UNIVERZITET U NOVOM SADU

PRIRODNO-MATEMATIČKI FAKULTET

DEPARTMAN ZA MATEMATIKU

I INFORMATIKU

Granične raspodele: zakoni velikih i malih brojeva

- master rad -

Mentor:

doc. dr Dora Seleši

Autor:

Bojana Jovanić

Novi Sad, januar 2011.

Sadržaj:

Predgovor

U ovom radu ćemo izučavati jedan od segmenata aktuarske matematike, zakon velikih i malih brojeva, kao i granične raspodele i njihove osobine.

U poglavlju 1. ćemo razmatrati osnovnu teoriju u vezi sa sumama nezavisnih slučajnih promenljivih, što uključuje jak zakon velikih brojeva u 1.1 i centralnu graničnu teoremu u 1.2. U 1.3 su date modifikacije CGT (asimptotska proširenja, velike devijacije i red konvergencije), dok su u 1.4 dati Brownovo i α-stabilno kretanje kao slabe granice procesa parcijalnih suma.
Javlja se i homogeni Poissonov proces kao specijalan proces prebrajanja u 1.5.2.
U poglavlju 2. ćemo se baviti klasičnom teorijom ekstremnih vrednosti. Glavni rezultat jeste Fisher-Tippetova teorema koja određuje oblik granične raspodele za centrirane i normalizovane maksimume. Takođe ćemo razmatrati Poissonovu aproksimaciju kao alat za proučavanje retkih događaja. Skoro sigurno ponašanje maksimuma razmatramo u poglavlju 2.5.
U poglavlju 3. ćemo razmatrati ponašanje nekoliko statistika višeg reda koje daju informacije o desnom repu funkcije raspodele.
U poglavlju 4. ćemo proučavati najduži niz uspeha jedinica, dužinu tog niza, kao i njegovo skoro sigurno ponašanje.

Ovom prilikom se zahvaljujem doc. dr Sanji Rapajić i doc. dr Danijeli Rajter-Ćirić na korisnim sugestijama i primedbama.

Posebno se zahvaljujem svom mentoru, doc. dr Dori Seleši, za stručno i profesionalno usmeravanje tokom izrade ovog rada, kao i za svo znanje koje sam stekla radeći sa njom.

Takođe moram da pomenem svoje roitelje, prijatelje i kolege i da im se zahvalim što me svakodnevno i na razne načine podsećaju da je od svih dobrih stvari znanje najbolje: zato što ga niko ne može ukrasti, zato što ga niko ne može kupiti, zato što je neuništivo.

Novi Sad, januar 2011. Bojana Jovanić

Uvod

Osiguranje predstavlja jedan oblik upravljanja rizikom čija je osnovna ideja povezivanje rizika velikog broja sličnih agenata u jedan fond, tako da zakon verovatnoće/zakon velikih brojeva obezbedi da se samo relativno mali broj nepovoljnih događaja ostvari u jednoj godini.

Prvi oblici osiguranja sreću se u prvobitnoj ljudskoj zajednici u okviru plemena, a kasnije i porodica. Prva opasnost sa kojom se sretao čovek je glad, a mera kojom je pokušao da se zaštiti bili su skromni obavezni prilozi pojedinca u žitu tokom rodnih godina.

Elementi osiguranja javljaju se u Vavilonu još pre četiri milenijuma: u slučaju gubitka broad, vlasniku se nadoknađivala šteta, a ako brod stigne na destinaciju, vlasnik je bio dužan da isplati određeni deo dobiti. Pisani tragovi o osiguranju postoje i u Hamurabijevom zakonu iz 2250. godine p.n.e. u vidu uredbe o međusobnoj obavezi učesnika trgovačkog karavana da nadoknade štetu koja bi nastala u slučaju pljačke.

Prva sačuvana polisa potiče iz Lombardije iz 1182. godine. Iz perioda od XII do XV veka sačuvano je više od 400 polisa, ali u to vreme polisa osiguranja nije uvek bila garancija dobijanja nadoknade, naročito u slučaju gubitka broda. Prvi zakoni u ovoj oblasti donose se u Barseloni 1435. i Firenci 1522. godine.

U Londonu su se vlasnici brodova, prodavci i kupci robe koja se prevozila brodovima sastajali u kafanama, a jedna od kafana u vlasništvu Edwarda Lloyda, postala je vodeća. Prva osiguravajuća kompanija u oblasti pomorskog osiguranja i nosi naziv po Edwardu Lloydu. Ne zna se tačno kada je Lloyd`s osnovan, ali je poznato da je postojao pre 1688 godine. Aktom Britanskog parlamenta iz 1871. udruženje osiguravača "Lloyds" dobilo je zvanični status korporacije osiguravača. Na taj način službeno je osnovana Lloydova berza. "Lloyds" danas nije kompanija, već predstavlja tržište osiguranja.

U XVII veku holandski državnik i matematičar Jan de Vitu postavio je matematičke osnove odredivanja životne rente. Dostignuća Newtona, Leibnitza i Pascala našla su veliku primenu u oblasti osiguranja. Od velikog značaja bila su i otkrića u matematičkoj statistici, pre svega zakon velikih brojeva (Bernoulli, Laplas i Gauss). Engleska akademija nauka je krajem XVIII veka stvorila pretpostavke za razvoj modemog osiguranja.

Zakon velikih brojeva je matematička premisa koja tvrdi da je predviđanje tačnije što je veća izloženost riziku, manje je odstupanje stvarnih gubitaka u odnosu na očekivane gubitke i veća je verodostojnost predviđanja. Ovaj zakon predstavlja osnovu za statističko očekivanje gubitka na osnovu kojeg se računaju premije za polise osiguranja. Osiguravajuća kompanija može da, od velike grupe osiguranika, prilično tačno predvidi broj osiguranika koji će pretrpeti gubitak.

Premije životnog osiguranja uključuju očekivane gubitke, kao i mala odstupanja. Na primer, ako osiguravajuća kompanija očekuje da u nekoj godini umre određeni broj osiguranika i umre ih tačno toliko, ili manje, ne postoji razlog za zabrinost aktuara. Međutim, ukoliko se desi suprotno, tj. umre veći broj osiguranika od očekivanog, situacija se drastično menja.

Osnovni pojmovi i teoreme

$\mathcal{R}_\alpha$	skup regularno varirajućih funkcija sa indeksom α
$\mathcal{R}_{-\infty}$	skup brzo varirajućih funkcija
$\overline{F}$	rep funcije raspodele F: $\overline{F} = 1 - F$
$F^{\leftarrow}(p)$	p-kvantil od F
G_α	α-stabilna raspodela
Γ	gama funkcija $\Gamma(\mathrm{x}) = \int_0^\infty \mathrm{t}^{\mathrm{x}-1}\mathrm{e}^{-\mathrm{t}}\mathrm{dt}$
H	raspodela ekstremnih vrednosti
iid promenljive	nezavisne promenljive sa identičnom raspodelom
s.s.	skoro sigurno
i.o.	beskonačno često
$s - \alpha - s$	simetrično α-stabilno kretanje
$\stackrel{P}{\rightarrow}$	teži u verovatnoći
$\stackrel{d}{\rightarrow}$	teži u raspodeli
$\stackrel{d}{=}$	jednako u raspodeli
$\asymp$	$a(x) \asymp b(x)$ kad $x \to x_0$ znači da je $0 < liminf_{x\to x_0} a(x)/b(x) \le limsup_{x\to x_0} a(x)/b(x) < \infty$
$\mathbb{C}$	prostor neprekidnih funkcija
$\mathbb{D}$	prostor cadlag funkcija

Teorema D1 (Konvergencija ka tipskim promenljivama)
Neka su $A, B, A_1, A_2, \ldots$ slučajne promenljive i $b_n > 0, \beta_n > 0$ i $a_n, \alpha_n \in \mathbb{R}$ konstante. Pretpostavimo da je

$$b_n^{-1}(A_n - a_n) \stackrel{d}{\rightarrow} A.$$

Tada relacija

$$\beta_n^{-1}(A_n - \alpha_n) \stackrel{d}{\rightarrow} B \quad \text{(D.1)}$$

važi ako i samo ako

$$\lim_{n\to\infty} b_n/\beta_n = b \in [0, \infty)\,, \ \lim_{n\to\infty} (a_n - \alpha_n)/\beta_n = a \in \mathbb{R}. \quad \text{(D.2)}$$

Ako (D.1) važi, onda je $B \stackrel{d}{=} bA + a$ i a, b su jedinstvene konstante za koje ovo važi.
Ako važi (D.1), A je nedegenerativno ako i samo ako je $b > 0$ i tada A i B pripadaju istom tipu, tj. važi $B \stackrel{d}{=} bA + a$. □

Tvrđenje D2 (Regularna varijacija za repove funkcija raspodele)
Pretpostavimo da je F funkcija raspodele takva da je $F(x) < 1$ za svako $x \geq 0$.

