

Noszály Csaba

April 4, 2013

## Nagy számok törvénye

Szerzőtársaimmal többindexes valségi változó sorozatokra általánosítottunk egy nagy számok erős törvényeinek bizonyítására szolgáló módszert, melyet fel is használtunk - többek között - logaritmikusan súlyozott, szuperadditív momentum-struktúrájú sorozatok illetve martingálok majdnem biztos konvergenciájának bizonyítására.

## Nagy számok törvénye

Cs.Noszály, T.Tómacs

A general approach to strong laws of large numbers for fields of random variables

Ann.Univ.Sci.Budap.Rol.Eötvös,Sect.Math. 43,61-78(2000)

O.Klesov, I.Fazekas, Cs.Noszály, T.Tómacs

Strong laws of large numbers for sequences and fields.

Theory Stoch.Process. 5(21),No.3-4,91-104(1999)

## Többdimenziós minták távolsága

Tusnády Gábor vetette fel a következő kérdést: Adott két minta,  $X_i, Y_i \in \mathbb{R}^2$  ( $i = 1, \dots, n$ ) esetén a

$$T_n = \min_{\pi \in S_n} \sum_{i=1}^n \|X_i - Y_{\pi i}\|$$

valségi változó szolgáltat-e elég információt az  $F_X$  és  $F_Y$  egyezéséről illetve különbözőségéről.

## Többdimenziós minták távolsága

Szimulációkat végeztem normális és egyenletes mintákat alapul véve. Ezen esetekben  $T_n$  jól közelíthető általánosított extrémérték eloszlással - sejtésem hogy ez sokféle eloszláspár esetén igaz és az eloszlások különbözőségéről az illesztett eloszlás paraméterei adhatnak információt.

# Többdimenziós minták távolsága

Cs. Noszály

Experiments on the distance of two-dimensional samples

Ann.Math.et Informaticae 39(2012)193–206

## Egy véletlen algoritmus

Szerzőtársaimmal egy optimumkeresési feladatot oldottunk meg  
Robbins-Monro típusú algoritmussal.

M. Barczy, Á. Nagy, Cs. Noszály, Cs. Vincze  
A stochastic algorithm for computing global minimizer of  
generalized conic functions  
benyújtva (Journal of Approximation Theory)

# Futamok

A Bernoulli kísérletsorozat egyszerűsége és alkalmazhatósága folytán fontos szerepet játszik a valószínűségszámításban a kezdetektől napjainkig. Az egyik manapság is aktív terület az ún. futamok vizsgálata.

## Futamok

Legyen  $X_1, \dots, X_n$  független, azonos eloszlású Bernoulli valószínűségi változó sorozat és  $p = P(X_1 = \text{fej})$ ,  $q = 1 - p$ . Ha megfigyelünk egy ilyen sorozatot, azaz vesszük a kapott *fej-írás* sorozatot, láthatunk benne egyforma elemekből álló összefüggő részsorozatokat: ezeket nevezzük *futamoknak*, melyek közül a maximális hosszúságúak, tovább nem bővíthetők az érdekesek.

## Futamok

Legegyszerűbb a csupa *fejekből* álló maximális futamokat vizsgálni.

Ekkor az  $L_n$  valségi változót figyeljük meg, ahol

$$L_n(\omega) = k \iff$$

$X_1(\omega), \dots, X_n(\omega)$ -ban  $k$ -hosszú a maximális *fej*-futam.

Ennek a valségi változónak a viselkedése: várható értéke, szórása, eloszlása érdekel minket.

# Futamok

consecutive  $k$  out of  $n$  systems

Példa futamok megjelenésére: "consecutive  $k$  out of  $n$  system":  
adott  $n$  komponens sorbarendezve. Ha tudjuk hogy a komponensek egymástól függetlenül  $p$  valószínűséggel romlanak el, illetve a teljes rendszer pont akkor romlik el, ha legalább  $k$  szomszédos komponens elromlik. A futamok eloszlásának ismeretében meg tudjuk mondani hogy milyen eséllyel romlik el a rendszer.

