infty} \int_{(\tau>n)} |\psi_n(\omega,\pi)| \, dP_\pi = 0 .$$

Nyilvánvaló továbbá, hogy

(7) $$E_\pi |\psi_\tau(\omega,\pi)| < \infty, \qquad \tau \in \mathfrak{m} .$$

A (6) és (7) feltételek alapján tetszőleges $\tau \in \mathfrak{m}$ -re

$$E_\pi \psi_\tau(\omega,\pi) = E_\pi \psi_0(\omega,\pi) = 0, \qquad 0 \le \pi \le 1 .$$

Ezért, ha $\tau \in \mathfrak{m}$

$$\rho(\pi,\tau) = P_\pi(\tau<\theta) + c E_\pi(\tau-\theta \mid \tau \ge \theta) P_\pi(\tau \ge \theta) =$$
$$= E_\pi \Big\{(1-\pi_\tau) + c\sum_{k=0}^{\tau-1} \pi_k + c\psi_\tau(\omega,\pi)\Big\} = E_\pi \Big\{(1-\pi_\tau) + c\sum_{k=0}^{\tau-1} \pi_k\Big\}$$

és

(8) $$\rho(\pi) = \inf_{\tau\in\mathfrak{m}} E_\pi\Big\{(1-\pi_\tau) + c\sum_{k=0}^{\tau-1} \pi_k\Big\} .$$

A $\{\pi_n, F_n, P_\pi\}$ sorozat szubmartingál: $E_\pi[\pi_{n+1} \mid \mathcal{F}_n] \ge \pi_n$. (Nyilván $E_\pi\{P(\theta \le n+1 \mid F_{n+1}) \mid \mathcal{F}_n\} = P(\theta \le n+1 \mid \mathcal{F}_n) \ge P(\theta \le n \mid F_n) = \pi_n$.)

A "martingál" konvergencia tétel értelmében 1 valószinüséggel $\lim \pi_n$ létezik. Továbbá

$$\lim_{n\to\infty} \pi_n \le 1 \quad \text{és} \quad \lim_{n\to\infty} E\pi_n = 1 .$$

A Fatou lemma alapján

$$1 = \lim E\pi_n \le E \lim \pi_n$$

és igy tetszőleges π-re

$$\lim \pi_n = 1 ,$$

ahonnan

$$\lim_{n\to\infty} \sum_{k=0}^{n} \pi_k = \infty \qquad (P_\pi \, m.m.) .$$

Ismét mindent az $\Omega' = \Omega \times [0,1]$, $\mathcal{F}' = \mathcal{F} \times B[0,1]$, $\pi_n'(\omega') = \pi_n(\omega,\pi)$ kiterjesztett téren vizsgálva egy Π' Markov folyamatot kapunk. $\mathfrak{m}$ legyen az $\{F_n'\}$ -re nézve Markov pontok összessége.

(9) $$\rho(\pi) = \inf_{\tau\in\mathfrak{m}} E_\pi\Big\{(1-\pi_\tau) + c\sum_{k=0}^{\tau-1} \pi_k\Big\} = \inf_{\tau\in\mathfrak{m}} E_\pi'\Big\{(1-\pi_\tau) + c\sum_{k=0}^{\tau-1} \pi_k\Big\} .$$

$\mathfrak{m}^\pi$ legyen az olyan $\tau \in \mathfrak{m}$ pontok összessége, amelyre minden n-re $\{\omega' : \tau(\omega') = n\} \in \mathcal{F}_n^\pi$, ahol $\mathcal{F}_n^\pi = \sigma\{\omega' : \pi_0, \dots, \pi_n\}$

Általános tételekből adódik, hogy (Sirjajev, II.fej.,7.3.§)

(1o) $$\rho(\pi) = \inf_{\tau\in\mathfrak{m}} E_\pi'\Big\{(1-\pi_\tau) + c\sum_{k=0}^{\tau-1} \pi_k\Big\} = \inf_{\tau\in\mathfrak{m}^\pi} E_\pi'\Big\{(1-\pi_\tau) + c\sum_{k=0}^{\tau-1} \pi_k\Big\}$$

Tehát a Bayes féle feladat visszavezethető egy Markov folyamat optimális megállitására (ehhez Sirjajev II.fej.8.§-a használható).

Legyen R^{π} azon $\tau \in \mathfrak{M}^{\pi}$ Markov pontok összessége, melyre

$$E_{\pi}\sum_{k=0}^{\tau-1} \pi_k < \infty, \quad 0 \le \pi \le 1.$$

Akkor (Sirjajev II.fej.3.§.)

$$\rho(\pi) = \inf_{\tau \in R^{\pi}} E'_{\pi}\left\{(1-\pi_{\tau}) + c\sum_{k=0}^{\tau-1}\pi_k\right\},$$

azaz elegendő csak erre az osztályra szoritkozni.

Legyen $g(\pi) = 1 - \pi$ és

$$Q_c\, g(\pi) = \min\{g(\pi), c\pi + Tg(\pi)\},$$

ahol, $Tg(\pi) = E'_{\pi}\, g(\pi_1)$.

Továbbá

$$v(\pi) = \lim_{N\to\infty} Q_c^N g(\pi).$$

Meg lehet mutatni, hogy $-v(\pi)$ a $-g(\pi)$ legkisebb $(1,c)$ ekszcesszív majoránsa (Sirjajev II.4.§. és II.8.§.). Speciálisan

$$-v(\pi) \ge -c\pi - Tv(\pi), \qquad 0 \le \pi \le 1,$$

és minden n-re

$$(11) \qquad -v(\pi) \ge -E'_{\pi}\left\{v(\pi_n) + c\sum_{k=0}^{n-1}\pi_k\right\}.$$

Innen (hosszabb számolással) belátható, hogy minden korlátos Markov-pontra (τ, $P_{\pi}(\tau \le N) = 1$, $N < \infty$)

$$(12) \qquad -v(\pi) \ge -E'_{\pi}\left\{v(\pi_{\tau}) + c\sum_{k=0}^{\tau-1}\pi_k\right\}.$$

Innen adódik, hogy ha $\tau \in R^{\pi}$ és $\tau_N = \min(\tau, N)$, $N \to \infty$,

$$-v(\pi) \ge \int_{(\tau<N)}\left[v(\pi_{\tau}) + c\sum_{k=0}^{\tau-1}\pi_k\right]dP'_{\pi} - \int_{\{\tau\ge N\}}\left[v(\pi_N) + c\sum_{k=0}^{N-1}\pi_k\right]dP'_{\pi}.$$

De $|v(\pi)| \le 1$ és $E'_{\pi}\sum_{k=0}^{\tau-1}\pi_k < \infty$ ha $\tau \in R^{\pi}$, ezért $N \to \infty$ esetén

$$\int_{(\tau\ge N)}\left|v(\pi_N) + c\sum_{k=0}^{N-1}\pi_k\right| dP'_{\pi} \le \int_{(\tau\ge N)} c\sum_{k=0}^{\tau-1}\pi_k \cdot dP'_{\pi} + P_{\pi}(\tau \ge N) \to 0.$$

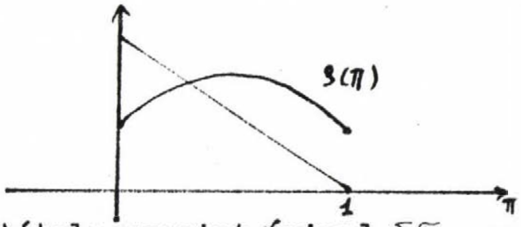
Tehát $-v(\pi)$ a legkisebb reguláris $(1,c)$ ekszcesszív majoránsa $-g(\pi)$-nek. (Sirjajev II.fej.16.tétele alapján)

$$(13) \qquad \rho(\pi) = \lim_{N\to\infty} Q_c^N g(\pi),$$

és

(14) $$g(\pi) = \min\{(1-\pi), c\pi + T\varrho(\pi)\}.$$

Egyszerü megmutatni, hogy $Q_c^N\, g(\pi)$ konkáv, igy $g(\pi)$ is az, tehát folytonos is.



Sirjajev (II.16.) tétele szerint (mivel $\sum \pi_k = \infty$), ha $c > 0$ a

$$\tau^*(\omega') = \inf(n \ge 0 : \varrho(\pi_n) = g(\pi_n)\}$$

optimális $\mathfrak{M}^{\pi}$-ben, azaz

$$\varrho(\pi) = \inf_{\tau \in \mathfrak{M}^{\pi}} E'_{\pi}\left\{(1-\pi_\tau) + c\sum_{k=0}^{\tau-1}\pi_k\right\} =$$

$$= E'_{\pi}\left\{(1-\pi_{\tau^*}) + c\sum_{k=0}^{\tau^*-1}\pi_k\right\}$$

Erre a τ^*-ra $\tau^* \in R^{\pi}$ azaz $E_{\pi}\sum_{k=0}^{\tau^*-1}\pi_k < \infty$.

A (14) egyenletből az $1-\pi$, $\varrho(\pi)$, $c\pi + T.\varrho(\pi)$ függvények folytonosságából és konkávságából következik, hogy létezik olyan A^*, melyre

(15) $$\tau^*(\omega') = \inf\{n : \varrho(\pi_n) = g(\pi_n)\} = \inf\{n \ge 0 : \pi_n \ge A^*\}$$

6. A "riasztás" feladata Brown-mozgás folyamatra

Legyenek megadva a következők:

1. $\{P_\pi,\ 0 \le \pi \le 1\}$ valószinüségek sokasága
2. $\theta = \theta(\omega)$ valószinüségi változó ($\theta \ge 0$), melynek eloszlása

$$(1)\qquad P_\pi\{\theta = 0\} = \pi \qquad P_\pi\{\theta \ge t \mid \theta > 0\} = e^{-\lambda t}, \qquad t > 0,$$

ahol λ fix, $0 < \lambda < \infty$, ismert és független π -től.

3. $\{w_t\},\ t \ge 0,$ standard Brown-mozgás folyamat, független (a P_π mértékek szerint) θ-tól, $w_0(\omega) \equiv 0$, $E_\pi \Delta w_t = 0$, $E_\pi \Delta w_t^2 = \Delta t$.
Megfigyeljük a ξ_t, $t \ge 0$, folyamatot

$$(2)\qquad d\xi_t = r\chi(t-\theta)\,dt + \sigma\, dw_t, \quad \xi_0(\omega) \equiv 0,$$
$$\sigma > 0, \quad r \ne 0, \quad \chi(t) = \begin{cases} 1 & t \ge 0, \\ 0 & t < 0, \end{cases}$$

amelynek alapján dönteni akarunk a folyamat lokális várható értékében történő esetleges változásról.

Bayes féle feladat megfogalmazása

Jelölje $\mathcal{F}_t = \sigma\{\omega, \pi : \pi, \xi_\Delta(\omega), s \le t\}$ a szokásos σ-algebrát. A Bayes féle veszteség legyen

$$(3)\qquad \rho(\pi) = \inf_{\tau \in \mathfrak{M}} \{P_\pi(\tau < \theta) + c\, E_\pi(\tau - \theta \mid \tau \ge \theta)\, P_\pi(\tau \ge \theta)\}$$

alaku, ahol $c > 0$. A $\tau^* \in \mathfrak{M}$ időpont Bayes féle, ha minden $0 \le \pi \le 1$-re

$$P_\pi(\tau^* < \theta) + c\, E(\tau^* - \theta \mid \tau^* \ge \theta)\, P_\pi(\tau^* \ge \theta) = \rho(\pi).$$

1. Tétel. Létezik Bayes féle megállási időpont $\tau^* = \tau^*(\omega, \pi)$ és

$$(4)\qquad \tau^*(\omega, \pi) = \inf\{t \ge 0 : \pi_t(\omega; \pi) \ge A^*\}$$

ahol

$$(5)\qquad \pi_t(\omega, \pi) = P_\pi(\theta(\omega) \le t \mid \mathcal{F}_t)$$

Az A^* konstans a

$$(6) \qquad C^{-1} = \int_0^{A^*} e^{-\Lambda[H(A^*)-H(x)]} \frac{dx}{x(1-x)^2}$$

egyenlet egyetlen gyöke,

$$C = c\left(\frac{r^2}{2\sigma^2}\right)^{-1}, \quad \Lambda = \lambda\left(\frac{r^2}{2\sigma^2}\right)^{-1}, \quad H(x) = \ln\frac{x}{1-x} - \frac{1}{x}$$

Továbbá

$$(7) \qquad \varrho(\pi) = \begin{cases} (1-A^*) + c\displaystyle\int_{1/A^*}^{1/\pi} \frac{e^{\Lambda x}(x-1)^{\Lambda}}{x^2}\left[\int \frac{e^{-\Lambda u}}{(u-1)^{2+\Lambda}}\,du\right]dx, & 0 \le \pi \le A^* \\ 1-\pi, & A^* \le \pi \le 1 \end{cases}$$

A tétel bizonyitásához szükség van egy sor segédtételre, melyeket az alábbiakban felsorolunk.