(a) Ako nizovi (a_n) i (x_n) zadovoljavaju $a_n/a_{n+1} \to 1, x_n \to \infty$ i ako za neku realnu funkciju g i svako λ iz gustog podskupa $(0,\infty)$, važi
$\lim_{n\to\infty} a_n \bar{F}(\lambda x_n) = g(\lambda) \in (0,\infty)$,
tada $g(\lambda) = \lambda^{-\alpha}$ za neko $\alpha \geq 0$ i $\bar{F}$ je regularno varirajuće.

(b) Pretpostavimo da je F apsolutno neprekidno sa gustinom f takvom da za neko $\alpha > 0, \lim_{x\to\infty} xf(x)/\bar{F}(x) = \alpha$. Tada $f \in \mathcal{R}_{-1-\alpha}$, pa i $\bar{F} \in \mathcal{R}_{-\alpha}$.

(c) Pretpostavimo da $f \in \mathcal{R}_{-1-\alpha}$ za neko $\alpha > 0$. Tada $\lim_{x\to\infty} xf(x)/\bar{F}(x) = \alpha$, što važi i ako $\bar{F} \in \mathcal{R}_{-\alpha}$ za neko $\alpha > 0$ i gustina f je monotona na (z,∞) za neko $z\epsilon\mathbb{R}$.

(d) Pretpostavimo da je X nenegativna slučajna promenljiva sa raspodelom repa $\bar{F} \in \mathcal{R}_{-\alpha}$ za neko $\alpha > 0$. Tada
$E(X^\beta) < \infty$ ako $\beta < \alpha$,
$E(X^\beta) = \infty$ ako $\beta > \alpha$.

(e) Pretpostavimo da je $\bar{F} \in \mathcal{R}_{-\alpha}$ za neko $\alpha > 0, \beta \geq \alpha$. Tada
$\lim_{x\to\infty} \frac{x^\beta \bar{F}(x)}{\int_0^x y^\beta dF(y)} = \frac{\beta-\alpha}{\alpha}$.

Obrnuto takođe važi u slučaju da je $\beta > \alpha$. Ako je $\beta = \alpha$, može se zaključiti da je $\bar{F}(x) = o\big(x^{-\alpha}L(x)\big)$ za neko $L \in \mathcal{R}_0$.

(f) Sledeće dve relacije su ekvivalentne:
(1) $\int_0^x y^2\, dF(y) \in \mathcal{R}_0$
(2) $\bar{F}(x) = o\big(x^{-2}\int_0^x y^2 dF(y)\big),\ \ x \to \infty$. □

Teorema D3 (Teorema o neprekidnom preslikavanju)
Neka je h preslikavanje iz metričkog prostora K u metrički prostor K' (oba prostora su snabdevena Borelovom σ-algebrom generisanom otvorenim skupovima). Neka je A_n niz slučajnih promenljivih sa vrednostima u prostoru K. Pretpostavimo da $A_n \xrightarrow{d} A$ u K i sa P_A označimo raspodelu od A. Tada $h(A_n) \xrightarrow{d} h(A)$ u prostoru K', pod pretpostavkom da skup tačaka prekida funkcije h ima P_A-meru nula. □

Teorema D4 (Osobine funkcija brze varijacije)

(a) Pretpostavimo da $h \in \mathcal{R}_{-\infty}$ nije rastuća funkcija. Tada za neko $z > 0$ i za svako $\alpha \in \mathbb{R}$

$$\int_z^\infty t^\alpha h(t)dt < \infty$$

i

$$\lim_{x\to\infty} \frac{x^{\alpha+1}h(x)}{\int_x^\infty t^\alpha h(t)dt} = \infty. \qquad \text{(D.3)}$$

Ako za neko $\alpha \in \mathbb{R}$, $\int_0^\infty t^\alpha h(t)\, dt < \infty$ i (D.3) važi, tada $h \in \mathcal{R}_{-\infty}$.

(b) Ako $h \in \mathcal{R}_{-\infty}$, tada postoje funkcije c i δ tako da $c(x) \to c_0 \in (0, \infty), \delta(x) \to -\infty$ kad $x \to \infty$ i za neko $z > 0$,
$h(x) = c(x) exp\left\{\int_z^x \frac{\delta(u)}{u} du\right\}, \quad x \geq z.$
Obrnuto takođe važi. □

Definicija D5 (Regularno varirajući nizovi)
Niz pozitivnih brojeva (c_n) je regularno varirajući sa indeksom $\alpha \in \mathbb{R}$ ako

$$\lim_{n\to\infty} \frac{c_{[tn]}}{c_n} = t^\alpha, t > 0$$ □

Definicija D6 (Mrežasta funkcija)
Funkcija raspodele F (za X) je mrežasta ako $\exists a, d \geq 0$ tako da
$\sum_{n=1}^\infty P\{X = a + nd\} = 1.$ □

Definicija D7 (Cramer-Lundbergov model, model obnavljanja)

Cramer-Lundbergov model je definisan uslovima (a)-(e):

(a) Proces iznosa zahteva:
Iznosi zahteva $(X_k)_{k\in N}$ su pozitivne iid slučajne promenljive koje imaju zajedničku ne-mrežastu funkciju raspodele F, konačno očekivanje $\mu = E(X_1)$, i varijansu $\sigma^2 = Var(X_1) \leq \infty$.

(b) Vremena zahteva:
zahtevi se dešavaju u slučajnim vremenskim trenucima
$0 < T_1 < T_2 < \cdots$ s.s.

(c) Proces dolaska zahteva:
Broj zahteva u intervalu $[0, t]$označavamo sa
$N(t) = sup\{n \geq 1: T_n \leq t\},\ t \geq 0,$
gde je $sup\emptyset = 0$.

(d) Vremena između dolazaka

$$Y_1 = T_1, Y_k = T_k - T_{k-1},\ k = 2,3, \ldots \tag{D.4}$$

su iid sa eksponencijalnom raspodelom sa konačnim očekivanjem $E(Y_1) = 1/\lambda$.

(e) Nizovi (X_k) i (Y_k) su međusobno nezavisni.
Model obnavljanja je dat sa (a)-(c), (e) i

(d')Vremena između dolazaka Y_k data u (D.3) su iid sa konačnim očekivanjem $E(Y_1) = 1/\lambda$. □

Tvrđenje D8 (Konvergencija uopštene inverzne funkcije)
Neka su $h, h_1, h_2, \ldots$ neopadajuće funkcije takve da $\lim_{n\to\infty} h_n(x) = h(x)$ za svaku tačku x u kojoj je funkcija h neprekidna. Tada je $\lim_{n\to\infty} h_n^{\leftarrow}(y) = h^{\leftarrow}(y)$ za svaku tačku y u kojoj je funkcija $h^{\leftarrow}$ neprekidna. □

Definicija D9 (Nejednakost Chebysheva)
$P(|X-\mu| \geq t) \leq \frac{\delta^2}{t^2},\ t > 0$
$E(X) = \mu$
$Var(X) = \delta^2.$ □

Definicija D10 (Hermiteov polinom)
Probabilistički Hermiteov polinom je definisan kao
$H_n(x) = (-1)^n e^{x^2/2} \frac{d^n}{dx^n} e^{-x^2/2}.$ □

Lema D11 (Lema o uklještenim nizovima)
Ako su definisana tri niza A_n, B_n, C_n za koje važi

$\lim_{x\to\infty} A_n = \lim_{x\to\infty} B_n = C.$

Sledi da je i
$\lim_{x\to\infty} C_n = C$. □

Definicija D12 (Cesaro sredina)
Cesaro sredina niza (a_n) su članovi niza c_n, gde je

$$c_n = \frac{1}{n}\sum_{i=1}^{n} a_i$$

aritmetička sredina prvih n elemenata niza (a_n).
Sledi da je $\lim_{n\to\infty} a_n = A \Longrightarrow \lim_{n\to\infty} c_n = A.$ □