# Futamok

egy intuitív teszt

Legyen  $X_i$  és  $Y_i$   $p$ -paraméterű Bernoulli és  $Z_i = 1$  ha  $X_i = Y_i$  és  $Z_i = 0$  egyébként. Ha  $X_i$  és  $Y_i$  független, akkor  $Z_i$  egy  $p^2 + q^2$  paraméterű Bernoulli. Ekkor  $L_n$  és a  $Z_i$  megfigyelt futamjai segítségével dönteni tudunk arról, hogy kiinduló feltevésünk a függetlenségről, a kapcsolat hiányáról mennyire felel meg a valóságnak.

# Futamok

aszimptotikus eredmények

Ha  $p = \frac{1}{2}$  akkor  $\lim_{n \rightarrow \infty} \frac{L_n}{\log_2 n} = 1$  majdnem biztosan.

P.Erdős-A.Rényi:

On a new law of large numbers.

Journ. Analyse Math. 22(1970) 103-111

# Futamok

aszimptotikus eredmények

$$\mathbb{E}(L_n) = \log_{\frac{1}{p}}(nq) + \frac{\gamma}{\log(\frac{1}{p})} - \frac{1}{2} + r_1(n) + \epsilon_1(n).$$

$\gamma$  az Euler konstans,  $|r_1(n)|$  kicsi és  $\epsilon_1 \rightarrow 0$ .

# Futamok

aszimptotikus eredmények

$$\mathbb{D}^2(L_n) = \left( \frac{\pi}{\sqrt{6} \log(\frac{1}{\rho})} \right)^2 + \frac{1}{12} + r_2(n) + \epsilon_2(n)$$

$|r_2(n)|$  kicsi és  $\epsilon_2 \rightarrow 0$ .

# Futamok

aszimptotikus eredmények

Ha  $\mu(x) = -\frac{\log x}{\log p}$  és  $P(W \leq t) = e^{-e^{-t}}$  akkor  $t$ -ben egyenletesen:

$$P(L_n - \mu(qn) \leq t) - P\left(\left[-\frac{W}{\log p} + \{\mu(qn)\}\right] - \{\mu(qn)\} \leq t\right) \rightarrow 0.$$

L.Gordon, M.F.Schilling, M.S.Waterman:

An extreme value theory for long head runs.

Probab. Theory. and Rel. Fields 72(1986) 279-287

# Futamok

well-matching

Ezen eredmények ún. szennyezett futamokra történő általánosításai is ismertek.  $T$ -szennyezett futam: olyan összefüggő részsorozat mely legfeljebb  $T$  írást is tartalmazhat. A fent említett intuitív tesztben is életszerűbb a használatuk.

# Futamok

well-matching

Az aszimptotikus eredmények a gyakorlatban nehezen alkalmazhatók, pontatlanok. Szükség van pontos eloszlásbeli eredményekre. A leghosszabb  $T$ -szennyezett futam eloszlásának pontos számolására dolgoztak ki egy Markov-láncokon alapuló algoritmust az alábbi cikkben:

# Futamok

well-matching

C.J.Chang,C.S.J. Fang,W.C. Chou,I.B. Lian  
On the tail probability of the longest well matching run.  
Stat. and Prob. Letters,63(2003) 267-274

# Futamok

well-matching

A leghosszabb  $T$ -szennyezett futam eloszlása "rekurzív" módon is kiszámolható. Például, ha  $k > 0$  és  $L_n$  a leghosszabb 1-szennyezett futam hossza és  $R(n) = P(L_n \leq k)$  és

$$R(n, v) = P(\{L_n \leq k\} \{X_{n-(v-1)} = \dots = X_n = fej\}),$$

akkor a megfelelő kezdeti feltételeket megadva:

# Futamok

well-matching

$$R(n, v) = \sum_{w+v+1 \leq k} R(n-k, w)$$

segítségével számolható az eloszlás. Hasonló, de egyre

bonyolultabb formulák adódnak  $T > 1$ -re.

# Futamok

well-matching

Jelenleg a fenti cikkbeli és ezen rekurzióval számolt eredmények, módszerek összehasonlítását végzem. A  $T = 1, 2$  esetet megvalósítottam, a  $T = 1$  esetben  $n < 100000$ -ra -  $T = 2$  esetben  $n < 10000$ -re 5 – 10 másodperces futási idővel asztali gépen. A  $T = 3$  eset folyamatban.