<u>1. Lemma</u> (Sirjajev, 173.o.) Legyen w_t, $t \in T$, standard Brown--mozgás folyamat, és tőle független mérhető, valós θ_t, melyre

$$E|\theta_t| < \infty, \quad \int_0^t E|\theta_s|^2\,ds < \infty, \quad t \ge 0.$$

Legyen η_t sztohasztikus differenciálja

$$d\eta_t = \theta_t \cdot dt + \sigma dw_t, \quad \eta_0(\omega) \equiv 0, \quad \sigma > 0,$$

akkor létezik olyan standard $\overline{w}_t$ Brown-mozgás folyamat, hogy

$$d\eta_t = \bar{\theta} \cdot dt + \sigma \cdot d\bar{w}_t, \quad \eta_0(\omega) \equiv 0,$$

ahol $\bar{\theta}_t = E(\theta_t | \mathcal{F}_t^{\eta})$, $\mathcal{F}_t^{\eta} = \sigma\{\omega : \eta_s(\omega),\ s \le t\}$.

A bizonyitás alapja, hogy

$$\bar{\eta}_t = \eta_t - \int_0^t E(\theta_s | \mathcal{F}_s^{\eta})\,ds$$

ugyancsak martingál, amelyre $E\bar{\eta}_t^{*2} < \infty$, $t \ge 0$, $E\left[(\bar{\eta}_t - \eta_s^2)^2 | \mathcal{F}_s^{\eta}\right] = \sigma^2 \cdot (t-s)$.

Innen következik, hogy $\bar{\eta}_t$ is Brown-mozgás folyamat.

<u>2. Lemma</u> A

$$\Pi_\pi = \{\pi_t(\omega,\pi), \mathcal{F}_t, P_\pi\}, \quad t \ge 0,$$

folyamat minden $0 \le \pi \le 1$-re Markov tipusu és kielégiti

$$(8) \qquad d\pi_t = \lambda(1-\pi_t)dt + \frac{r}{\sigma}\pi_t(1-\pi_t)d\bar{w}_t, \quad \pi_0 = \pi,$$

ahol $\overline{w}_t$ valamilyen standard Brown-mozgás folyamat.

A bizonyitáshoz szükség van az 1.lemmára, mely szerint van olyan

$\bar{w}_t$ standard Wiener folyamat, hogy

(9) $$d\xi_t = r\pi_t \cdot dt + \sigma \cdot d\bar{w}_t, \quad \xi_0 = 0.$$

Legyen $\mu_s(\cdot)$, $s \geq 0$, a valós függvények terén az $\{\eta_t^s\}$ $t \geq 0$, folyamatnak megfelelő mérték, ahol

$$d\eta^s = r\chi(t-s)\cdot dt + \sigma \cdot dw_t, \quad \sigma > 0, \ \eta_0 = 0.$$

(Ha $s = \infty$, $\mu(\cdot) = \mu_\infty(\cdot)$.)
Legyen továbbá

$$\nu_t^\pi(\cdot) = \pi\mu_0(\cdot) + (1-\pi)\int_0^t \lambda e^{-\lambda s}\mu_s(\cdot)\, ds$$

és

$$\nu^\pi(\cdot) = \nu_\infty^\pi(\cdot).$$

A Bayes formula alapján

$$\pi_t(\omega,\pi) = P_\pi(\theta \leq t \mid \xi_0^t) = \frac{d\nu_t^\pi}{d\nu^\pi}(\xi_0^t),$$

ez a Radon-Nikodym deriváltat jelenti a ξ_0^t "pontban". ν_t^π, ν^π és μ abszolut folytonosak egymásra nézve és

$$\frac{d\nu_t^\pi}{d\nu^\pi}(\xi_0^t) = \frac{d\nu_t^\pi}{d\mu}(\xi_0^t) \Bigg| \frac{d\nu^\pi}{d\mu}(\xi_0^t) =$$

$$= \frac{\pi \frac{d\mu_0}{d\mu}(\xi_0^t) + (1-\pi)\int_0^t \lambda e^{-\lambda s}\frac{d\mu_s}{d\mu}(\xi_0^t)\, ds}{\frac{d\mu}{d\mu}(\xi_0^t) + (1-\pi)\int_0^t \lambda e^{-\lambda s}\frac{d\mu_s}{d\mu}(\xi_0^t)\, ds + (1-\pi)\int_t^\infty \lambda e^{-\lambda s}\frac{d\mu_s}{d\mu}(\xi_0^t)\, ds}.$$

Már tudjuk, hogy

$$\frac{d\mu_s}{d\mu}(\xi_0^t) = \exp\left\{\frac{r}{\sigma^2}\left[(\xi_t - \xi_s) - \frac{r}{2}(1-s)\right]\right\}, \quad 0 \leq s \leq t,$$

és

$$\frac{d\mu_s}{d\mu} \equiv 1, \text{ ha } s > t.$$

Ilymódon

(10) $$\pi_t(\omega,\pi) = \frac{\pi e^{\frac{r}{\sigma^2}\left[\xi_t - \frac{r}{2}t\right]} + (1-\pi)\int_0^t e^{\frac{r}{\sigma^2}\left[(\xi_t - \xi_s) - \frac{r}{2}(t-s)\right]}\lambda e^{-\lambda s}\, ds}{\pi e^{\frac{r}{\sigma^2}\left[\xi_t - \frac{r}{2}t\right]} + (1-\pi)\int_0^t e^{\frac{r}{\sigma^2}\left[(\xi_t - \xi_0) - \frac{r}{2}(t-s)\right]}\cdot\lambda e^{-\lambda s}\, ds + (1-\pi)e^{-\lambda t}}$$

ahonnan

(11) $$1 - \pi_t(\omega,\pi) = \frac{(1-\pi)e^{-\lambda t}}{\pi e^{\frac{r}{\sigma^2}\left[\xi_t - \frac{r}{2}t\right]} + (1-\pi)\int_0^t e^{\frac{r}{\sigma^2}\left[\xi_t - \xi_0 - \frac{r}{2}(t-s)\right]}\lambda e^{-\lambda s} + (1-\pi)e^{-\lambda t}}$$

A Kolmogorov egyenlőtlenség szerint

$$P_\pi\left\{\sup_{s\leq t}|\bar{w}_s| \geq c\right\} \leq \frac{t}{c^2}, \quad t > 0.$$

és $P_\pi\left\{\sup_{s\leq t}|\bar{w}_s| < \infty\right\} = 1$, ahonnan (9) szerint

(12) $\quad P_\pi \left\{ \sup_{s \le t} |\xi_s| < \infty \right\} = 1 .$

Hacsak $0 \le \pi \le 1$ (12) alapján

(13) $\quad P_\pi \left\{ \pi_s < 1 ;\ s \le t \right\} = 1$

Legyen $\quad \varphi_t = \dfrac{\pi_t}{1 - \pi_t}$ akkor (13) alapján

(14) $\quad P_\pi \left\{ \varphi_s < \infty ,\ s \le t \right\} = 1 , \quad 0 \le \pi \le 1 , \quad t \ge 0 .$

De (1o), (11) szerint

(15) $\quad \varphi_t = \dfrac{\pi}{1-\pi}\, e^{\lambda t} \cdot e^{\frac{r}{\sigma^2}\left[\xi_t - \frac{r}{2} t\right]} + \lambda e^{\lambda t} \int_0^t e^{\frac{r}{\sigma^2}\left[(\xi_t - \xi_s) - \frac{r}{2}(t-s)\right]} e^{-\lambda s}\, ds ,$

ahonnan az Itó szabály szerinti sztohasztikus differenciálással $0 \le \pi \le 1$ esetén

$$d\varphi_t = \lambda (1 + \varphi_t)\, dt + \varphi_t\, d\xi_t , \qquad \varphi_0 = \frac{\pi}{1-\pi} \, .$$

Az Itó formula alapján (helyettesitéssel való differenciálás) $\pi_t = \dfrac{\varphi_t}{1+\varphi_t}$ -re ((14) alapján meg lehet tenni)

(16) $\quad d\pi_t = (1-\pi_t)\left(\lambda - \dfrac{r^2}{\sigma^2}\pi_t^2\right) dt + \dfrac{r}{\sigma^2}\pi_t (1-\pi_t)\, d\xi_t$

(9) és (16)-ból adódik (8).

<u>A tétel bizonyitása</u>

A $\pi' = \left\{ \pi_t(\omega'), \mathcal{F}'_t, P'_\pi \right\}, t \in T, \quad \omega' = (\omega, \pi), \quad \pi_0(\omega, \pi) = \pi$
$\mathcal{F}'_t = \sigma\left\{ \omega' : \pi_0(\omega'), \xi_s(\omega'), s \le t \right\}, \quad \xi_s(\omega') = \xi_s(\omega)$, jelölések bevezetése után, ahol a P'_π mérték értelmezése a következő: $P'_\pi(A) = P_\pi(A_\pi) \cdot \chi_{A^\omega}(\pi)$, ahol $A_\pi = \{\omega : (\omega, \pi) \in A\}$, $A^\omega = \{\pi : (\omega, \pi) \in A\}$ az uj téren vizsgáljuk a feladatot.
Legyen $\mathfrak{M}^\pi$ azon Markov pontok, $\tau \in \mathfrak{M}$, osztálya, melyre

$\left\{ \omega' : \tau(\omega') \le t \right\} \in \mathcal{F}^\pi_t ,$ ahol $\mathcal{F}^\pi_t = \sigma\left\{ \omega' : \pi_s(\omega') , s \le t \right\} .$

Meg kell mutatni (mint ahogyan azt a diszkrét esetben tettük), hogy

$$(17)\quad \rho(\pi) = \inf_{\tau \in \mathfrak{M}^{\pi}} E_{\pi}' \left\{ (1-\pi_{\tau}) + c\int_0^{\tau} \pi_s \, ds \right\} =$$
$$= \inf_{\tau \in \mathfrak{R}^{\pi}} E_{\pi}' \left\{ (1-\pi_{\tau}) + c\int_0^{\tau} \pi_s \, ds \right\},$$

ahol $\mathfrak{R}^{\pi}$ azon τ Markov-pontok összessége ($\tau \in \mathfrak{M}^{\pi}$), melyekre

$$(18)\quad E_{\pi}' \int_0^{\tau} \pi_s \, ds < \infty, \quad 0 \le \pi \le 1.$$

Mivel $E_{\pi}' \int_0^{\tau} \pi_s \, ds = E_{\pi} \max(\tau - \theta, 0)$

és $E_{\pi}\theta < \infty$, $0 \le \pi \le 1$, az $\mathfrak{R}^{\pi}$ osztály megegyezik azon $\mathfrak{M}^{*} = \{\tau\}$ osztállyal, ahol $E_{\pi}'\tau < \infty$, $0 \le \pi \le 1$.

Jelölje

$$g(\pi) = 1 - \pi$$
$$Q_{cn}\, g(\pi) = \min\left\{ g(\pi), E_{\pi}' \left[g(\pi_{2^{-n}}) + c\int_0^{2^{-n}} \pi_s \, ds \right] \right\},$$

mennyiségeket

$Q_{c,n}^{N}$ legyen a $Q_{c,n}$ N-edik hatványa.

Nyilván

$$E_{\pi}' \int_0^{\Delta} \pi_s \, ds = \pi \cdot \Delta + (1-\pi)\int_0^{\Delta} P_{\pi}\{\theta \le s \mid \theta > 0\}\, ds =$$
$$= \Delta - \frac{1-\pi}{\lambda}\left(1 - e^{-\lambda\Delta}\right), \quad \Delta \ge 0.$$

Ezért

$$Q_{c,n}\, g(\pi) = \min\left\{ g(\pi), E_{\pi}' g(\pi_{\Delta}) + c\left[\Delta - \frac{1-\pi}{\lambda}(1 - e^{\lambda\Delta})\right]\right\}, \quad \Delta = 2^{-n}.$$

Belátható (lásd Sirjajev, 217.o.), hogy $Q_{c,n}g(\pi)$ és $Q_{c,n}^{N}g(\pi)$ felülről konkávok, továbbá

$$\rho(\pi) = \lim_{n \to \infty} \lim_{N \to \infty} Q_{c,n}^{N}\, g(\pi)$$

és $\rho(\pi)$ is konkáv és folytonos (0,1)-ben.

Mivel

$$(19)\quad \lim_{t \to \infty} \int_0^{t} \pi_s \, ds = \infty \quad (P_{\pi}\text{ m.m.})$$

és $c > 0$ a Sirjajev II.5.tétel 1.megjegyzése szerint létezik $\mathfrak{R}^{\pi}$-ben optimális Markov-pont:

$$\tau^{*}(\omega) = \inf\{t \ge 0 : \rho(\pi_t) = g(\pi_t)\} = \inf\{t \ge 0 : \pi_t \ge A^{*}\}.$$

Az A^{*} és $\rho(\pi)$ meghatározása a következőképpen történik.