Definicija D13 (Nepotpuna beta funkcija)
Nepotpuna beta funkcija je definisana kao

$$I_x(a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^x z^{a-1}(1-z)^{b-1} dz$$

□

Definicija D14 (Stein-Chenov metod)
Ovaj metod daje granicu za totalnu udaljenost varijacije između raspodele za sumu ne-iid Bernoullijeve promenljive Y_i , $i \in 1$ i raspodele za Poissonovu promenljivu sa parametrom $\lambda = \sum_{i\in I} E(Y_i)$.
Neka je $N(W) = \sum_{i=1}^{n-h+1} Y_i$ i neka je $B_i \subset I$ okolina od i.
Chen-Steinova teorema kaže da je
$d_{TV}(\alpha\big(N(W)\big), \mathcal{P}(\lambda)) \leq 2(b_1 + b_2 + b_3)$

$$b_1 = \sum_{i\in I}\sum_{j\in B_i} E(Y_i)E(Y_j)$$

$$b_2 = \sum_{i\in I}\sum_{j\in B_i\setminus\{i\}} E\big(Y_iY_j\big)$$

$b_3 = \sum_{i \in I} E\left|E(Y_i - E(Y_i)\middle|\sigma(Y_j, j \in B_i^c))\right|$. □

Definicija D15 (Tejlorova ekspanzija za funkciju $f(x)$ i datu tačku a)

$$T_n(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n$$
$$= \sum_{k=0}^{n}\left(\frac{f^k(a)}{k!}(x-a)^k\right)$$ □

Definicija D16 (Ergodičan niz)

Neka je $A = \{a_j\}$ beskonačan strogo rastući niz pozitivnih celih brojeva. Tada se, za dat ceo broj q ovaj niz naziva ergodičnim sa modom q ako za svaki ceo broj $1 \leq k \leq q$, $\lim_{t\to\infty} \frac{N(A,t,k,q)}{N(A,t)} = \frac{1}{q}$, gde je $N(A,t) = card\{a_j \in A: a_j \leq t\}$ i $card$ je broj elemenata skupa, tako da je $N(A,t)$ broj elemenata u nizu A koji su manji ili jednaki t i $N(A,t,k,q) = card\{a_j \in A: a_j \leq t, a_j \ mod\ q = k\}$, pa $N(A,t,k,q)$ broj elemenata u A manji od t koji su ekvivalentni $k \ mod\ q$, tj. niz je ergodičan ako ima uniformnu raspodelu $mod\ q$ kad niz ide u ∞. □

Definicija D17 (Cadlag funkcija)

Cadlag funkcija je funkcija koja je neprekidna zdesna i postoje (konačni) limesi sa leve strane u svakoj tački. □

Teorema D18 (Centralna granična teorema)

Neka je (X_n) iid niz sa matematičkim očekivanjem$E(X_1) = m$ i konačnom disperzijom $D(X_1) = \sigma^2 > 0$. Ako je $S_n = \sum_{k=1}^{n} X_k$, onda za svako $x \in \mathbb{R}$, kad $n \to \infty$ važi

$$P\left\{\frac{S_n - nm}{\sigma\sqrt{n}} \leq x\right\} \to \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-t^2/2}dt.$$

Ovo tvrđenje može biti formulisano i kao:

$$\frac{S_n - E(S_n)}{\sqrt{D(S_n)}} \xrightarrow{D} N(0,1).$$

□

Tvrđenje D19 (Hinčinov zakon iteriranog logaritma)

Neka je Y_n iid slučajna promenljiva sa očekivanjem 0 i disperzijom 1. Neka je

$$S_n = Y_1 + \cdots + Y_n.$$

Tada je

$$\limsup_{n\to\infty} \frac{S_n}{\sqrt{n\, ln \ln n}} = \sqrt{2}, \qquad \text{s.s.}$$

□

Definicija D20 (Regularna varijacija sa indeksom ∞ / brza varijacija)

Kažemo da je funkcija $U: \mathbb{R}_+ \to \mathbb{R}_+$ regularno varirajuća sa indeksom ∞ ako za svako $x > 0$ važi

$$\lim_{t\to\infty} \frac{U(tx)}{U(t)} = x^{\infty} := \begin{cases} 0, & x < 1, \\ 1, & x = 1, \\ \infty, & x > 1. \end{cases}$$

Slično, funkcija $U: \mathbb{R}_+ \to \mathbb{R}_+$ je regularno varirajuća sa indeksom $-\infty$ ako

$$\lim_{t\to\infty} \frac{U(tx)}{U(t)} = x^{-\infty} := \begin{cases} \infty, & x < 1, \\ 1, & x = 1, \\ 0, & x > 1. \end{cases}$$

□

Definicija D21 (Metrizabilan metrički prostor)
Metrizabilan metrički proctor je prostor čija je topologija generisana nekom metrikom putem pravila: tačka pripada zatvaranju skupa ako i samo ako je njena udaljenost od skupa jednaka nuli. □

Definicija D22 (Nepotpuna gama funkcija)
Nepotpuna gama funkcija ima oblik

$$\Gamma(a, x) = \int_x^\infty t^{a-1} e^{-t} dt,$$

gde je a kompleksan broj čiji je realan deo veći od nule. □

Definicija D23 (Stirlingova formula)

$$N! \approx \left(\frac{N}{e}\right)^N \sqrt{2\pi N} = S_t(N).$$

□

1. Fluktuacija suma

1.1 Zakon velikih brojeva

Zakon velikih brojeva je fundamentalna teorema iz oblasti teorije verovatnoće i statistike.
U svom najjednostavnijem obliku ovaj zakon tvrdi da se relativna verovatnoća slučajnog događaja približava verovatnoći ovog događaja kada se slučajni eksperiment ponavlja veliki broj puta. Formalnije, radi se o konvergenciji slučajne promenljive u „jakom“ (skoro sigurna konvergencija) i u „slabom“ smislu (konvergencija verovatnoće).
Verovatnoća da bačeni novčić pokaže pismo ili glavu iznosi $1/2$. Što se više ponavlja ovaj eksperiment, to je verovatnije da će broj ishoda kada „padne glava“ (relativna verovatnoća ishoda „glava“), biti blizak vrednosti $1/2$. S druge strane, vrlo je verovatno da će apsolutna razlika između broja ishoda „glava“ i polovine broja bacanja novčića rasti.

Postoji i ukorenjeno pogrešno shvatanje zakona velikih brojeva. Ovaj zakon ne tvrdi da će oni ishodi koji se do sada nisu pojavljivali, od sada pojavljivati češće da bi uravnotežili raspodelu verovatnoća.

Neka je niz ishoda bacanja novčića: glava, pismo, glava, glava. Ishod „glava“ se pojavio tri puta, a ishod „pismo“ jedanput. „Glava“ se dakle pojavljivala sa relativnim učešćem $3/4$, dok je ova vrednost za „pismo“ $1/4$. Posle novih 96 bacanja novčića ishod je bio 49 „pisama“ i 51 „glava“. Apsolutna razlika glava i pisama je posle 100 bacanja ostala ista i kao posle 4, ali je relativna razlika znatno smanjena. Tako dolazimo do definicije Zakona velikih brojeva – vrednost relativnog učešća $51/100 = 0.51$ teži očekivanoj vrednosti 0.5.

Zakon velikih brojeva ima veliki praktični značaj za industriju osiguranja. On omogućava da se naprave dugoročne prognoze iznosa odštetnih zahteva. Što je veći broj osiguranih osoba i dobara, i ako se podrazumeva da su svi izloženi jednakom riziku, to je manji uticaj slučaja. Ovaj zakon ipak, ne može da da nikakvu prognozu o tome ko će konkretno uložiti odštetni zahtev.