# Futamok

simple-bound

Érdekes terület egyszerű alsó és felső becslések keresése

$R(n) = R(n, k) = P(L_n < k)$ -ra. Fu és Muselli a következőket

látták be:

$$(1 - p^k)^{n-k+1} \leq R(n) \leq (1 - qp^k)^{n-k+1}$$

$$(1 - p^k)^{n-k+1} \leq R(n) \leq (1 - p^k)^{\lfloor \frac{n}{k} \rfloor}$$

# Futamok

simple-bound

Fu, J.C.

Reliability of a large consecutive-k-out-of-n:F system  
IEEE Trans. Reliability R-34, 127–130 (1985)

M. Muselli

Useful inequalities for the longest run distribution  
Statistics and Probability Letters 46 (2000) 239–249

# Futamok

simple-bound

A fenti egyenlőtlenségek  $T$ -szennyezett sorozatra történő általánosításán, illetve a leghosszabb *fej* vagy *írás* futamot tartalmazó változatán dolgozom.

# Példa

Muselli

Legyen  $n, k > 0$ . Ekkor

1.  $R(n) = 1$ , ha  $n < k$ ;
2.  $R(k) = 1 - p^k$ ;
3.  $R(k + 1) = 1 - (2qp^k + p^{k+1})$ ;
4.  $R(n + 1) = R(n) - R(n - k)qp^k$ , ha  $n > k$ ;
5.  $R$  nemnövekvő ( $n$ -ben).

# Példa

Muselli

1. Nincs rossz sorozat.
2. Csak a csupa *fej* a rossz.
3. A csupa *fej* és a valamelyik végén *írást* tartalmazó egyébként csupa *fej* a rossz.

# Példa

Muselli

4. Ha egy  $n$  hosszú jó sorozat után *írást* rakunk akkor az jó, ha *fejet* akkor, gond lehet hogy létrejön egy  $k$ -hosszú tiszta *fejsorozat* a végén: ennek a valószínűsége  $R(n - k)qp^k$ , amit le is vontunk.
5. Lásd az előző pontot. Vagy: minél hosszabb a sorozat, annál kevésbé esélyes, hogy  $k$  alatt marad a leghosszabb tiszta *fejsorozat* hossza.

# Példa

Muselli

1.  $R(n) \leq (1 - p^k)R(n - k)$ , ha  $0 < k < n$ .
2.  $(1 - p^k)R(n - 1) \leq R(n)$ , ha  $0 < n$ .

# Példa

Muselli

1. Egy  $n$ -hosszú jó sorozatot felbontva egy  $k$  és egy  $n - k$ -hosszúra, mindkettő jó-sorozat és a jobboldali szorzatban minden  $n$ -hosszú jó sorozatot figyelembe vettünk.
2.  $n = 1, \dots, k$ -ra világos az állítás az eddigiekből.  $n = k + 1$ -re:

$$\begin{aligned}(1 - p^k)^2 &\leq 1 - (2qp^k + p^{k+1}) \\ 1 - 2p^k + p^{2k} &\leq 1 - (2qp^k + p^{k+1}) \\ p^{2k} &\leq p^{k+1}\end{aligned}$$

# Példa

Muselli

Mivel  $n > k + 1$  esetén:

$$R(n+1) = R(n) - R(n-k)qp^k,$$
$$\frac{R(n+1)}{R(n)} = 1 - \left(\frac{R(n-k)}{R(n)}q\right)p^k,$$

elég azt látni, hogy

$$R(n-k)q \leq R(n),$$

ami azért igaz mert az  $n$ -hosszú jó sorozatoknál minden olyan sorozatot is figyelembe vettünk amelyeknek az eleje egy  $n - k$ -hosszú jó sorozat és *írással* folytatódik.

# Példa

Muselli

Az előző egyenlőtlenségeket alkalmazva kapjuk, hogy:

$$(1 - p^k)^{n-k+1} \leq R(n) \leq (1 - p^k)^{\lfloor \frac{n}{k} \rfloor}$$

# Viszlát