Legyen az infinitézimális operátor

$$\mathcal{A}f(x) = \lim_{t \downarrow 0} \frac{T_t f(x) - f(x)}{t} \quad \text{és} \quad \mathfrak{O}f(x_0) = \lim_{U \downarrow x} \frac{M_{x_0} f(x_{\sigma(U)}) f(x_0)}{M_{x_0} \sigma(U)},$$

ahol $\sigma(U)$ U-ból való kilépés időpontja, és

$$T_t f(x) = \int_E f(y)\, P(t, x, dy).$$

Legyen $\varrho(\pi)$ folytonosan differenciálható A^*-ban, akkor

$$\begin{aligned} &\mathfrak{O}\cdot\varrho(\pi) = -c\pi, & 0 \le \pi \le A^*, \\ (20) \quad &\varrho(\pi) = g(\pi), & A^* \le \pi \le 1, \\ &\frac{d\varrho(\pi)}{d\pi}\Big|_{\pi = A^*} = \frac{dg(\pi)}{d\pi}\Big|_{\pi = A^*}. \end{aligned}$$

Ha $F^* = f(\pi)$, $0 \le \pi \le 1$, a nemnegativ, konkáv, kétszer folytonosan differenciálható függvények osztálya. A (2o) megoldását itt keressük.

Ha $f \in F^*$ $\mathfrak{O}f(\pi) = \mathfrak{D}f(\pi)$ (Dünkin 5.7. tétel), akkor

$$(21) \quad \mathfrak{D} = \lambda(1-\pi)\frac{d}{d\pi} + \frac{r^2}{2\sigma^2}[\pi(1-\pi)]^2 \frac{d^2}{d\pi^2}.$$

Azaz keressük azt az f függvényt és A konstanst, melyre

$$\begin{aligned} (22) \quad &\lambda(1-\pi) f'(\pi) + \varrho[\pi(1-\pi)] f''(\pi) = -c\pi, & 0 \le \pi < A, \\ &f(\pi) = 1-\pi, & A \le \pi \le 1 \\ &f'(A) = -1, & \varrho = r^2/2\sigma^2. \end{aligned}$$

Megmutatjuk, hogy F^*-ban van f^* megoldás, amely egyértelmü és

$$(23) \quad \frac{df^*(\pi)}{d\pi}\Big|_{\pi=0} = 0.$$

Legyen $C = \frac{c}{\varrho}$, $\Lambda = \frac{\lambda}{\varrho}$, $\psi(\pi) = f'(\pi)$, akkor (67)-ből

$$(24) \quad \psi'(\pi) = -\frac{C\pi + \Lambda(1-\pi)\psi(\pi)}{[\pi(1-\pi)]^2}$$

amelynek szinguláris pontja van 0-ban.

Létezik a $\psi^*(\pi)$ szeparátrix (melyre $\psi^*(0) = 0$)

$$(25) \quad \psi^*(\pi) = -C\int_0^{\pi} e^{-\Lambda[H(\pi) - H(y)]} \frac{dy}{y(1-y)^2},$$

ahol $\quad H(y) = \ln\frac{y}{1-y} - \frac{1}{y}$

Ha A^* a

$$(26) \quad \psi^*(A^*) = -1$$

megoldása

$$(27)\qquad f^*(\pi)=\begin{cases}(1-A^*)-\int_{\pi}^{A^*}\psi^*(x)\,dx, & 0\le \pi<A^*,\\ 1-\pi, & A^*\le\pi\le 1,\end{cases}$$

nemnegativ, konkáv és megoldása (22)-nek.
Megmutatható, hogy F^*-ben egyetlen megoldás (Sirjajev, 219.o.).

Még megmutatjuk, hogy $f^*(\pi)$ a $\rho(\pi)$ rizikóval azonos.

Dünkin egy tétele (5.1. tétel) szerint, ha $\pi<A^*$ minden $\tau\in\mathcal{Q}^{\pi}$-re

$$E'_{\pi}f^*(\pi_t)-f^*(\pi)=-c\,E'_{\pi}\int_0^t\pi_s\,ds .$$

Akkor (18) szerint

$$(28)\qquad \rho(\pi)=\inf_{\tau\in\mathcal{Q}^{\pi}}E'_{\pi}\left\{g(\pi_\tau)+c\int_0^{\tau}\pi_s\,ds\right\}\ge$$

$$\ge\inf_{\tau\in\mathcal{Q}^{\pi}}E'_{\pi}\left\{f^*(\pi_\tau)+c\int_0^{\tau}\pi_s\,ds\right\}+\inf_{\tau\in\mathcal{Q}^{\pi}}E'_{\pi}\left\{g(\pi_\tau)-f^*(\pi_\tau)\right\}=$$

$$=f^*(\pi)+\inf_{\tau\in\mathcal{Q}^{\pi}}E'_{\pi}\left\{g(\pi_\tau)-f^*(\pi_\tau)\right\}$$

De $g(\pi)\ge f^*(\pi)$, $0\le\pi\le 1$ esetén, és a

$$\tau^*=\inf\left\{t\ge 0:\pi_t\ge A^*\right\}$$

pontra

$$E'_{\pi}\left\{g(\pi_{\tau^*})-f^*(\pi_{\tau^*})\right\}=0$$

Tehát $\rho(\pi)\ge f^*(\pi)$.

De

$$E'_{\pi}\left\{g(\pi_{\tau^*})+c\int_0^{\tau^*}\pi_s\,ds\right\}=E'_{\pi}\left\{f(\pi_{\tau^*})+c\int_0^{\tau^*}\pi_s\,ds\right\}=f^*(\pi),\quad 0\le\pi<A^*,$$

azaz $\rho(\pi)=f^*(\pi)$ ebben az intervallumban.

Hasonlóan bizonyitható az állitás $\pi\ge A^*$ esetén. Ezzel a tétel bizonyitása kész.

A variációs feladatra a megoldás innen adódik.
Igaz továbbá a diszkrét időben bizonyitott tétel megfelelője is.

2. Tétel. Legyen $0<\alpha<1$, $0\le\pi<1$, $0<\lambda<\infty$, $\mathfrak{M}(\alpha;\pi)$ azon Markov pontok összessége, melyekre $P_{\pi}(\tau<\theta)\le\alpha$, akkor a

$$\tilde\tau_\alpha=\inf\left\{t\ge 0:\pi_t(\omega,\pi)\ge\tilde A_\alpha\right\}$$

(ahol $\tilde A_\alpha=1-\alpha$) Markov pont optimális, abban az értelemben, hoyg minden $\tau\in\mathfrak{M}$-re

$$E_{\pi}(\tilde\tau_\alpha-\theta\,|\,\tilde\tau_\alpha\ge\theta)\le E_{\pi}(\tau-\theta\,|\,\tau\ge\theta).$$

A tétel bizonyitására nem térünk ki, az megtalálható Sirjajev idézett könyvében.

Adott α hibás riasztási valószinüség esetén

(29) $R(\alpha;\lambda) = E_0(\tilde{\tau}_\alpha - \theta / \tilde{\tau}_\alpha \geq \theta)$

viselkedése (legyen $\pi_0 = 0$) különböző aszimptotikus esetekben külön érdekességgel bir.

Ha $\lambda \to 0$, azaz $E_0\theta = 1/\lambda \to \infty$ és $\alpha \to 1$, de

$$\frac{1-\alpha}{\lambda} = T$$

a következő összefüggést kapjuk

$$R(T) = \lim_{\substack{\alpha \to 1 \\ \lambda \to 0}} R(\alpha;\lambda) = \frac{(\rho T)^{-1}}{\rho} \int_{(\rho T)}^{\infty} \frac{e^y}{y^2} [-E_i(y)]\, dy$$

ahol

$$-E_i(-y) = \int_y^{\infty} \frac{e^{-z}}{z} dz, \quad \rho = \frac{r^2}{2\sigma^2}$$

Ha $T \gg 1$ innen adódik, hogy

$R(T) = \frac{1}{\rho}\left\{\ln \rho T - 1 - C + O\left(\frac{1}{\rho T}\right)\right\}$, ahol C az Euler féle konstans (C=0,577...)

7. Számitógépek hierarchikus lap tárolási eljárásainak optimalizálásáról

Ahhoz hogy a számitógépek memória lehetőségeit jól és gyorsan lehessen kihasználni, gyakran alkalmaznak lineáris hierarchikus tárolási eljárásokat. Az ilyen hierarchikus információ tárolási eljárások esetén meghatározott számu szó alkot egy "lapot". A hierarchia minden szintjén lehetőség van lapok tárolására, és az egyes szinteken belül elhelyezhető lapok száma különböző lehet. Egy futó programban a tárolási helyekre történő hivatkozások minden lépésben kétféle módon történhetnek. Vagy az első szinten elhelyezkedő lapra irányul a hivatkozás, amikor is az elérés közvetlenül megtörténik, vagy egy alsóbb szinten elhelyezkedő lapra, amikor ezt a lapot át kell helyezni automatikusan az első szintre. Az utóbbi esetben más lapokat alsóbb szintre kell helyezni; az alsóbb szinten lévő lapra hivatkozást szokás lapolási hibának (page fault) nevezni. A tárolási eljárást lineárisan hierarchikusnak szokás nevezni, ha a keresett lapot a megtalálási szintről az összes közbenső szinteken keresztül lehet eljuttatni a legelső szintre.

A memória elvi bővitésének lehetőségét először a virtuális memória rendszerek megvalósitása adta (lásd Kilburn*), melyet először az ATLAS rendszernél készitettek. A buffer-tárolós memóriákat az IBM 360-as rendszerben alkalmazták. A lapok cserélési eljárásai közül az u.n. "megkeresési lapolás"(demand paging) - amikor is csak akkor van csere a szintek között, ha lapolási hiba fordul elő - a legelterjedtebb. A cserélési algoritmusok közül a következőket emlitjük meg:

1. Az első bekerült kerül ki (FIFO: first in - first out) algoritmus, amely az első szinten legrégebben bennlévő lapot helyezi alacsonyabb szintre lapolási hiba esetén.
2. A legritkábban használt kerül ki (LFU: least frequently used) algoritmus, amely az első szinten - egy bizonyos perióduson belül - legkevesebbszer használt lapot helyezi alacsonyabb szintre lapolási hiba esetén.
3. A legrégebben használt kerül ki (LRU: least recently used) algoritmus, amely az első szinten legrégebben hivatkozott lapot helyezi alacsonyabb szintre lapolási hiba esetén.

---

*Kilburn,T.,Edwards,D.,Lanigan,M.,Sümner,F.: One level storage system. IEEE Transactions on Electronic Comp.,EC-11(1962)2,223-235.

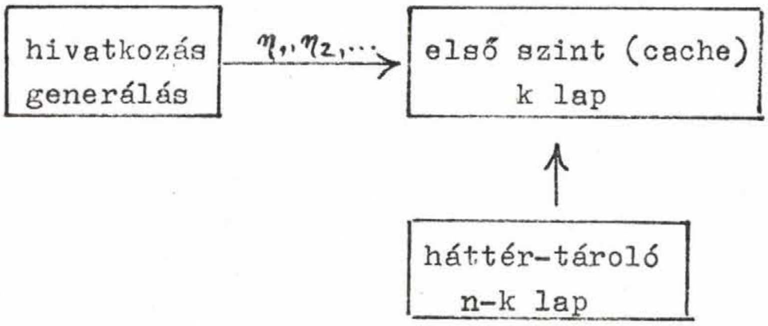
4. Optimális algoritmus, amely azt a lapot helyezi az első szintről alacsonyabbra, amelyre a jövőben a legkésőbb fognak hivatkozni. Ez utóbbi nyilván használhatatlan a gyakorlatban, mivel a program jövőbeni viselkedését kellene ismerni.
5. Véletlen (RR: random replacement) algoritmus, mely az első szint lapjai közül bármelyiket egyenlő valószinüséggel helyezi alsóbb szintre lapolási hiba esetén.

Jelen paragrafus célja annak a megmutatása, hogy bizonyos egyszerü feltevések esetén megadhatók olyan tárolási illetve cserélési eljárások, amelyek a lapolási hibák átlagos számát minimalizálják. A feladat megoldásánál a legjelentősebb korlátozás, hogy az egymásutáni hivatkozások függetlenek. Ez a feltevés a gyakorlatban ritkán teljesül, és csak mint durva közelités használható. Opderbeck és Chu* dolgozatában a relativ gyakoriságokon alapuló lap tárolási algoritmust vizsgálják. Ez az algoritmus áll legközelebb a Bayes-féle feltevésen alapuló optimális eljáráshoz, igy nyilvánvaló, hogy szimulációs eredményeik alátámasztják a jelen részben bizonyitandó optimális eljárás jóságát. Idézett dolgozatukban megmutatják, hogy az általuk gyakorisági helyettesitési algoritmusnak (PFF: page fault frequency replacement algorithm) nevezett eljárás jobb, mint az LRU (legrégebben használt) algoritmuson alapuló eljárás. Mérési eredményeik egyben azt is igazolják, hogy a lapok független hivására vonatkozó feltevések igen általános feltételek mellett jó közelitésnek tekinthetők. Megmutatjuk azokat a feltételeket, melyek mellett a megoldás a "kétpisztolyos bandita" (illetve a többpisztolyos bandita) problémakör megoldására vezethető vissza. Kitérünk a különböző más feltétel melletti feladat megfogalmazásokra, azok lehetséges megoldásaira és a közelitésekre is.