U ovom poglavlju, $X_1, X_2, \ldots$ predstavlja niz iid nedegenerativnih slučajnih promenljivih, definisan na prostoru verovatnoće $[\Omega, \mathcal{F}, P]$ sa zajedničkom funkcijom raspodele F.
Ukoliko želimo da steknemo grubu ideju o fluktuaciji X_n, možemo da zahtevamo konvergenciju niza (X_n).
Nažalost, za skoro sve $\omega \in \Omega$, ovaj niz ne konvergira. Međutim, možemo da dobijemo neke informacije o tome kako se ponaša „sredina“ od X_n. Zato razmatramo kumulativne sume

$$S_0 = 0,\, S_n = X_1 + \cdots + X_n,\, n \geq 1$$

i aritmetičke (ili uzoračke) sredine

$$\bar{X}_n = n^{-1}S_n,\ n \geq 1.$$

Intuitivno, jasno je da bi aritmetička sredina trebalo da ima neku vrstu stabilnosti za dovoljno veliko n, pa očekujemo da će za veliko n individualne vrednosti X_i imati manje uticaja na red $\bar{X}_n$, tj. niz $(\bar{X}_n)$ se stabilizuje u okolini fiksirane vrednosti (konvergira) kada $n \to \infty$. Ovo nazivamo zakonom velikih brojeva.

Prva granična teorema koja je dokazana u teoriji verovatnoća je Bernoullijev zakon velikih brojeva. Ovde ćemo dati i neka uopštenja te teoreme.

1) **Bernoullijev zakon velikih brojeva**
   Neka je (X_n) niz nezavisnih slučajnih promenljivih koje imaju raspodelu
   $X_n\colon \begin{pmatrix} 0 & 1 \\ 1-p & p \end{pmatrix}$, $n = 1,2,\dots$ i $E(S_n) = np$.
   Tada važi nejednakost
   $P\left\{\left|\frac{S_n}{n} - p\right| \geq \varepsilon\right\} \leq \frac{p(1-p)}{n\varepsilon^2}$.

2) **Chebyshev zakon velikih brojeva**
   Neka je (X_n) niz nezavisnih slučajnih promenljivih i $C > 0$ konstanta takva da za svaki prirodan broj n važi $Var(X_n) \leq C$.
   Ako je $S_n = \sum_{k=1}^{n} X_k$, onda važi $\frac{S_n - E(S_n)}{n} \xrightarrow{P} 0$, kad $n \to \infty$.

3) **Hinchinov zakon velikih brojeva**
   Neka je (X_n) niz iid slučajnih promenljivih sa konačnim matematičkim očekivanjem m. Tada $\frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow{P} m$, kad $n \to \infty$.

Pretpostavimo sada da je $\sigma^2 = Var\,(X)$ konačno i pišemo $\mu = E(X)$. Iz nejednakosti Chebysheva zaključujemo da za $\varepsilon > 0$,

$$P(|\bar{X}_n - \mu| > \varepsilon) \leq \varepsilon^{-2} Var(\bar{X}_n) = (n\varepsilon^2)^{-1}\sigma^2 \to 0,\ n \to \infty,$$

pa

$$\bar{X}_n \xrightarrow{P} \mu,\ n \to \infty.$$

Ova relacija se naziva slabim zakonom velikih brojeva, ili jednostavnije, zakonom velikih brojeva za niz (X_n). Ako tumačimo indeks od X_n kao vreme n, onda je $\bar{X}_n$ prosek po vremenu. S druge strane, očekivanje

$$E(X) = \int_{\Omega} X(\omega)dP(\omega)$$

je težinski prosek na prostoru verovatnoće Ω. Prema tome, zakon velikih brojeva nam govori da, kroz dug period, prosek po vremenu $\bar{X}_n$ konvergira ka $E(X)$.
Videli smo da (X_n) zadovoljava slab zakon velikih brojeva ako je disperzija σ^2 konačna. Ovaj uslov se može značajno oslabiti:

Teorema 1.1.1 (Kriterijum za slab zakon velikih brojeva)
Slab zakon velikih brojeva $\bar{X}_n \xrightarrow{P} 0$ važi ako i samo ako su zadovoljena sledeća dva uslova

$$nP(|X| > n) \to 0,$$
$$E(X)I_{\{|X|\leq n\}} \to 0.$$

□

Nadalje koristimo indikator promenljivu događaja na skupu A

$$I_A(\omega) = \begin{cases} 1, & \omega\epsilon A \\ 0, & inače. \end{cases}$$

Sada se postavlja pitanje koji su uslovi neophodni da bi se osiguralo da $\bar{X}_n$ ne konvergira samo u verovatnoći, već sa verovatnoćom 1, tj. skoro sigurno. Takav rezultat se onda naziva jakim zakonom velikih brojeva (JZVB) za niz (X_n). Postojanje prvog momenta je neophodan uslov za JZVB: ako $\bar{X}_n \xrightarrow{s.s.} a$ za neku konačnu konstantu a, važi da je

$$n^{-1}X_n = n^{-1}(S_n - S_{n-1}) \xrightarrow{s.s.} a - a = 0,$$

pa za $\varepsilon > 0$,

$$P(n^{-1}|X_n| > \varepsilon \quad i.o.) = 0.$$

Ovo, zajedno sa Borel-Cantellijevom lemom (2.5), implicira da za $\varepsilon > 0$,

$$\sum_{n=1}^{\infty} P(|n^{-1}X_n| > \varepsilon) = \sum_{n=1}^{\infty} P(|X| > \varepsilon n) < \infty,$$

što znači da je $E|X| < \infty$. Sledeći uslov je takođe dovoljan za JZVB:

Teorema 1.1.2 (Kolmogorov jak zakon velikih brojeva)
Jak zakon velikih brojeva

$$\bar{X}_n \xrightarrow{s.s.} a$$

važi za niz (X_n) i neku realnu konstantu a ako i samo ako $E|X| < \infty$. Štaviše, ako (X_n) zadovoljava JZVB, onda je $a = \mu$. □

Kolmogorov JZVB važi i za pozitivne (negativne) slučajne promenljive sa beskonačnim očekivanjem, tj. u tom slučaju imamo

$$\bar{X}_n \xrightarrow{s.s.} E(X) = \infty \ (= -\infty).$$

Teorema 1.1.3 (Marcinkiewicz-Zygmundov jak zakon velikih brojeva)
Pretpostavimo da $p \in (0,2)$. Jak zakon velikih brojeva

$$n^{-\frac{1}{p}}(S_n - an) \xrightarrow{s.s.} 0 \qquad (1.1)$$

važi za neku realnu konstantu a ako i samo ako $E|X|^p < \infty$. Ako (X_n) zadovoljava JZVB (1.1), onda možemo da izaberemo

$$a = \begin{cases} 0, & p < 1 \\ \mu, & p \in [1,2). \end{cases}$$

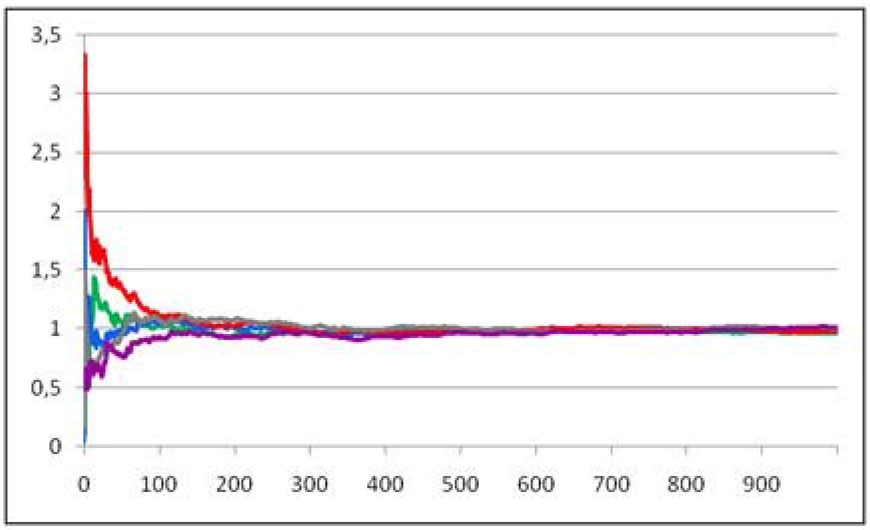
Štaviše, ako $E|X|^p = \infty$ za neko $p \in (0,2)$, onda za svako realno a,

$\limsup_{n\to\infty} n^{-1/p}|S_n - an| = \infty$ s.s. □

Ova teorema daje potpunu karakterizaciju JZVB sa normalizujućim funkcijama n-tog stepena. Pod uslovima ove teoreme dobijamo skoro sigurnu aproksimaciju prvog reda od $\bar{X}_n$:

$\bar{X}_n = \mu + o(n^{\frac{1}{p}-1})$ s.s. (1.2)

što važi ako $E|X|^p < \infty$ za neko $p \in [1,2)$.
Teorema 1.1.3 nam omogućava da izvedemo elementarnu vezu između velikih fluktuacija suma S_n, sabiraka X_n i maksimuma $M_1 = |X_1|$, $M_n = \max(|X_1|, \dots, |X_n|)$, $n \geq 2$.