Mindenekelőtt egy, a programozás szempontjából elvi jelentőségü feladat megfogalmazásával kezdjük. Multiprogramozásu gép egy programját vizsgálva vetődik fel ezen program két lapjának két különböző fokozaton való elhelyezkedési problémája. Amennyiben a második fokozaton lévő lapra történő hivatkozás esetén a két lapot azonnal ki kell cserélni - semmilyen optimalizálási feladat nem merül fel. Tegyük fel, hogy lehetőség van mindkét lap első szinten való tartására a hivatkozás befejezéséig, és csak ezután (azaz az ujabb hivatkozás

* Opderbeck, H., Chu, W. W.: Performance of the page fault frequency replacement algorithm in a multiprogramming environment. IFIP Congress (1974), Inform. Proc., 235-241.

előtt) helyezzük a lapok valamelyikét a második fokozatra. Az Aho, Denning, Ullmann* féle kétszintes tárolási eljárás (lásd még Franaszek, Wagner**) esetén, amikor az elérési idők átlagát akarjuk minimalizálni, a fenti megvalósitás reális. Ezért röviden ismertetjük egy ilyen rendszer vázlatát (lásd 1.ábra).



1. ábra. Kétszintes tárolás

Az első szint (cache vagy puffer) egy kisebb, de gyors elérésü memória egység, a második szint (háttér-tároló) lassabb elérésü memória egység. Az $\eta_1, \eta_2, \dots$ hivatkozási sor generálása a lehetséges $A_1, A_2, \dots A_n$ lapokra vonatkozik ($\eta_i = j$, ha az i-edik hivatkozás az A_j lapra történik). Az összes elérés az első szinten történik, ha a hivatkozott lap az első szinten van: az elérési idő T_1 mp., ha a második szinten, akkor előbb az első szintre kell hozni, és az elérési idő T_2 mp. (általában $T_2 \gg T_1$). Az első szinten egy szabad hely rendszerint rendelkezésre áll az esetleges csere (helyettesités) lebonyolitásához (lásd Aho, Denning, Ullmann*). Az itt ismertetett eljárás annyiban uj, hogy a hivatkozás végén kell dönteni a második szintre történő kivitelről (esetleges cseréről).

Visszatérve a két lap esetére az előbbi megfogalmazás alapján jutunk a következő optimalizálási feladatra. Az A_1 és A_2 programrészekre (lapokra) való hivatkozások történjenek függetlenül egymásután, de hivatkozásaik valószinüségét nem ismerjük. A könnyebb elérésü szint sorszáma legyen 1, a nehezebbé 2. Minden hivatkozás után lehetőség van annak eldöntésére, hogy az A_1 vagy az A_2 programrész kerüljön a könnyebb elérésü helyre. N számu független hivatkozás esetén olyan döntési eljárást kivánunk kidolgozni az A_1 ill.

*Aho,A.V.,Denning,P.J.,Ullmann,J.D.: Principles of optimal page replacement. Journ.ACM, 18 (1971)1, 8o-93.

**Franaszek,P.A., Wagner,T.J.: Some distribution-free aspects of paging algorithm performance. Journ.ACM, 21(1974)1,4o-53.

A_2 elhelyezésére, amely minimalizálja a nehezebb elérésü helyekre való hivatkozások számát (azaz a lapolási hibák számát). Feltételezzük, hogy az esetleges cserék nem kerülnek költségbe.

Ha az A_1 és A_2 programrészek hivatkozási valószinüségei p_1 és $1-p_1$ ismertek, és $p_1 > 1-p_1$, az A_1 részt kell a könnyebb elérésü 1. helyre helyezni, mivel az 1. helyre történő átlagos hivatkozások száma ekkor Np_1, és ez nagyobb, mint az ellenkező elhelyezés esetén. Ha az A_1 és A_2 hivatkozási valószinüségei nem ismertek első közelitésben, feltesszük hogy az egyes programrészekre való hivatkozási valószinüségek p_1 illetve p_2 és $p_1 > p_2$ ismertek ($p_1 + p_2 = 1$, bár ez nem szükséges kikötés), de ismeretlen a hozzárendelés sorrendje.

Jelölje η_t a megfigyelési folyamatot ($t=1,2,\ldots,N$), lehetséges értékei 1,2 és η_t megadja, hogy a t időpontban az $A_1(\eta_t=1)$ vagy az $A_2(\eta_t=2)$ programrészre történt-e hivatkozás.

Az $X_t^{(d)}$ (ahol $d=1$, ha az A_1, és $d=2$, ha az A_2 programrész van a 2./nehezebb/ elérésü helyen) valószinüségi változó a d döntés esetén megadja, hogy a t időpontban melyik helyre történt hivatkozás:

$$X_t^{(d)} = \begin{cases} 1 & \text{ha az 1. helyen lévő programrészre történik hivatkozás,} \\ 0 & \text{ha a 2. helyen lévő programrészre történik hivatkozás.} \end{cases}$$

Minden kisérletnél módunkban áll választani (döntést hozni) az előző kisérletek eredményei alapján arról, hogy az $X^{(1)}$ (az A_1 rész kerül a 2. helyre) vagy az $X^{(2)}$ változót figyeljük meg.

Nyilvánvaló, hogy

$$X_t^{(d)} = \begin{cases} 0 & \text{ha } \eta_t = d, \\ 1 & \text{ha } \eta_t \neq d. \end{cases}$$

A W valószinüségi változó lehetséges értékei legyenek

1 ha (A_1, A_2) hivási valószinüségei (p_2, p_1),

2 ha (A_1, A_2) hivási valószinüségei (p_1, p_2).

Legyen továbbá az u.n. nem megfigyelhető W komponens apriori eloszlása

$$P(W=1)=\xi_1, \quad P(W=2)=\xi_2=1-\xi_1.$$

Feltevéseink szerint tetszőleges t időpontban

$$P\{X_t^{(1)}=1\mid W=1\}=p_1, \quad P\{X_t^{(1)}=0\mid W=1\}=p_2,$$
$$P\{X_t^{(2)}=1\mid W=2\}=p_1, \quad P\{X_t^{(2)}=0\mid W=2\}=p_2,$$
$$P\{X_t^{(1)}=1\mid W=2\}=p_2, \quad P\{X_t^{(1)}=0\mid W=2\}=p_1,$$
$$P\{X_t^{(2)}=1\mid W=1\}=p_2, \quad P\{X_t^{(2)}=0\mid W=1\}=p_1.$$

Az optimalizálási feladat megfogalmazása: a $\delta=(d_0, d_1, \ldots, d_{N-1})$ döntési sorozat olyan δ^* megválasztása, melyre

$$E\left(X_1^{(d_0)}+X_1^{(d_1)}+\ldots+X_N^{(d_{N-1})}\right)=\max.$$

A feladat megfogalmazása és megoldása megtalálható DeGroot* könyvében (14.7.§), ezt ismertetjük a következő pontban. A dinamikus programozás u.n. Bellman egyenlete segitségével történő megoldást (lásd pl. Prohorov-Rozanov** 363.o.) ismertetjük az alábbiakban. Ez a nyereség egy más megfogalmazását igényli. A nyereségfüggvény V(x) definiciója az x megfigyelés esetén a következő

$$V(x)=\begin{cases}1 & \text{ha} \quad x=1,\\ 0 & \text{ha} \quad x=0.\end{cases}$$

Az optimalizálási feladat a W változó t időpontban adott $\xi(t)$ apriori (vagy a t-1 utáni aposteriori) eloszlás esetén

$$E_{t,\xi,x}\sum_{s=t}^{N} V\left(X_s^{(d_{s-1})}\right)=V\left(t,\xi(t), X_t=x, \delta^{[t,N]}=(d_{t-1},\ldots,d_{N-1})\right)$$

maximalizálása, azaz

$$V(t,\xi(t),x)=\sup_{\delta^{[t,N]}} V\left(t,\xi(t), X_t=x, \delta^{[t,N]}\right)$$

megadása. Nyilvánvaló, hogy

$$V(N,\xi(N),x)=V(x),$$ (független a ξ eloszlástól),

---

* DeGroot, M.H.: Optimal Statistical Decisions. McGraw-Hill, 197o.

** Prohorov,J.V., Rozanov,J.A.: Probability theory. Nauka,Moszkva,1967.

$$V(N-1,\xi(N-1),x)=V(x)+\max_{d_{N-1}}\left\{P(W=1)P(X_N^{(d_{N-1})}=1\mid W=1)V(N,\xi(N),1)+\right.$$

$$\left.+P(W=2)P(X_N^{(d_{N-1})}=1\mid W=2)V(N,\xi(N),1)\right\},$$

ahol $V(N,\xi(N),1)=1.$

Az utóbbi összefüggés a $d_{N-1}=1$ esetben a következőt adja

$$V(N-1,\xi(N-1),x,d_{N-1}=1)=V(x)+\{\xi_1(N)p_2+\xi_2(N)p_1\}$$

mig a $d_{N-1}=2$ esetben

$$V(N-1,\xi(N-1),x,d_{N-1}=2)=V(x)+\{\xi_1(N)p_1+\xi_2(N)p_2\}$$

adódik. Különbségük (feltevésünk szerint $p_1>p_2$)

$$\xi_1(N)(p_1-p_2)-\xi_2(N)(p_1-p_2)=(\xi_1(N)-\xi_2(N))(p_1-p_2)$$

alapján a $\xi_1(N)>\xi_2(N)$ esetben $d_{N-1}=1$, mig a $\xi_1(N)<\xi_2(N)$ esetben a $d_{N-1}=2$ döntés az optimális. Tehát csak a $\xi(N)$ aposteriori eloszlás alapján kell dönteni az eljárás optimalizálásáról ebben a lépésben. A továbbiakban szükség van a $P\{W=1\mid X_t^{(d_{t-1})}=x\}$ valószinüségek meghatározására. A Bayes tétel alapján (ahol $P\{W=i\}$ jelenti a t időpontbeli apriori valószinüséget)

$$P\{W=1\mid X_t^{(d_{t-1})}=x\}=\frac{P\{X_t^{(d_{t-1})}=x\mid W=1\}P\{W=1\}}{P\{X_t^{(d_{t-1})}=x\mid W=1\}P\{W=1\}+P\{X_t^{(d_{t-1})}=x\mid W=2\}P\{W=2\}}$$

és látható, hogy a W változó t-beli aposteriori valószinüsége kifejezhető az apriori valószinüségek segitségével. Az aposteriori valószinüségek segitségével tetszőleges t-re

$$V(t,\xi(t),x)=V(x)+\max_{d}\left\{\xi_1(t+1)P\{X_t^{(d_{t-1})}=1\mid W=1\}V(t+1,\xi(t+1),1)+\right.$$

$$\left.+\xi_2(t+1)P\{X_t^{(d_{t-1})}=1\mid W=2\}V(t+1,\xi(t+1),1)\right\}.$$

Innen könnyen leolvasható, hogy az optimális döntési eljárás:

$$d_t=1 \quad \text{ha} \quad \xi_1(t+1)>\xi_2(t+1),$$
$$d_t=2 \quad \text{ha} \quad \xi_1(t+1)<\xi_2(t+1).$$

A döntési eljárás tehát az u.n. rövidlátó politika: a t-edik lépésben azt a programrészt (lapot) helyezzük a 2.szintre (nehezebb elérésü helyre), amelynek hivási valószinüsége, p_2, aposteriori valószinüsége, $\xi(t)$, nagyobb.

Feladatok.

1. Jelölje $\sigma_0, \sigma_1, \ldots, \sigma_t, \ldots$ az első szinten lévő lapokat, azaz $\sigma_t = \sigma_t(A_{i_1}, \ldots, A_{i_k})$ ha az $i_1, \ldots, i_k$ sorszámu lapok vannak az első szinten. σ_t lehetséges állapotainak száma $\binom{n}{k}$ a sorrend figyelembevétele nélkül.

Ismert $p_1, p_2, \ldots, p_n$ hivatkozási valószinüségek esetén a legegyszerübb (és egyben optimális) stratégia, ha σ_t választása a következő

$$\sigma_t = \sigma_0(A_1, A_2, \ldots, A_k)$$

és ha második szinten lévő lapra történik hivatkozás, akkor behozva a lapot a hivatkozás után ismét visszahelyezzük a második szintre (az ujabb kihelyezés árát nem véve figyelembe). Ez az eljárás ugyan megkeresési lapolási eljárás, de nem megkeresési cserélési (helyettesitő) lapolási eljárás.

A legegyszerübb megkeresési helyettesitő lapolási eljárást Aho et al. 1.) cikke javasolja. Ez a következő:

$$\sigma_t = \sigma_0(A_1, A_2, \ldots, A_{k-1}, \eta_t), \quad \text{ha} \quad \eta_t \bar{\in} \sigma_{t-1},$$
$$\sigma_t = \sigma_0(A_1, A_2, \ldots, A_{k-1}, \eta_{t-1}) \quad \text{ha} \quad \eta_t \in \sigma_{t-1},$$

ahol η_t a t időpontban hivatkozott lapot jelenti ($\eta_t=0$ ha nem történt hivatkozás a második szintre a t időpontig).

Mutassuk meg, hogy σ_t homogén Markov láncot alkot abban az esetben, ha η_t egy független azonos eloszlásu sorozat. A

$$P\{\sigma_t = j \mid \sigma_{t-1} = i\} = P\{\sigma_t = \sigma_0(A_1, \ldots, A_{k-1}, A_j) \mid \sigma_{t-1} = \sigma_0(A_1, \ldots, A_{k-1}, A_i\} = P_{ij} \qquad i, j \geq k$$

átmenetvalószinüségek a következők:

$$P_{ij} = \begin{cases} p_1 + p_2 + \ldots + p_{k-1} + p_i & \text{ha} \quad j = i, \\ p_j & \text{ha} \quad j \neq i. \end{cases}$$

Határozzuk meg a σ_n Markov lánc stacionér eloszlását.

Mutassuk meg, hogy a lapolási hiba valószinüsége a következő:

$$\sum_{i=k}^{n} p_i \left(1 - \frac{p_i}{\sum_{j=k}^{n} p_j}\right)$$

2. Legyen

$$Q = \sum_{j=k}^{n} p_j ,$$

1.) Aho, A., Denning, P., Ullman, J.: Principles of optimal page replacement. Jorn. ACM, 18 (1971) 80-93.

$$Q_1 = \sum_{j=k-1}^{n} p_j \left(1 - \frac{p_j}{\sum_{k-1}^{n} p_i}\right).$$

Mutassuk meg, hogy

$$1 \le \frac{Q_1}{Q_0} \le 2 .$$

Ha Q_{LRU} jelenti a lapolási hiba valószinüségét az LRU algoritmus (legritkábban használt kerül ki) esetén, mutassuk meg, hogy

$$\frac{Q_{LRU}}{Q_0} \le 1 + [(k-1)(1-Q_0)/(1+(k-2)Q_0]$$

tetszőleges $(p_1, \ldots, p_n)$ esetén [lásd Franaszek, Wagner már idézett dolgozatát].

3. Tekintsük 3 lap esetét. Legyenek az A_1, A_2, A_3 lapok hivatkozási valószinüségei $p_1 > p_2 > p_3$ ismeretlenek abban az értelemben, hogy a hozzárendelési sorrend nem ismert. Bevezetve a w valószinüségi változót, mely megadja a hozzárendelés sorrendjét

$$(A_1,\ A_2,\ A_3)$$

$w = 1$	$p_3,\ p_2,\ p_1$	$P(w=1) = \xi_{11}$	
$w = 2$	$p_3,\ p_1,\ p_2$	$P(w=2) = \xi_{12},$	$\xi_{11} + \xi_{12} = \xi_1,$
$w = 3$	$p_2,\ p_3,\ p_1$	$P(w=3) = \xi_{21},$	
$w = 4$	$p_1,\ p_3,\ p_2$	$P(w=4) = \xi_{22},$	$\xi_{21} + \xi_{22} = \xi_2,$
$w = 5$	$p_2,\ p_1,\ p_3$	$P(w=5) = \xi_{31},$	
$w = 6$	$p_1,\ p_2,\ p_3$	$P(w=6) = \xi_{32}$	$\xi_{31} + \xi_{32} = \xi_3.$

Feltéve, hogy $p_1 = p_2 = \frac{1-\varepsilon}{2}$, $p_3 = \varepsilon$ mutassuk meg, hogy az optimális döntési eljárás csak a max $(\xi_1(t),\ \xi_2(t),\ \xi_3(t))$ aposteriori valószinüségektől függ. (Lásd Arató 1.).).

4. Legyen az $X^{(1)}$ és $X^{(2)}$ valószinüségi változók eloszlása a következő:

$$P\{X^{(1)} = 1\} = w_1, \qquad P\{X^{(1)} = 0\} = 1 - w_1,$$
$$P\{X^{(2)} = 1\} = w_2, \qquad P\{X^{(2)} = 0\} = 1 - w_2,$$

ahol tudjuk, hogy $w_2 = 1/2$. A w_1 legyen valószinüségi változó

1.) Arató M.: Számitógépek hierarchikus laptárolási eljárásainak optimalizálásáról. MTA SZTAKI Közlemények (1975).

$P(w_1 = 3/4) = \xi = 1 - P(w_1 = 1/4)$.

Lehetőségünk van n megfigyelést végezni az $X^{(1)}$ vagy $X^{(2)}$ változókra vonatkozólag. Az

$$E(X_1^{(d_1)} + \ldots + X_n^{(d_n)})$$

összeget akarjuk maximalizálni fix n esetén. Nyilvánvaló, hogy a ξ eloszlásra csak $X^{(1)}$ megfigyelése alapján kapunk információt, igy az optimális eljárás r ($r = 0,1,2,\ldots,n$) darab $X^{(1)}$ változó, majd n-r darab $X^{(2)}$ változó megfigyelésében áll.

Legyen

$$T(n,r) = E(X_1^{(1)} + \ldots + X_r^{(1)} + X_{r+1}^{(2)} + \ldots + X_n^{(2)})$$

Határozzuk meg T(n,r) értékét és adjuk meg, hogy ξ milyen értékére lesz r=n. [Utmutatás: Ha $r = 0$ $T(n,0) = \frac{n}{2}$. Legyen $1 \leq r \leq n$ és jelölje A_j azt az eseményt, hogy pontosan j darab $X^{(1)}$ változó értéke volt 1. Nyilván

$$T(n,r) = \sum_{j=0}^{r} E(Z|A_j)P(A_j),$$

ahol

$$P(A_j) = P(A_j|w_1 = 3/4)P(w_1 = 3/4) + P(A_j|w_1 = 1/4)P(w_1 = 1/4) =$$

$$= \xi \cdot \binom{r}{j}\left(\frac{3}{4}\right)^j\left(\frac{1}{4}\right)^{r-j} + (1-\xi)\binom{r}{j}\left(\frac{1}{4}\right)^j\left(\frac{3}{4}\right)^{r-j}$$

és

$$E(Z|A_j) = j + \frac{n-r}{2}.]$$

Lásd még R. Bradt, S. Johnson, S. Karlin: On sequential designs for maximizing the sum of n observation, Annals Math. Stat., 27 (1956) 1060-1074. cikkét, ahol alkalmazások is találhatók.

5. Szerencsejáték feladat. Legyen a játékos nyerési valószinüsége minden lépésben p ($0 < p < 1$). Kezdő összege legyen Y_0. Az első lépésben x_1 ($\leq Y_0$) összeget tehet fel a játékos, amit vesztés esetén elveszit, nyerés esetén $Y_0 + x_1 = Y_1$ összegig játszhat a következő lépésben, A játékosnak a j-edik lépésben Y_{j+1} összegü pénze van és x_j ($\leq Y_{j+1}$) összeget tehet meg nyerésre. Az egyes lépések kimenetelei egymástól függetlenek. Feltehetjük továbbá, hogy a játékos haszna Y_j összeg esetén $U(Y_j)$. A feladat adott n esetén $E(U(Y_n))$ maximalizálása.

Legyen $j=0,1,2,\ldots,n$ esetén $V_j(y)$ az $E(U(Y_n))$ maximumát

az $Y_{n-j} = y$ feltétel mellett, azaz $V_j(y) = E(U(Y_n)| Y_{n-j} = y)$.
Nyilván $y \geq 0$ esetén

$$V_0(y) = \sup E(U(Y_n)| Y_n = y) = U(y)$$

és $j=0,1,2,\dots,n-1$ esetén

$$(*) \quad V_{j+1}(y) = \sup[pV_j(y+x) + qV_j(y-x)]$$

A $(*)$ összefüggésből $V_n(Y_0) = \sup E(U(Y_n))$ meghatározható. Továbbá az optimális eljárás $j+1$ megmaradó lépés esetén, amikor is $Y_{n-j-1} = y$, egy olyan x összeg tevése, amelyre $(*)$ jobboldala maximális.

a.) Mutassuk meg, hogy

$$U(y) = y^{\alpha} \quad (\alpha > 1), \quad p < 1/2$$

esetén az optimális eljárás a következő:

1) ha $p > 2^{-\alpha}$ a játékos minden lépésben megteszi a rendelkezésre álló összeget,

2) ha $p \leq 2^{-\alpha}$ nem tesz meg semmit.

b) Mutassuk meg, hogy

$$U(y) = \log y \quad (y > 0), \quad p > 1/2,$$

esetén az optimális eljárás, ha a játkos a rendelkezésére álló összeg $(p-q)$-ad részét teszi meg. Ekkor

$$E(\log Y_n) = \log Y_0 + n(p\log p + q\log q + \log 2).$$

6. Általánositsuk a kettős választás lehetőségét a következő módon. A $p_1(x)$ és $p_2(x)$ sürüségfüggvények a B_1 illetve B_2 kisérlethez tartozhatnak a következő elrendezésben:

w \	B_1	B_2
1	$p_2(x)$	$p_1(x)$
2	$p_1(x)$	$p_2(x)$

$$P(w=1) = \xi_1, \quad P(w=2) = \xi_2 = 1 - \xi_1.$$

Azt mondjuk, hogy a kisérlet <u>sikeres</u> volt, ha a $p_2(x)$ sürüségfüggvényü változóra történt megfigyelés. Nyilván $w=1$ esetén a B_1 kisérlet a sikeres és B_2 a hibás, mig $w=2$ esetén B_2 a sikeres és B_1 a hibás kisérlet. Sikeres kisérlet esetén a nyereség 1, egyébként 0. Ismerve ξ_1 értékét és a $p_1(x)$, $p_2(x)$ sürüségfüggvényeket adjuk meg N (fix) kisérlet esetén azt az optimális politikát, mely maximalizálja a <u>sikeres</u> kisérletek átlagos számát. [Utmutatás: Lásd: D.Feldman: Contributions to the "two-armed bandit" problem, Annals Math. Stat., 33 (1962) 847-856.]

7. A 4. és 6. feladat általánositásaként vizsgáljuk azt az esetet, amikor a w_1 és w_2 valószinüségek mindketten ismeretlenek és tetszőlegesek. [Utmutatás: D. Berry: A Bernoulli two-armed bandit. Annals Math. Stat., 43 (1972) 871-897.]

8. A fejezet eredményeit és feladatait általánositsuk Brown-mozgás folyamatra. [Utmutatás: H. Chernoff: Optimal Stochastic Control, Sankhya, Ser.A.30 (1968) 221-252.]

8. A lineáris filtráció vizsgálata diszkrét Gauss folyamatok esetén

A lineáris filtráció Kálmán-Bucy egyenleteit vizsgáljuk abban a speciális esetben, amikor az ismeretlen idősor lineáris konstans együtthatós differencia egyenletnek tesz eleget, s a megfigyelhető folyamat a nem megfigyelhetőből ugyancsak lineáris leképzéssel adódik, egy additiv zaj hozzáadásával. A tárgyalás ebben az esetben egészen elemi, s nem kiván speciális valószinüségelméleti felkészülést.

Az $(\Omega,\mathcal{F},P)$ valószinüségi mezőn egymással egyszerü sztohasztikus kapcsolatban lévő (θ_n,η_n) valószinüségi változó sorozatokat vizsgálunk, ahol η_n a megfigyelhető θ_n pedig a nem megfigyelhető komponens és θ_n négyzetes középben legjobb becslését kivánjuk megadni. Kálmán és Bucy* mutatták meg, hogy egyszerü esetekben a megoldás a kiinduló egyenletekhez hasonló sztohasztikus differencia (illetve a folytonos esetben differenciál) egyenletet elégit ki, azaz az optimális megoldásra egy algoritmus adható.

Legyen a θ_n n=1,2,..., sorozat Markov tipusu Gauss, mégpedig elégitse ki a

(1) $\theta_n = A_{n-1}\theta_{n-1} + F_n\varepsilon_n, \quad \theta_0 = 0, \quad (E\theta_n = E\varepsilon_n = 0)$

egyenletet, ahol ε_n egy független, egységnyi szórásu Gauss sorozat, és ε_n független $\mathcal{F}_\theta^{n-1}$ -től is, ahol $\mathcal{F}_\theta^{n-1}$ a θ változók által generált σ-algebra $\mathcal{F}_\theta^{n-1} = \sigma(\omega : \theta_{n-1}(\omega), \dots, \theta_0(\omega))$. Feltesszük, hogy θ_n közvetlenül nem megfigyelhető, csak az η_n sorozaton keresztül, ahol

(2) $\eta_n = C_n\theta_n + G_n w_n, \quad n = 1, 2, \dots,$

ahol w_n olyan független, egységnyi szórásu Gauss sorozat, mely - az egyszerüség kedvéért - az $\{\varepsilon_n\}$ sorozattól is független. Az A_n, C_n, F_n, G_n nem azonosan 0 mátrixok, csak az időtől függnek, s függetlenek az egyenletekben szereplő valószinüségi változóktól (a legegyszerübb esetben a folyamatok egydimenziósak).