Grafik 1. *Prikaz konvergencije jakog zakona velikih brojeva: pet trajektorija procesa* $(\frac{S_n}{n})$ *za iid standardnu eksponencijalnu promenljivu* X_n

Posledica 1.1.4 Pretpostavimo da je $p \in (0,2)$. Sledeći uslovi su ekvivalentni:

1) $E|X|^p < \infty (= \infty)$ (1.3)
2) $\lim sup_{n\to\infty} n^{-1/p}|X_n| = 0 \ (= \infty)$ s.s. (1.4)
3) $\lim sup_{n\to\infty} n^{-1/p}M_n = 0 \ (= \infty)$ s.s. (1.5)
4) $\lim sup_{n\to\infty} n^{-1/p}|S_n - an| = 0 \ (= \infty)$ s.s. (1.6)

Ovde a mora da bude izabrano kao u Teoremi 1.1.3. □

Dokaz (1.3) je zadovoljeno akko

$$\sum_{n=1}^{\infty} P(|X| > \varepsilon n^{1/p}) < \infty \ (= \infty) \quad \forall \varepsilon > 0.$$

Borel-Cantellijev argument govori da je ovo ekvivalentno sa

$$P\big(|X_n| > \varepsilon n^{1/p} \quad i.o.\big) = 0 \ (= 1) \ \ \forall \varepsilon > 0.$$

Iz ovog i Teoreme 1.1.3 vidimo da su uslovi (1.3), (1.4) i (1.6) ekvivalentni.
Ekvivalencija (1.4) i (1.5) je posledica elementarne relacije

$$n^{-1/p}|X_n| \le n^{-1/p} M_n \le \max\left(\frac{|X_1|}{n^{1/p}}, \dots, \frac{|X_{n_0}|}{n^{1/p}}, \frac{|X_{n_{0+1}}|}{(n_0+1)^{1/p}}, \frac{|X_{n_{0+2}}|}{(n_0+2)^{1/p}}, \frac{|X_n|}{n^{1/p}}\right)$$

za svako fiksirano $n_0 \le n$. □

1.2 Centralni granični problemi

Centralna granična teorema predstavlja familiju teorema u kojima se dokazuje da niz kumulativnih suma S_n , normiran pogodno izabranim konstantama, ima normalnu raspodelu kad $n \to \infty$, tj. centralna granična teorema tvrdi da zbir velikog broja slučajnih sabiraka, pri čemu je udeo svakog pojedinačnog sabirka u celom zbiru mali, ima približno normalnu raspodelu.

U prethodnom delu smo videli da sume S_n niza iid (X_n) divergiraju skoro sigurno kada su normalizovane sa $n^{1/2}$. Međutim, i dalje možemo da dobijemo informacije o rastu $n^{-1/2}S_n$ ako se prebacimo na konvergenciju u raspodeli (slaba konvergencija). Postavlja se pitanje o mogućim (nedegerativnim) graničnim zakonima za sume S_n kada su one normalizovane i centrirane, tj. o raspodelama koje zadovoljavaju identitet:

$$c_1X_1 + c_2X_2 \overset{d}{=} b(c_1,c_2)X + a(c_1,c_2) \tag{1.7}$$

za sve nenegativne brojeve c_1, c_2 i odgovarajuće realne brojeve $b(c_1, c_2) > 0$ i $a(c_1, c_2)$?
Mogući granični zakoni za sume iid slučajnih promenljivih su samo raspodele koje zadovoljavaju (1.7) za sve nenegativne brojeve c_1, c_2. Mnoge klase raspodela su zatvorene u odnosu na konvoluciju, ali (1.7) zahtev je stroži. Na primer, konvolucija dve Poissonove raspodele je Poissonova raspodela. Medjutim, Poissonova raspodela ne zadovoljava (1.7).
Definicija 1.2.1 (Stabilna raspodela i slučajna promenljiva)
Slučajna promenljiva (raspodela, funkcija gustine) je stabilna ako zadovoljava (1.7) za iid X,X_1, X_2, za sve nenegativne brojeve c_1, c_2 i odgovarajuće realne brojeve
$b(c_1, c_2) > 0$ i $a(c_1, c_2)$. □

Sada razmatramo sumu S_n iid stabilnih slučajnih promenljivih. Na osnovu (1.7), imamo za neke realne konstante a_n i $b_n > 0$ i $X = X_1$,

$$S_n = X_1 + \dots\dots + X_n \overset{d}{=} b_nX + a_n,\ n \ge 1$$

što još možemo zapisati i kao

$$b_n^{-1}(S_n - a_n) \overset{d}{=} X.$$

Zaključujemo da ako je raspodela stabilna, onda je ona granična raspodela za sume iid slučajnih promenljivih. Druge granične raspodele ne postoje.

Teorema 1.2.2 (Granična osobina zakona stabilnosti)
Klasa stabilnih (nedegenerativnih) raspodela se podudara sa klasom svih mogućih (nedegenerativnih) graničnih zakona za (normalizovane i centrirane) sume iid slučajnih promenljivih. □

Najuobičajeniji način analiziranja klase stabilnih raspodela je utvrđivanje njihovih karakterističnih funkcija.

Teorema 1.2.3 (Spektralna reprezentacija zakona stabilnosti)
Stabilna raspodela ima karakterističnu funkciju oblika

$$\phi_x(t) = E(exp\{iXt\}) = exp\{i\gamma t - c|t|^\alpha(1 - i\beta sign(t)z(t,\alpha))\} \quad t \in \mathbb{R} \tag{1.8}$$

gde je γ realna konstanta, $c > 0$, $\alpha \in (0,2]$, $\beta \in [-1,1]$ i

$$z(t,\alpha) = \begin{cases} tg\left(\frac{\pi\alpha}{2}\right), & \alpha \neq 1 \\ -\frac{2}{\pi}\ln|t|, & \alpha = 1. \end{cases}$$

□

Definicija 1.2.4 Broj α u karakterističnoj funkciji (1.8) se naziva karakterističnim eksponentom, a odgovarajuća raspodela α- stabilnom raspodelom. □

Dalje nas zanima koji uslovi impliciraju da normalizovane i centrirane sume S_n slabo konvergiraju ka G_α ako je G_α data α-stabilna raspodela?
Problem može da predstavlja način odabira konstanti $a_n \in \mathbb{R}$ i $b_n > 0$, tako da

$$b_n^{-1}(S_n - a_n) \xrightarrow{d} G_a. \tag{1.9}$$

Teorema D1 osigurava da je granični zakon jedinstveno određen do na pozitivne linearne transformacije.
Definicija 1.2.5 (Domen atrakcije)
Kažemo da slučajna promenljiva X (raspodela od X) pripada domenu atrakcije α-stabilne raspodele G_α ako postoje konstante $a_n \in \mathbb{R}$, $b_n > 0$ tako da važi (1.9). Pišemo $X \in DA(G_\alpha)$ (ili $F \in DA(G_\alpha)$) i kažemo da (X_n) zadovoljava CGT sa granicom G_α. □

Ako nas zanima da li X (ili F) pripada nekom α-stabilnom zakonu, jednostavo pišemo $X\epsilon DA(\alpha)$ (ili $F\epsilon DA(\alpha)$).