Jelölje $\mathcal{F}_\eta^n = \sigma\{\omega : \eta_n(\omega), \eta_{n-1}(\omega), \dots, \eta_0(\omega)\}$ az η változók által generált σ-algebrák sorozatát. Bevezetjük a következő valószinüségi változókat

---

* Kalman, R.E., Bucy, R.C.: New results in linear filtering and prediction theory. Trans.ASME Journ.Basic.Engr., 83 D (1961)95-108.

(3) $\hat{\theta}_n = E(\theta_n | \mathcal{F}_\eta^n)$, $n = 1, 2, \ldots,$

(4) $e_n = \theta_n - \hat{\theta}_n$, $d_n = Ee_n^2$, és e_n független $\mathcal{F}_\eta^n$-től $n = 1, 2, \ldots,$

(5) $\xi_n^0 = \eta_n - E(\eta_n | \mathcal{F}_\eta^{n-1})$, $n = 1, 2, \ldots.$

A feltételes várható érték definiciója alapján $\hat{\theta}_n$ és ξ_n^0 is olyan Gauss eloszlásu változók, melyek $\eta_1, \eta_2, \ldots \eta_n$ lineáris kombinációi, továbbá ξ_n^0 független $\mathcal{F}_\eta^{n-1}$-től és igy egy független sorozatot alkot. A definiciók felhasználásával adódik ξ_n^0 következő előállitása

(6) $\xi_n^0 = C_n\theta_n + G_n w_n - E(C_n\theta_n + G_n w_n | \mathcal{F}_\eta^{n-1}) =$

$= C_n\theta_n + G_n w_n - C_n A_{n-1}\hat{\theta}_{n-1} = C_n\theta_n + G_n w_n - C_n A_{n-1}(\theta_{n-1} - e_{n-1})$

$= C_n F_n \epsilon_n + G_n w_n + C_n A_{n-1} e_{n-1}.$

A $\xi_1^0, \xi_2^0, \ldots, \xi_n^0$ változók az $\eta_1, \eta_2, \ldots, \eta_n$ változók olyan lineáris kombinációi, melyekből egyértelmüen kölcsönösen meghatározhatók a ξ_i^0 és η_i változók. Ugyanis a leképezés mátrixa olyan háromszög mátrix, melynek fődiagonálisában 1-ek vannak. Ez azt jelenti, hogy a ξ_i^0 által generált σ-algebra sorozat megegyezik az η_i által generált σ-algebrák sorozatával: $\mathcal{F}_{\xi^0}^n = \mathcal{F}_\eta^n$, $n=1,2,\ldots$

Tekintsük a következő valószinüségi változó sorozatot

(7) $\xi_n^1 = \hat{\theta}_n - A_{n-1}\hat{\theta}_{n-1} = (\hat{\theta}_n - \theta_n) + A_{n-1}(\theta_{n-1} - \hat{\theta}_{n-1}) + (\theta_n - A_{n-1}\theta_{n-1}) =$

$= -e_n + A_{n-1}e_{n-1} + F_n\epsilon_n$, $n = 1, 2, \ldots,$

könnyen látható, hogy (7)-ben értelmezett ξ_n^1 sorozat ugyancsak független Gauss sorozat. Ugyanis (7) jobb oldalán minden változó független F_η^{n-1}-től, és igy - mivel ξ_m^1 $\mathcal{F}_\eta^m$ mérhető - ξ_n^1 független ξ_m^1-től is, ha $m<n$.

Mivel ξ_n^1 $\mathcal{F}_\eta^n$-mérhető és egyben $\mathcal{F}_{\xi^0}^n$ mérhető, nemcsak az $\eta_1, \eta_2, \ldots,$ η_n változók, hanem a $\xi_1^0, \ldots, \xi_n^0$ változók lineáris függvénye is. Továbbá ϵ_n és $\mathcal{F}_\eta^{n-1}$ (tehát $\mathcal{F}_{\xi^0}^{n-1}$) függetlensége miatt

(8) $\xi_n^1 = \hat{\theta}_n - A_{n-1}\hat{\theta}_{n-1} = E(\theta_n | \mathcal{F}_\eta^n) - A_{n-1}E(\theta_{n-1} | \mathcal{F}_\eta^{n-1}) =$

$= E(\theta_n | \mathcal{F}_{\xi^0}^n) - A_{n-1}E(\theta_{n-1} | \mathcal{F}_{\xi^0}^n) = E(\theta_n - A_{n-1}\theta_{n-1} | \mathcal{F}_{\xi^0}^n) =$

$= E(F_n\epsilon_n | \mathcal{F}_{\xi^0}^n) = K_n\xi_n^0.$

Innen ξ_n^0-al való szorzással

(9) $E(\xi_n^1 \xi_n^0) = K_n E(\xi_n^0)^2$.

A K_n együttható meghatározása a következőképpen történik. Egyrészt (6) alapján

$$(10)\quad E(\xi_n^0 \xi_n^0) = E[C_n F_n \epsilon_n + G_n w_n + C_n A_{n-1} e_{n-1}]^2 = \\ = (C_n F_n)^2 + G_n^2 + (C_n A_{n-1})^2 d_{n-1},$$

ahol felhasználtuk ϵ_n és w_n függetlenségét, valamint ezek függetlenségét (4) és (1) alapján e_{n-1}-től. Továbbá (7) és (6) alapján

$$(11)\quad E(\xi_n^1 \xi_n^0) = E(-e_n + A_{n-1} e_{n-1} + F_n \epsilon_n)(C_n F_n \epsilon_n + G_n w_n + C_n A_{n-1} e_{n-1}) = \\ = E(A_{n-1} e_{n-1} + F_n \epsilon_n)(C_n F_n \epsilon_n + G_n w_n + C_n A_{n-1} e_{n-1}) = \\ = C_n (F_n)^2 + A_{n-1} C_n A_{n-1} d_{n-1},$$

felhasználva e_n és ξ_n^0 függetlenségét, valamint a már emlitett többi változó függetlenségét. (9)-ből (10) és (11) alapján

$$(12)\quad K_n = \frac{E(\xi_n^1 \xi_n^0)}{E(\xi_n^0)^2} = [C_n (F_n)^2 + A_{n-1} C_n A_{n-1} d_{n-1}][(C_n F_n)^2 + G_n^2 + (C_n A_{n-1})^2 d_{n-1}]^{-1}.$$

A (8) összefüggésből a $\hat{\theta}_n$ sorozatra a következő rekurziót kapjuk

$$(13)\quad \hat{\theta}_n = A_{n-1} \hat{\theta}_{n-1} + K_n \xi_n^0$$

ahol ξ_n^0 az (5) szerinti független sorozat.

A d_n szórásnégyzetek sorozatára egy differencia egyenlet irható fel.

Ugyanis $d_0 = 0$

$$d_n = E(\theta_n - \hat{\theta}_n)^2$$

továbbá

$$(14)\quad d_n = E[(\theta_n - \hat{\theta}_n)\theta_n - (\theta_n - \hat{\theta}_n)\hat{\theta}_n] = \\ = E(\theta_n)^2 - E(\theta_n \hat{\theta}_n) = E(\theta_n)^2 - E(E(\theta_n \hat{\theta}_n | F_\eta^n)) = E(\theta_n)^2 - E(\hat{\theta}_n)^2.$$

Viszont (1) alapján

$$E(\theta_n)^2 = A_{n-1}E(\theta_{n-1})^2 A_{n-1} + (F_n)^2$$

és (13) alapján (1o) felhasználásával

$$E(\hat{\theta}_n)^2 = A_{n-1}E(\hat{\theta}_{n-1})^2 A_{n-1} + K_n E(\xi_n^0)^2 K_n .$$

Behelyettesitve (14)-be a két utolsó összefüggést és K_n (12) szerinti értékét

(15) $$d_n = A_{n-1}E(\theta_{n-1})^2 A_{n-1} + (F_n)^2 - [A_{n-1}E(\hat{\theta}_{n-1})^2 A_{n-1} + \frac{[E(\xi_n^1 \xi_n^0)]}{E(\xi_n^0)^2} =$$

$$= A_{n-1}d_{n-1}A_{n-1} + (F_n)^2 - \frac{[C_n(F_n)^2 + A_{n-1}C_nA_{n-1}d_{n-1}]^2}{(C_nF_n)^2 + G_n^2 + (C_nA_{n-1})^2 d_{n-1}}$$

adódik.

A fenti tárgyalás akkor is igaz marad, ha θ és η vektor folyamatok. A megfelelő képletek felirását az olvasó is elvégezheti a számitások végig követésével.

Ha az (1) és (2) lineáris egyenletek helyett a következő egyenletek teljesülnek

(16) $$\theta_n = A(\theta_{n-1}, \eta_{n-1}, n) + F_n \epsilon_n \qquad \theta_0 = 0$$

és

(17) $$\eta_n = C(\theta_{n-1}, \eta_{n-1}, n) + G_n w_n ,$$

ahol, mint az előbbiekben ϵ_n és w_n független Gauss sorozatok, melyek egymástól is függetlenek, a nemlineáris filtrációra jutunk. Ekkor a (θ_n, η_n) pár egy nemlineáris sztohasztikus egyenletrendszer megoldása. Az előbbiekkel ellentétben, ha A(x,y,n) és C(x,y,n) nemlineáris függvények θ_n és η_n nem lesznek többé Gauss eloszlásu változók.

Megmutatható, hogy a

$$\hat{\theta}_\eta = E(\theta_n \mid F_\eta^n)$$

sorozat ekkor is kielégit egy (7)-hez hasonló differencia egyenletet. A megfelelő, de időben folytonos eset tárgyalása megtalálható Lipcer és Sirjajev* összefoglaló dolgozatában.

* Lipcer, P.S., Sirjajev, A.N.: Nyelinejnaja filtracija diffuzionnüh processzov. Trudü MIAN 104 (1968)

9. Markov láncok vezérlése

Legyen $\xi(0) \longrightarrow \xi(1) \longrightarrow \xi(2) \longrightarrow \dots$ egy Markov lánc, melynek állapottere $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n, \dots$ $(E \ni \varepsilon_i)$.
Feltesszük, hogy a $\{p_{ij}\}$ átmenetvalószinüségek függnek egy d vezérlő paramétertől is: $p_{ij}(d)$, $(d \in D)$.

A vezérlés célja: Az E egy állapotába vagy állapothalmazába vezérelni a $\xi(0) = \varepsilon_0$ állapotból kiinduló folyamatot. A vezérlés csak nagy valószinüséggel történhet. A P valószinüség függ a vezérléstől.

A vezérlés módja: minden ε_i és n értékhez egy $d(\varepsilon_i,n)$ vezérlési (döntési) paraméter érték tartozik.

Az egész program a $d = d(x,t)$ (ahol $x=\varepsilon_1, \varepsilon_2, \dots$; $t = 0,1,\dots$) döntésfüggvénnyel van megadva. Ha $d(x,t)$ adott a folyamat lefolyása a következő: ha $\xi(t)$ $t=n$ esetén ε_i-ben van $p_{ij}=p_{ij}(d)$ (ahol $d=d(\varepsilon_i,n)$) valószinüséggel ε_j-be megy át.
A kiválasztott állapotba jutás valószinüsége d-től függ: $P=P(d)$.
$d^o = d^o(x,t)$ optimális, ha

$$P(d^0) = \max_d P(d).$$

Példa. Fix n értékre maximális valószinüséggel E_0-ba vezetni a folyamatot.

$$P(d) = P\{\xi(n) \in E_0\}.$$

Legyen $P\{k,i,d\}$ a következő

$$P\{k,i,d\} = P\{\xi(n) \in E \mid \xi(k) = \varepsilon_i, d\}$$

A teljes valószinüség tétele alapján

$$P\{k,i,d\} = \sum_j p_{ij}(d) P\{k+1,j,d\}, \qquad d = d(\varepsilon_i, k)$$

Ha k=n

$$P(n,j,d) = \begin{cases} 1, & \text{ha } \varepsilon_j \in E_0, \\ 0, & \text{egyébként,} \end{cases}$$

és

$$P(n-1,i,d) = \sum_{j:\varepsilon_j \in E_0} P(n,j,d) p_{ij}(d).$$

Ez csak $d(\varepsilon_i,n-1)$-től függ, legyen d^o az az érték, amely minimalizálja ezt a kifejezést:

$$P^0(n-1,i) = P(n-1,i,d^0) = \max_{d\in D} P(n-1,i,d)\cdot$$

(feltesszük, hogy elérik maximumukat).

$k = n-2$ esetén

$$P(n-2,i,d) = \sum_j p_{ij}(d)P(n-1,j,d),$$

adódik, ahol $p_{ij}(d)$ csak a $d=d(\varepsilon_i,n-2)$-től mig $P(n-1,j,d)$ csak $d(\varepsilon_j,n-1)$-től függ. A vezérlésben $d^0(\varepsilon_j,n-1)$-et véve $P(n-2,i,d)$ növekszik, és

$$\tilde{P}(n-2,i,d) = \sum p_{ij}(d)P^0(n-1,j)$$

csak $d(\varepsilon_i,n-2)$ változtatásával növelhető. Legyen $d^0\in D$ az az érték, mely maximalizálja a

$$P^0(n-2,i) = \max_{d\in D} \tilde{P}(n-2,i,d) = \tilde{P}(n-2,i,d^0).$$

kifejezést. Ha $d(x,t)$ az $x=\varepsilon_1$, ε_2, ... ; $t=n-2$, $n-1$ értékekre a fenti $d^0(x,t)$ akkor $P(k,i,d)$ a maximális $P^0(k,i)$ érték $k=n-2,n-1$ és minden $i=1,2,\dots$-re.

Ebből a d^0-ból kiindulva

$$\tilde{P}(n-3,i,d) = \sum p_{ij}(d)P^0(n-2,j)$$

melyet megint maximalizálunk.

A $$P^0(k,i) = \max_{d\in D}\sum_j p_{ij}(d)P^0(k+1,j)$$

összefüggést, melynek alapján keressük a $d^0=d^0(x,t)$ optimális döntésfüggvényt, Bellman egyenletnek szokás nevezni. A fenti eljárást a következő példákkal világitjuk meg.

1. Példa. Legyenek $\alpha_1 \le p_{11}(d) \le \beta_1$, $\alpha_2 \le p_{21}(d) \le \beta_2$ és a Markov lánc két állapotu: ε_1, ε_2. Azt a valószinüséget kivánjuk maximalizálni, amely a 2. lépésben ε_1-ben tartózkodását adja meg.

Nyilván

$$P^0(1,1) = \beta_2, \quad P^0(1,2) = \beta_2.$$

és

$$P^0(0,1) = \max_d [p_{11}(d)\beta_2 + p_{12}(d)\beta_2] = \max_d [p_{11}(d)(\beta_1-\beta_2)+\beta_2]$$

Innen látszik, hogy $\beta_1>\beta_2$ esetén $p_{11}=\beta_1$, mig $\beta_1<\beta_2$ esetén a $p_{11}=\alpha_1$ választás a jó.

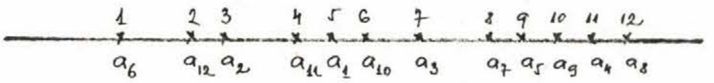
2. Példa. <u>A legjobb elem kiválasztása</u>. (Elemi tárgyalás)

Tegyük fel adva van m darab elem, amelyek közül a legjobbat kell kiválasztani. Nem azt kell megállapitani, melyik a legjobb, hanem olyan eljárást kidolgozni, amelynek alapján legnagyobb valószinüséggel a legjobb elemet választjuk ki.

Az eljárás a következő: megismerkedve az első elemmel vagy megállapodunk ennél, vagy ugy döntünk, hogy még egy elemet megvizsgálunk. A második elemnél ugyanigy dönthetünk: vagy megállapodunk ennél az elemnél, vagy további elemet vizsgálunk meg. Megismerkedve egy elemmel vagy kiválasztjuk, vagy további vizsgálat mellett döntünk. A korábban megvizsgált elemekhez visszatérni nincs lehetőségünk.

Annak a megkötésnek, hogy korábban megvizsgált elemhez nincs lehetőség visszatérni, igen sok esetben van értelme (de a megvizsgált elemet megsemmisitjük). Az a megszoritás, hogy <u>m</u> ismert, már kevésbé tünik természetesnek.

Ábrázoljuk m=12 esetén az elemeket az egyenesen, nagyság szerinti sorrendben megszámozva

(1) 

Jelölje a kiválasztási sorrendben a_1, a_2, ... az első, második és i.t. elemet. Mivel a megvizsgálás teljesen véletlenszerüen történik, az a_1 $\frac{1}{m}$ valószinüséggel lehet az elemek bármelyike: példánkban annak valószinüsége, hogy a_1 a legnagyobb (12-ik)$\frac{1}{12}$. Az a_2 is egyenlő ($\frac{1}{m-1}$) valószinüséggel lehet a megmaradó elemek bármelyike. (Képzeljük el, hogy a megismert pontokban lámpák gyulnak.) A feladat az n-edik pont felismerése a lehető legnagyobb valószinüséggel, amikor megjelenik.

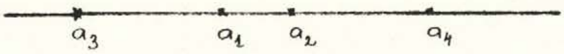
Ha az eljárás egy meghatározott helyen való megállást ir elő (pl. az első megvizsgált elem, vagy az utolsó megvizsgált elemnél való kiválasztást), annak valószinüsége $\frac{1}{m}$ (s ha $n \longrightarrow \infty$ ez a valószinüség 0).

Hogy van ($m \longrightarrow \infty$ esetén is) jó stratégia, azt a következő meggondolás is mutatja: vizsgáljunk meg $\frac{m}{2}$ elemet, s ezután válaszszuk ki azt, mely minden előzőnél nagyobb. Annak valószinüsége, hogy a második legnagyobb elem az első $\frac{m}{2}$ megvizsgált között van, s a legnagyobb elem a második $\frac{m}{2}$ között lesz, nyilván

$$\frac{m/2}{m} \cdot \frac{m/2}{m-1} > \frac{1}{4}$$

A: a második legjobb elem az első $\frac{m}{2}$ közé
B: a legjobb elem a második $\frac{m}{2}$ közé esik
$P(AB) = P(A\ B) . P(B)$

Természetes kiválasztási szabály: ott nem megállni, amelynél volt jobb.



Annak valószinüsége, hogy a_{k+1} a $[-\infty, a^*)]$, $[a_1^*, a_2^*]$,...., $[a_k^*, \infty]$ intervallumok közül valamelyikbe esik:

$$\frac{1}{(k+1)!}\Big/\frac{1}{k!} = \frac{1}{k+1} \qquad [P(A/B) = \frac{P(AB)}{P(B)}$$

A: $a_1,\dots,a_{k+1}$ meghatározott sorrendje

B: $a_1,\dots,a_n$ meghatározott sorrendje

Legyen $\xi(0) = 1$, $\xi(1)$ az a sorszám, melyhez tartozó elem jobb az összes előbbinél, $\xi(2)$ a következő összes előbbinél jobb elem sorszáma. (A fenti példában $\xi(0) = 1$, $\xi(x) = 2$, $\xi(2) = 4$. Az első rajzon $\xi(0) = 1$, $\xi(1) = 3$, $\xi(2) = 4$, $(3) = 8$.)

Ha B: a k megvizsgált közül a k-adik a legjobb,

A: a k-adik az összes közül a legjobb.

Akkor

$$P(A|B) = \frac{P(AB)}{P(B)} = \frac{P(A)}{P(B)} = \frac{\frac{(m-1)!}{m!}}{\frac{(k-1)!}{k!}} = \frac{k}{m}$$

Hasonlóan látható, hogy

$$P\{\xi(i+1)=\ell \mid \xi(i)=k\} = \frac{P\{\xi(i)=k, \xi(i+1)=\ell\}}{P\{\xi(i)=k\}} = \frac{k}{\ell(\ell-1)}$$

azaz, homogén Markov lánc.

Meg kell mutatni, hogy ha a_k maximális elem (azaz $a_k^*=a_k$), akkor a döntés (megállni vagy folytatni a megfigyelést) csak az a_k elem sorszámától (k) függ, s nem függ az $a_1, a_2, \dots a_{k-1}$ elemek elhelyezkedésétől. Azaz az $a_k, a_{k+1},\dots,a_m$ elemek elrendezésének feltételes valószinüsége az $a_1, a_2,\dots, a_k$ feltétel mellett nem függ az $a_1,\dots,a_{k-1}$ elemek elhelyezkedésétől.

Ehhez elegendő belátni, hogy nemcsak az $a_k=a_k^*$, $a_{k+1},\dots, a_m$ pontok elhelyezkedése, hanem $a_1^* < a_2^* \dots$ a_{k-1}^*-hez viszonyitott elhelyezkedésük sem függ az $a_1^*,\dots, a_k^*$ pontok elhelyezkedésétől.

Már láttuk, hogy a_{k+1} egy $[-\infty, a_1^*]$,...., $[a_k^*,\infty]$ intervallumba esésének valószinüsége $\frac{1}{k+1}$, a_{k+2}-nek az előbbiek valamelyikébe esése $\frac{1}{k+2}$ valószinüségü, igy $a_k, a_{k+1},\dots,a_m$ feltételes valószinüsége

$$\frac{1}{k+1}\cdot\frac{1}{k+2},\dots\frac{1}{m}$$

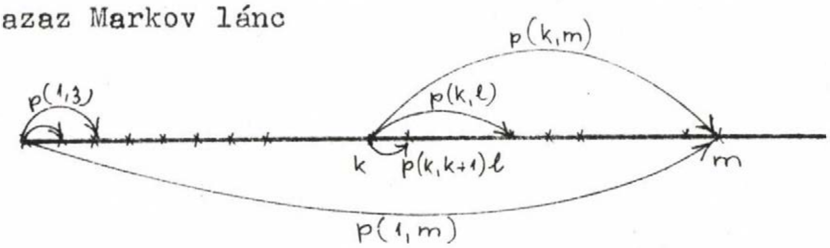
és ez független az $a^*_1, \dots, a^*_k$ megjelenési sorrendjétől.

A feladatot igy teljesen a $\xi(i)$ sorozat megfigyelésére vezettük vissza.

Másrészt (éppen most láttuk be)

$$P\{\xi(i+1)=\ell/\xi(i)=k,\ \xi(i-1)=k_1,\dots,\ \xi(0)=1\}=P\{\xi(i+1)=\ell/\xi(i)=k\}=p(k,\ell)$$

azaz Markov lánc



Állapottér

Véletlen bolyongás m állapottal (m elnyelőhely). A valószinüségekre

$\sum\limits_{\ell=k+1} p(k,\ell) \le 1$ $\quad (1-\sum\limits_{\ell} p(k,\ell)$ hogy k után m lesz több maximális pont, azaz $a_k = \max(a_1,\dots,a_m))$.

A

$$p(k,\ell)=\begin{cases}0 & ,\text{ha } \ell \le k\\ \frac{k}{\ell(\ell-1)} & ,\text{ha } (1 \le k < \ell \le m).\end{cases}$$

Az optimális módszer (stratégia) kiválasztása minden k-ra annak megadása, hogy megállni vagy sem. Mivel k lehetséges értékei $1,2,\dots,m$, azon részhalmazt kell megadni, ahol meg kell állni Γ az $(1,2,\dots,m)$ részhalmazával meg lehet adni a stratégiát. Ezen stratégiák száma véges 2^n (az üres és teljes halmazt is beleértve), ezek közül kell kiválasztani a legjobbat.

A stratégiák közül érdekes a következő (k legyen fix szám)

$$\tau=\xi(i), \text{ ha } \quad \xi(0)<\dots<\xi(i-1)<k \quad \text{és} \quad \xi(i)\ge k.$$

Választhatjuk τ értékét $\xi(i+1)$-nek is Ez nem optimális stratégia. Adott k esetén legyen $q(k)$ az "eltünés" feltételes valószinüsége a megállásnak a $\xi(i)=k$ feltétel mellett (hogy a k-adik abszolut legjobb is)

$$q(k)=1-\sum_{\ell} p(k,\ell)=1-\sum_{\ell=k+1}^{m}\frac{k}{\ell(\ell+1)}=1-k\sum\left(\frac{1}{\ell-1}-\frac{1}{\ell}\right)=\frac{k}{m}$$

Jelölje $q'(k)$ annak valószinüségét, hogy a "legjobb" a választás (a nyereséget), ha $\xi(i)=k$ és a $\xi(i+1)$-nél állnak meg. Nyilván

$$q'(k)=\sum_{\ell=k+1} p(k,\ell)\, q(\ell)=\sum \frac{\ell}{m}\,\frac{k}{\ell(\ell-1)}=\frac{k}{m}\left(\frac{1}{k}+\frac{1}{k+1}+\dots+\frac{1}{m-1}\right)=q(k)\left(\frac{1}{k}+\right.$$

$$\left.+\dots+\frac{1}{m-1}\right) \quad (k<m) \quad q'(k)=0 \text{ ha } k=m.$$

Az $\frac{1}{n}+\dots+\frac{1}{m-1}$ összeg monoton csökkenők, növelésével $\frac{q'(k)}{q(k)}$ és monoton csökkenő $\frac{q'(m)}{q(m)}=0$. Igy $q'(k)\leq q(k)$ valamilyen k_m, $k_m+1, \dots, m$ szakaszon.