Teorema 1.2.6 (Karakterizacija domena atrakcije)

a) Funkcija raspodele F pripada domenu atrakcije normalne raspodele ako i samo ako je $\int_{|y|\leq x} y^2 dF(y)$
sporo varirajuće.

b) Funkcija raspodele F pripada domenu atrakcije α-stabilne raspodele za neko $\alpha < 2$ ako i samo ako

$$F(-x)=\frac{c_1+o(1)}{x^\alpha}L(x),\ 1-F(x)=\frac{c_2+o(1)}{x^\alpha}L(x),\ \ x\to\infty$$

gde je L slabo varirajuća funkcija i c_1, c_2 su nenegativne konstante takve da $c_1 + c_2 > 0$. □

Posmatramo slučaj $\alpha = 2$ detaljnije. Ako $E(X^2) < \infty$, onda

$$\int_{|y|\le x} y^2 dF(y) \to E(X^2),\ x \to \infty,$$

pa $X \in DA(2)$. Štaviše, na osnovu Teoreme D2(f) zaključujemo da je spora varijacija od $\int_{|y|\le x} y^2 dF(y)$ ekvivalentna uslovu repa

$$G(x) = P(|X| > x) = o\left(x^{-2}\int_{|y|\le x} y^2 dF(y)\right),\ \ x \to \infty \tag{1.10}$$

Posledica 1.2.7 (Domen atrakcije za normalnu raspodelu)
Slučajna promenljiva X se nalazi u domenu atrakcije za normalanu raspodelu ako i samo ako je ispunjen jedan od sledećih uslova:

a) $E(X^2) < \infty$
b) $E(X^2) = \infty$ i (1.10). □

Situacija je potpuno drugačija za $\alpha < 2$: $X \in DA(\alpha)$ implicira

$$G(x)= x^{-\alpha}L(x),\ \ x > 0 \tag{1.11}$$

za sporo varirajuću funkciju L i

$$x^2G(x)/\int_{|y|\le x} y^2 dF(y) \to \frac{2-\alpha}{\alpha},\ \ x \to \infty \tag{1.12}$$

Relacija (1.11) i Posledica 1.2.7 pokazuju da je domen atrakcije za normalnu raspodelu uopšteniji nego domen atrakcije α-stabilnog zakona sa eksponentom $\alpha < 2$.
Vidimo da $DA(2)$ sadrži najmanje sve raspodele koje imaju konačan drugi moment.

Posledica 1.2.8 (Momenti raspodele u $DA(\alpha)$)
Ako $X \in DA(\alpha)$, tada

$$E|X|^\delta < \infty \text{ za } \delta < \alpha,$$

$$E|X|^\delta = \infty \text{ za } \delta > \alpha \text{ i } \alpha < 2.$$

Specijalno,

$$Var(X) = \infty \text{ za } \alpha < 2,$$

$$E|X| < \infty \text{ za } \alpha > 1,$$

$$E|X| = \infty \text{ za } \alpha < 1.$$

□

Tvrđenje 1.2.9 (Normalizujuće konstante u CGT)
Normalizujuće konstante u CGT za $F \in DA(\alpha)$ mogu biti izabrane kao jedinstveno rešenje jednačine:

$$Q(b_n) = n^{-1}, n \geq 1. \tag{1.13}$$

Specijalno, ako je $\sigma^2 = Var(X) < \infty$ i $E(X) = 0$, onda

$$b_n \sim n^{1/2}\,\sigma,\ n \to \infty.$$

Ako je $\alpha < 2$, možemo alternativno izabrati (b_n) tako da

$$b_n = inf\{y: G(y) < n^{-1}\},\ n \geq 1 \tag{1.14}$$

(1.14) implicira da

$$G(b_n) \sim n^{-1}, n \to \infty$$

i da, u smislu (1.11),

$$b_n = n^{\frac{1}{2}} L(n),\ n \geq 1,$$

za odgovarajuću sporo varirajuću funkciju L. □

Tvrđenje 1.2.10 (Centrirajuće konstante u CGT)
Centrirajuće konstante a_n u CGT (1.9) mogu biti izabrane kao

$$a_n = n \int_{|y| \leq b_n} y dF(y) \tag{1.15}$$

gde je b_n dato u Tvrđenju 1.2.8. Specijalno, možemo da uzmemo da je $a_n = \tilde{a}n$, gde je

$$\tilde{a} = \begin{cases} \mu, & \alpha \in (1,2] \\ 0, & \alpha \in (0,1) \\ 0, & \alpha = 1 \ \ i\ F\ je\ simetrično. \end{cases} \tag{1.16}$$

□

Teorema 1.2.11 (Uopštena CGT)
Pretpostavimo da $F \in DA(\alpha)$ za neko $\alpha \in (0, 2]$.

a) Ako je $E(X^2) < \infty$, onda
$$\left(\sigma n^{1/2}\right)^{-1}(S_n - \mu n) \xrightarrow{d} \phi$$
za standardnu normalnu raspodelu ϕ sa očekivanjem 0 i disperzijom 1.
b) Ako je $E(X^2) = \infty$ i $\alpha = 2$ ili ako je $\alpha < 2$, onda
$$\left(n^{1/2}\, L(n)\right)^{-1}(S_n - a_n) \xrightarrow{d} G_\alpha$$
za α-stabilnu raspodelu G_α, odgovarajuću sporo varirajuću funkciju$L$ i centrirajuće konstante kao u (1.15).

Specijalno,

$\left(n^{1/\alpha}\, L(n)\right)^{-1} (S_n - \tilde{a}n) \xrightarrow{d} G_\alpha$, gde je $\tilde{a}$ definisano u (1.16). □

Primetimo da je moguće da normalizujuće konstante u CGT budu specijalnog oblika $b_n = cn^{1/\alpha}$ za neke konstante c. To se dešava npr. ako je $E(X^2) < \infty$, ili ako je X α-stabilno.

Definicija 1.2.12 (Domen normalne atrakcije)
Kažemo da X (ili F) pripada domenu normalne atrakcije α-stabilne raspodele G_α ($X \in DNA(G_\alpha)$ ili $F \in DNA(G_\alpha)$) ako $X \in DA(G_\alpha)$ i ako u CGT možemo da odaberemo normalizaciju $b_n = cn^{1/\alpha}$ za neku pozitivnu konstantu c. □

Posledica 1.2.13 (Karakterizacija DNA)

a) Realicija $F \in DNA(2)$ važi ako i samo ako $E(X^2) < \infty$
b) Za $\alpha < 2, F \in DNA(\alpha)$ ako i samo ako
$$F(-x) \sim c_1 x^{-\alpha} \;\; i \;\; 1 - F(x) \sim c_2 x^{-\alpha}, x \to \infty,$$
za nenegativne konstante c_1, c_2 tako da $c_1 + c_2 > 0$

Specijalno, svaka α-stabilna raspodela je u sopstvenom DNA. □

Prema tome, vidimo da $F \in DNA(\alpha)$, $\alpha < 2$ ustvari znači da se odgovarajući rep $G(x)$ ponaša kao stepena funkcija ili kao Pareto raspodela. Primetimo da funkcija raspodele F sa repom $G(x) \sim cx^{-\alpha}$ koji je sličan Paretovoj raspodeli, za neko $\alpha \geq 2$ je u $DA(2)$, a ako $\alpha > 2$, onda $F \in DNA(2)$.

1.3 Poboljšanja centralne granične teoreme

U ovom poglavlju ćemo se ograničiti na slučaj kada je $E(X^2) < \infty$.
Interesuje nas kako možemo utvrditi i poboljšati kvalitet aproksimacije u CGT.

Berry-Esseenova Teorema

Neka ϕ označava funkciju raspodele standardne normalne raspodele. Pišemo:
$$G_n(x) = P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq x\right), \; x \in \mathbb{R}.$$
Znamo da je
$$\Delta_n = sup_{x\in\mathbb{R}} |G_n(x) - \phi(x)| \to 0. \qquad (1.17)$$
U Teoremi 1.2.11 smo formulisali samo slabu konvergenciju, tj. konvergenciju G_n u svakoj tački u kojoj je funkcija ϕ neprekidna. Međutim, ϕ je neprekidna u svakoj tački, stoga (1.17) važi.
Možemo pokazati da brzina kojom Δ_n konvergira ka 0 može biti proizvoljno mala ako ne zahtevamo više od konačnog drugog momenta od X.
Tipična brzina konvergencije je $1/\sqrt{n}$ ako postoji treći moment od X. Dajemo neuniformnu verziju dobro poznate Berry-Esseenove teoreme:

Teorema 1.3.1 (Berry-Esseenova teorema)
Pretpostavimo da $E|X|^3 < \infty$. Onda je

$$|G_n(x) - \phi(x)| \leq \frac{c}{\sqrt{n}\,(1+|x|)^3} \frac{E|X-\mu|^3}{\sigma^3}, \tag{1.18}$$

za svako x, gde je c univerzalna konstanta. Specijalno,

$$\Delta_n \leq \frac{c}{\sqrt{n}} \frac{E|X-\mu|^3}{\sigma^3} \tag{1.19}$$

□

Iz (1.18) vidimo da kvalitet aproksimacije može biti značajno poboljšan za veliko x. Štaviše, na brzinu u (1.18) i (1.19) utiče red veličine količnika $\frac{E|X-\mu|^3}{\sigma^3}$ i konstante c, što je od krucijalnog značaja ako je n malo.
Brzine u (1.18) i (1.19) su optimalne u smislu da postoje nizovi (X_n) takvi da $\Delta_n \asymp (1/\sqrt{n})$.
S druge strane, Berry-Esseenova ocena je prilično pesimistična i može se unaprediti kada su zadovoljeni specijalni uslovi za X, npr. postojanje glatke gustine, funkcije generatrise momenta itd.

Asimptotska ekspanzija

Kao što je gore navedeno, Berry-Esseenova ocena (1.19) je optimalna za određene funkcije raspodele F. Međutim, u nekim slučajevima, funkcija raspodele G_n se može aproksimirati standardnom normalnom funkcijom raspodele ϕ i nekim dodatnim članovima. Tada aproksimirajuća funkcija nije funkcija raspodele. Čest metod aproksimacije se naziva Edgeworthova ili asimptotska ekspanzija. Formalno pišemo

$$G_n(x) = \phi(x) + \sum_{k=1}^{\infty} n^{-k/2}\, Q_k(x), \; x \in \mathbb{R}, \tag{1.20}$$

gde su Q_k izrazi koji uključuju Hermiteove polinome, tačan oblik izraza koji zavisi od momenta od X.
U praksi možemo uzeti samo konačan broj sabiraka Q_k u obzir. Da bismo stekli utisak, razmatramo prva dva sabirka: neka

$$\varphi(x) = (2\pi)^{-1/2}\, exp\{-x^2/2\}, \; x \in \mathbb{R},$$

označava funkciju gustine standardne normalne funkcije raspodele ϕ. Onda, za $x \in \mathbb{R}$,

$$Q_1(x) = -\varphi(x)\frac{H_2(x)}{6}\frac{E(X-\mu)^3}{\sigma^3},$$

$$Q_2(x) = -\varphi(x)\left\{\frac{H_5(x)}{72}\left(\frac{E(X-\mu)^3}{\sigma^3}\right)^2 + \frac{H_3(x)}{24}\left(\frac{E(X-\mu)^4}{\sigma^4} - 3\right)\right\}, \tag{1.21}$$

gde H_i označava Hermiteov polinom stepena i:

$$H_2(x) = x^2 - 1,$$

$$H_3(x) = x^3 - 3x,$$

$$H_5(x) = x^5 - 10x^3 + 15x.$$

Treba naglasiti da Q_k nestaje u slučaju kada je X Gaussova, a

$$E(X-\mu)^3/\sigma^3, \quad E(X-\mu)^4/\sigma^4$$

predstavljaju koeficijente asimetrije i spljoštenosti od X, respektivno. Ove dve veličine mere blizinu funkcije raspodele F u odnosu na ϕ.

Teorema 1.3.2 (Asimptotska ekspanzija u apsolutno neprekidnom slučaju)
Pretposatvimo da je $E|X|^k < \infty$ za neke cele brojeve $k \geq 3$. Ako je F apsolutno neprekidno, onda

$$(1+|x|)^k \left| G_n(x) - \phi(x) - \sum_{i=1}^{k-2} \frac{Q_i(x)}{n^{i/2}} \right| = o\left(\frac{1}{n^{(k-2)/2}}\right),$$

uniformno po x. Specijalno,

$$G_n(x) = \phi(x) + \sum_{i=1}^{k-2} \frac{Q_i(x)}{n^{i/2}} + o\left(\frac{1}{n^{(k-2)/2}}\right),$$

uniformno po x. □

Velika odstupanja

CGT može biti poboljšana ako se posmatra G_n za x koje se uzima iz određenih oblasti (u zavisnosti od n) ili ako je $x = x_n \to \infty$.

Teorema 1.3.3 (Cramerova teorema o velikim odstupanjima)
Pretpostavimo da funkcija generatrise momenta $M(h) = E\,(exp\{hX\})$ postoji u okolini koordinatnog početka. Onda je

$$\frac{1-G_n(x)}{1-\phi(x)} = exp\left\{\frac{x^3}{\sqrt{n}}\,\lambda\left(\frac{x}{n}\right)\right\}\left[1 + O\left(\frac{x+1}{\sqrt{n}}\varphi(x)\right)\right],$$

$$\frac{G_n(-x)}{\phi(-x)} = exp\left\{\frac{-x^3}{\sqrt{n}}\,\lambda\left(\frac{-x}{n}\right)\right\}\left[1 + O\left(\frac{x+1}{\sqrt{n}}\varphi(x)\right)\right],$$

uniformno za pozitivno $x = o(\sqrt{n}\,)$. Ovde je $\lambda(z)$ stepeni red koji konvergira u određenoj okolini koordinatnog početka, a čiji koeficijenti zavise isključivo od momenta od X. □

Stepeni red $\lambda(z)$ se naziva Cramerov red. Umesto utvrđivanja opštih koeficijenata ovog reda, razmatra se određeni slučaj Teoreme 1.3.3.

Posledica 1.3.4 Pretpostavimo da su uslovi Teoreme 1.3.3 zadovoljeni. Tada

$$1 - G_n(x) = \big(1-\phi(x)\big)\exp\left\{\frac{x^3}{6\sqrt{n}}\,\frac{E(X-\mu)^3}{\sigma^3}\right\} + O\left(\frac{1}{\sqrt{n}}\,\varphi(x)\right),$$

$$G_n(-x) = \phi(-x)\exp\left\{\frac{-x^3}{6\sqrt{n}}\,\frac{E(X-\mu)^3}{\sigma^3}\right\} + O\left(\frac{1}{\sqrt{n}}\,\varphi(x)\right),$$

za $x \geq 0,\ x = O\left(n^{1/6}\right)$. Specijalno, ako je $E(X-\mu)^3 = 0$, onda

$G_n(x) - \phi(x) = O\left(\frac{1}{\sqrt{n}}\,\varphi(x)\right)\,,\ x \in \mathbb{R}$ □

Rezultati velikih odstupanja se mogu protumačiti kao poboljšanja stepena konvergencije u CGT. Zaista, neka je $x = x_n \to \infty$, tako da je $x_n = O(n^{1/6})$. Tada, iz Posledice 1.3.4 zaključujemo da je

$$P\left(\left|\frac{S_n-\mu n}{\sigma\sqrt{n}}\right| > x_n\right) = 2\big((1-\phi(x_n)\big) + O\left(\frac{1}{\sqrt{n}}\,exp\left\{\frac{x_n^2}{2}\right\}\right),$$

Gde je x_n izabrano tako da $\frac{S_n-\mu n}{\sigma\sqrt{n}\,x_n} \xrightarrow{P} 0.$ (1.22)

Teorema 1.3.5 (Heydova teorema o velikim odstupanjima)
Neka je $X \in DA(\alpha)$ simetrično i $\alpha \in (0,2)$. Neka je (b_n) niz takav da $b_n \uparrow \infty$ i $P(X > b_n) \sim 1/n$ i $M_1 = X_1,\ M_n = max(X_1, X_2 \dots X_n),\ n \geq 2$ su maksimumi uzoraka. Tada

$\lim_{n\to\infty} \frac{P(S_n > b_n x_n)}{nP(X > b_n x_n)} = \lim_{n\to\infty} \frac{P(S_n > b_n x_n)}{P(M_n > b_n x_n)} = 1$ za svaki niz $x_n \to \infty$. □

1.4 Funkcionalna CGT: Pojava Brownovog kretanja

Neka je (X_n) iid niz sa $0 < \sigma^2 < \infty$. U ovom poglavlju potapamo niz parcijalnih suma (S_n) u proces na $[0,1]$ i razmatramo granični proces koji je ustvari Brownovo kretanje. Prvo razmatramo proces $S_n(\cdot)$ na $[0,1]$ tako da $S_n(n^{-1}k) = \frac{1}{\sigma\sqrt{n}}\,(S_k - \mu k),\ k = 0 \dots\dots n$ i definišemo grafik procesa $S_n(\cdot)$ u svakoj tački $[0,1]$ linearnom interpolacijom između tačaka $\big(k/n, S_n(k/n)\big)$. Ovaj grafik je samo jedna „izlomljena linija“, a trajektorije su neprekidne funkcije. Pretpostavimo da je (X_n) niz iid standardnih normalnih slučajnih promenljivih. Onda su priraštaji $S_n(k/n) - S_n(l/n)$ za $l < k$ Gaussovi sa očekivanjem nula i disperzijom $(k-l)/n$. Štaviše, proces ima nezavisne priraštaje kada je ograničen na tačke $(k/n),\ k = 0, \dots . n$.