A $\Gamma=\{k_m, k_m+1, \dots m\}$ halmaznak optimális stratégia felel meg. Azaz a megfigyelést mindaddig folytatni kell, mig $\xi(i)<k_m$, és megállni, ha $\xi(i)\geq k_m$.

Legyen $m\geq 3$, akkor

$$q'(1)=q(1)\left(1+\frac{1}{2}+\dots\frac{1}{m-1}\right)>q(1)$$

azaz az elsőt nem fogadjuk el, mint optimálisat.

Ha az A stratégia a $k<k_m$-nél megállást irna elő, az A' stratégia mely eggyel továbbá lépne jobb lenne nála, ha $p_A(k)$ az A esetén a k-ba jutás valószinüsége (nyilván $p_A(k)>0$) a nyereségek $p_A(k)\, q(k)$ ill. $p_A(k)\, q'(k)$.

Indukcióval belátható, hogy (k-ban visszafelé) az optimális A stratégia a (k_m+1, k_m+2, ..., k) pontokban azonnali megállást követel.

Ezen a szakaszon $q'(k)<q(k)$. Ha az A stratégia m-nél továbblépést javasol, akkor A', mely megegyezik A-val, csak m-ben megállást javasol - jobb lesz, mert nyeresége $p_A(m)$-mel nagyobb, mint A-é. Azaz k=m esetén igaz az állitás, tegyük fel k+1, k+2,..., m esetén már igaz ($k\geq k_m+1$). Ha A a k kihagyását (folytatást) javasolna, akkor az A', mely megegyezik A-val, csak k-ban megállást javasol, igy A' jobb, mint A. Valóban az indukció miatt A a k+1-ben való megállás, A' a k-ban való megállás. Az A nyeresége és A' nyeresége közti különbség $p_A(k)\, q(k) - p_A(k)\, \dot{q}(k)<0$, azaz k-ban meg kell állni.

Ha $q'(k_m)<q(k_m)$ akkor k_m is Γ-hoz tartozik (egyenlőség esetén is vegyük Γ-hoz, igy egyértelmü).

A k_m kiválasztása az a legkisebb k, melyre $q'(k)$, azaz

$$\frac{1}{k}+\frac{1}{k+1}+\dots+\frac{1}{m-1}\leq 1 \text{ vagy } \left(\frac{1}{k_m}+\dots+\frac{1}{m-1}\leq 1<\frac{1}{k_m-1}+\frac{1}{k_m}+\dots\frac{1}{m-1}\right).$$

Az optimális stratégia a valószinüségre a következőt adja

$$p_m = \frac{k_m - 1}{m}\left(\frac{1}{k_m - 1} + \frac{1}{k_m} + \dots + \frac{1}{m-1}\right).$$

Bizonyitás. Annak valószinüsége, hogy $k_m - 1$ megvizsgált elem után a legelső, mely jobb, mint a többi, k sorszámu lesz (a_k - szélső elem) $\frac{k_m - 1}{k(k-1)}$ (a_k a szélső elem, s a legközelebbi hozzá az a_1, ..., $a_{k_m} - 1$ elemek bármelyike). A siker feltételes valószinüsége $\frac{k}{m}$ és igy

$$p_m = \sum_{k_m} \frac{k_m - 1}{k(k-1)} \cdot \frac{k}{m} = \frac{k_m - 1}{m}\left(\frac{1}{k_m - 1} + \frac{1}{k_m} + \dots + \frac{1}{m-1}\right).$$

m=10 esetén

k	$\frac{1}{k}$	$\frac{1}{k}+\dots+\frac{1}{m-1}$	k	$\frac{1}{k}$	$\frac{1}{k}+\dots+\frac{1}{m-1}$
9	0,111	0,111	4	0,250	0,996
8	0,125	0,236	3	0,333	1,329
7	0,143	0,379	2	0,500	
6	0,167	0,546	1		
5	0,200	0,746			

$$p_{10} = 0,399$$

Ezután vizsgáljuk meg a feladatot, mint Markov lánc irányitását. Az állapotot ε_1, ..., ε_m és ε_{m+1} jelenti, ha már megvolt a legjobb elem. Az átmenetek valószinüségei

$$p_{ij} = \begin{cases} 0 & \text{ha} \quad j \le i \\ \frac{i}{j(j-1)} & \text{ha} \quad i < j \le m \\ \frac{i}{m} & \text{ha} \quad j = m+1 \end{cases}, \quad \text{megállás esetén} \quad p_{ij} = \begin{cases} 1 & i = j \\ 0 & i \ne j \end{cases}$$

d lehetséges értékei: 0(megállás), 1 (folytatás).

Olyan döntésfüggvény kell, melyre a legjobb (legnagyobb) elem kiválasztásának valószinüsége maximális.

Legyen p_i annak valószinüsége, hogy a folyamatot az ε_i állapotban állitjuk meg. A

$$P = \sum_i \frac{i}{m} p_i$$

valószinüséget maximalizálni akarjuk. Ha P(k,d) annak valószinüsége, hogy valahol ε_k-ban van folyamat, és a legjobbat fogjuk kiválasztani, akkor

$$P(k,d) = \sum_{j=k}^{m} p_{kj}(d) P(j,d)$$

Mivel

$$P(m,d) = \begin{cases} 1 & \text{, ha } \quad d^0(\varepsilon_m) = 0, \\ 0 & \text{, ha } \quad d(\varepsilon_m) = 1, \end{cases}$$

továbbá

$$P(m-1,d) = \begin{cases} \frac{m-1}{m} & \text{, ha } \quad d(\varepsilon_{m-1}) = 0 \\ \frac{m-1}{m(m-1)} & \text{, ha } \quad d(\varepsilon_{m-1}) = 1 \end{cases}$$

adódik, hogy

$$P^0(m-1) = \frac{m-1}{m}$$

A meggondolást folytatva a már ismert eredményre jutunk. U.i., ha $\varepsilon_k, \varepsilon_{k+1}, \dots, \varepsilon_m$-ben $d^0 = d^0(x) = 0$ milyen legyen $d^0(\varepsilon_{k-1})$?

Nyilván

$$P(k-1,d) = \begin{cases} \frac{k-1}{m} & \text{, ha } \quad d(\varepsilon_{k-1}) = 0 \\ \frac{k}{m} \quad \frac{k-1}{(k-1)k} + \frac{k-1}{m} \quad \frac{k-1}{k(k+1)} + \dots + \frac{k-1}{m(m-1)} & \text{, ha } \quad d(\varepsilon_{k-1}) = 1, \end{cases}$$

ahonnan látható, hogy

$$d^0(\varepsilon_{k-1}) = \begin{cases} 0 & \text{, ha } \quad \frac{1}{k-1} + \frac{1}{k} + \dots + \frac{1}{m-1} \leq 1 \\ 1 & \text{egyébként.} \end{cases}$$

p_m és k_m-re közelités található nagy m esetén

$$\ln \frac{m}{k} < \frac{1}{k} + \frac{1}{k+1} + \dots + \frac{1}{m-1} < \ln \frac{m-1}{k-1}$$

igy

$$\ln \frac{m}{k_m} < 1 < \ln \frac{m-1}{k_m - 2},$$

ahonnan

$$\frac{m}{e} < k_m < \frac{m}{e} + \left(2 - \frac{1}{e}\right)$$

és

$$\frac{1}{k_m - 1} + \frac{1}{k_m} + \dots + \frac{1}{m-1} \longrightarrow 1$$

miatt

$$p_m < \lim \frac{k_m - 1}{m} \longrightarrow \frac{1}{e} \sim 0,368$$

Jelen dolgozat az 5.11.3 számu intézeti kutatási téma keretében került kidolgozásra

A TANULMÁNYOK sorozatban eddig megjelentek:

1/1973 Pásztor Katalin: Módszerek Boole-függvények minimális vagy nem redundáns, $\{\wedge, \vee, \neg\}$ vagy {NOR} vagy {NAND} bázisbeli, zárójeles vagy zárójel nélküli formuláink előállitására

2/1973 Вашкеви Иштван: Расчленение многосвязных промышленных процессов с помощью вычислительных машин

3/1973 Ádám György: A számitógépipar helyzete 1972 második felében

4/1973 Bányász Csilla: Indentification in the Presence of Drift

5/1973* Gyürki J.- Laufer J.- Grint M.- Somló J.: Optimalizáló adaptiv szerszámgépirányitási rendszerek

6/1973 Szelke Erzsébet - Tóth Károly: Felhasználói Kézikönyv (USER MANUAL) a Folytonos Rendszerek Szimulációjára készült ANDISIM programnyelvhez

7/1973 Legendi Tamás: A CHANGE nyelv/multiprocesszor

8/1973 Klafszky Emil: Geometriai programozás és néhány alkalmazása

9/1973 R. Narasimhan: Picture Processing Using Pax

10/1973 Dibuz Ágoston - Gáspár János - Várszegi Sándor: MANU - WRAP hátlaphuzalozó. MSI-TESTER integrált áramköröket mérő, TESTOMAT - C logikai hálózatokat vizsgáló berendezések ismertetése

11/1973 Matolcsi Tamás: Az optimum-számitás egy új módszeréről

12/1973 Makroprocesszorok, programozási nyelvek. Cikkgyüjtemény az NJSzT és SzTAKI közös kiadásában.
Szerkesztette: Legendi Tamás

13/1973 Jedlovszky Pál: Uj módszer bonyolult retifikáló oszlopok vegyészmérnöki számitására

14/1973 Bakó András: MTA kutatóintézeteinek bérszámfejtése számitógéppel

15/1973 Ádám György: Kelet-nyugati kapcsolatok a számitógépiparban

16/1973 Fidrich Ilona - Uzsoky Miklós: LIDI-72 listakezelő rendszer a DIgitális Osztályon, 1972. évi változat

17/1974 Gyürki József: Adaptiv termelésprogramozó rendszer (APS) termelő mühelyek irányitására

18/1974 Pikler Gyula: MINI-számitógépes interaktiv alkatrészprogramiró rendszer NC szerszámgépek automatikus programozásához

19/1974 Gertler, J.-Sedlak, J.: Software for process control

20/1974 Vámos, T.-Vassy, Z.: Industrial pattern Recognition Experiment - A Syntax Adided Approach

21/1974 A KGST I.-15-1.: Diszkrét rendszerek automatikus tervezése c. témában 1973. februárban rendezett szeminárium előadásai

22/1974 Arató, M.-Benczúr, A.-Krámli, A.-Pergel, J.: Stochastic Processes, Part I.

23/1974 Benkó Sándor-Renner Gábor: Erősen telitett mágneses körök tervezési módszere

24/1974 Kovács György-Franta Lászlóné: Programcsomag elektronikus berendezések hátlaphuzalozásának tervezésére

25/1974 Járdán R. Kálmán: Háromfázisú tirisztoros inverterek állandósult tranziens jelenségei és belső impedanciája

26/1974 Gergely József: Numerikus módszerek sparse mátrixokra

27/1974 Somló János: Analitikus optimalizálás

28/1974 Vámos Tibor: Tárgyfelismerési kisérlet nyelvi módszerekkel

29/1974 Móritz Péter: Vegyészmérnöki számitási módszerek fázisegyensúlyok és kémiai egyensúlyok vizsgálatára

30/1974 Vámos, T.-Vassy, Z.: THE BUDAPEST ROBOT - Progmatic intelligence -

31/1975 Nagy István: Frekvenciásos, középfrekvenciás inverter elmélete

32/1975 Singer-Borossay-Koltai: Gázhálózatok optimális irányitása különös tekintettel a Fővárosi Gázmüvek hálózataira

33/1975 Vámos, T.-Vassy, Z.: Limited and Pragmatic Robot Intelligence

Mérő, L.-Vassy, Z.: A Simplified and Fastened Version of the Hueckel Operator for Finding Optimal Edges in Pictures

Галло В.: Программа для распознавания геометрических образов, основанная на лингвистическом методе описания и анализа геометрических структур

34/1975 Nemes László: Pattern Identification Method for Industrial Robots by Extracting the main Features of Objects

35/1975 Garádi János-Krámli András-Ratkó István-Ruda Mihály: Statisztikai és számitástechnikai módszerek alkalmazása kórházi morbiditás vizsgálatokban

36/1975 Renner Gábor: Elektromágneses tér számitása nagyhőmérsékletü anyagban

37/1975 Edgardo Felipe: Specification problems of a process control display

38/1975 Hajnal Andrásné: Nemlineáris egyenletrendszerek megoldási módszerei

39/1975* A. Abd. El Sattar: Control of induction motor by Three ...

40/1975 Gerhardt Géza: QDT Grafikus interaktiv szubrutinok a CDC 3300 - GD'71 grafikus konfigurációra

41/1975 Arató M.-Benczur A.-Krámli A.-Pergel J.: Stochastic Processes, Part II.

A * -gal jelölt kivételével a sorozat kötetei az Intézet könyvtárában megrendelhetők /Budapest, XIII. Victor Hugo u. 18-22./