Definicija 1.4.1 (Brownovo kretanje)
Neka je $(B_t), t \in [0,1]$ stohastički proces koji zadovoljava sledeće uslove:

(a) Počinje od nule, $B_0 = 0$ s.s.
(b) Ima nezavisne priraštaje: za bilo koju particiju $0 \leq t_0 < t_1 < \dots < t_m \leq 1$ i bilo koje m, slučajne promenljive $B_{t1} - B_{t0}, \dots, B_{tm} - B_{tm-1}$ su nezavisne.
(c) Za svako $t \in [0,1]$, B_t ima Gaussovu raspodelu sa očekivanjem nula i disperzijom t.
(d) Uzoračke putanje/trajektorije su neprekidne sa verovatnoćom 1.

Ovaj proces se naziva Brownovo kretanje ili Wienerov proces na $[0,1]$. □

Posledica ove definicij jeste da priraštaji $B_t - B_s,\ \ t > s$ imaju $\mathcal{N}(0, t-s)$ raspodelu. Brownovo kretanje na $[0, T]$ i na $[0, \infty)$ je definisano modifikacijom Definicije 1.4.1.

Možemo dati i definiciju Brownovog kretanja, kao procesa sa stacionarnim, nezavisnim priraštajima i skoro sigurnim neprekidnim trajektorijama. Može se pokazati da iz ovih osobina sledi da priraštaji moraju imati normalnu raspodelu.
Pišemo $\mathbb{C}[0,1]$ za vektorski prostor neprekidnih funkcija koji ima supremum normu za $x \in \mathbb{C}[0,1]$, $\|x\| = sup_{0\leq t\leq 1}|x(t)|$.
Uvodimo drugi proces na $[0,1]$ koji se poklapa sa $S_n(\cdot)$ u tačkama k/n , $k = 0,1,\dots,n$. Lakši je za konstruisanje, ali je teoretski dosta zahtevniji:

$$\tilde{S}_n(t) = \frac{1}{\sigma\sqrt{n}}\big(S_{[nt]} - \mu[nt]\big),\ \ 0\leq t \leq 1,$$

gde $[y]$ označava ceo deo realnog broja y. Ovaj proces ima nezavisne priraštaje koji su Gaussovi ako je X Gaussova promenljiva. Njegove trajektorije nisu neprekidne, ali imaju moguće skokove u tačkama k/n. U svakoj tački $[0,1)$ su neprekidni zdesna i u svakoj tački $(0,1]$ postoji limes sleva. Prema tome, proces $S_n(\cdot)$ ima cadlag trajektorije, tj. one pripadaju prostoru $\mathbb{D}[0,1]$. Prostor $\mathbb{D}[0,1]$ cadlag funkcija može da bude snabdeven različitim metrikama da bi se na njemu definisala slaba konvergencija.
Međutim, naš granični proces će biti Brownovo kretanje koji uzima vrednosti u $\mathbb{C}[0,1]$ tako da je dozvoljeno da uzmemo supremum normu kao odgovarajuću metriku na $\mathbb{D}[0,1]$.

Definicija 1.4.2 (Proces sa nezavisnim, stacionarnim priraštajima)
Neka je $\xi = (\xi_t)_{0\leq t\leq 1}$ stohastički proces. Tada ξ ima nezavisne priraštaje ako su za svako $0\leq t_0 < \dots < t_m \leq 1$ i svako $m \geq 1$ slučajne promenljive $\xi_{t1} - \xi_{t0},\dots,\xi_{tm} - \xi_{tm-1}$, nezavisne. Za ovaj proces kažemo da ima stacionarne priraštaje ako za sve
$0 \leq s \leq t \leq 1$, slučajne promenljive $\xi_t - \xi_s$ i ξ_{t-s} imaju istu raspodelu.
Proces sa nezavisnim, stacionarnim priraštajima i trajektorijama u $\mathbb{D}[0,1]$ se još naziva Levyjev proces. □

Definicija 1.4.3 (α-stabilno kretanje)
Za stohastički proces $(\xi_t)_{0\leq t\leq 1}$ sa uzoračkim putanjama na $\mathbb{D}[0,1]$ se kaže da je α-stabilno kretanje ako važe sledeći uslovi:

(a) Počinje od nule: $\xi_0 = 0$ skoro sigurno.
(b) Ima nezavisne, stacionarne priraštaje.
(c) Za svako $t \in [0,1]$, ξ_t ima α-stabilnu raspodelu sa fiksiranim parametrima $\beta\epsilon[-1,1]$ i $\gamma = 0$ u spektralnoj reprezentaciji (1.8).

□

Lema 1.4.4 Za α-stabilno kretanje $(\xi_t)_{0\leq t\leq 1}$ važi $\xi_t - \xi_s = (t-s)^{1/\alpha}\xi_1\ \ 0 \leq s \leq t \leq 1$. □

Na osnovu leme 1.4.4 možemo da izvedemo konačno dimenzionalne raspodele α-stabilnog kretanja:

$$(\xi_{t1},\xi_{t2},\dots,\xi_{tm}) = \big(\xi_{t1},\xi_{t1} + (\xi_{t2} - \xi_{t1}),\dots,\xi_{t1} + (\xi_{t2} - \xi_{t1}) + \dots + (\xi_{tm} - \xi_{tm-1})\big) =$$

$$\Big(t_1^{1/\alpha}\,Y_1, t_1^{1/\alpha}\,Y_1 + (t_2 - t_1)^{1/\alpha}Y_2,\dots,t_1^{1/\alpha}\,Y_1 + (t_2 - t_1)^{1/\alpha}Y_2 + \dots + (t_m - t_{m-1})^{1/\alpha}Y_m\Big)$$

za bilo koji realan broj 0≤ $t_1 < \cdots < t_m \leq 1$ i iid α-stabilne slučajne promenljive $Y_1, \ldots, Y_m$ takve da $Y_1 = \xi_1$.

1.5 Slučajne sume

1.5.1 Slučajno indeksirani nizovi

Slučajne, (tj. slučajno indeksirane) sume su osnova matematike u osiguranju. Ukupan iznos štete jednog portfolija se obično modelira slučajnim sumama

$$S(t) = S_{N(t)} = \begin{cases} 0, & N(t) = 0, \\ X_1 + \cdots + X_{N(t)}, & N(t) \geq 0, \end{cases} \quad t \geq 0,$$

gde je $\left(N(t)\right)_{t\geq 0}$ stohastički proces na $[0,\infty)$ takve da su slučajne promenljive $N(t)$ nenegativne celobrojne vrednosti.
$\left(N(t)\right)$ je po pretpostavci generisan nizom $(T_n)_{n\geq 1}$ nenegativnih slučajnih promenljivih tako da $0 \leq T_1 \leq T_2 \leq \cdots$ s.s. i

$$N(t) = sup\{n \geq 1 : T_n \leq t\}, \quad t \geq 0. \tag{1.23}$$

Znamo, $supA = 0$ ako je $A = \emptyset$. Tada se ovo zove proces prebrajanja. Slučajna promenljiva X_n može biti definisana kao individualan iznos štete, a $S(t)$ je ukupan iznos šteta u portfoliju do vremena t. U finansijskom smislu $N(t)$ na primer može da predstavlja (slučajan) broj promena pozicije u deviznom portfoliju koji je zasnovan na visokofrekventnim posmatranjima. Tada $S(t)$ predstavlja ukupan prinos u $[0, t]$.

Primer 1.5.2 (Homogeni Poissonov proces i mešoviti Poissonov proces)
U Cramer-Lundbergovom modelu (Definicija D7) se pretpostavlja da su (X_n) i $\left(N(t)\right)$ nezavisni i da je $\left(N(t)\right)$ homogeni Poissonov proces sa parametrom $\lambda > 0$ tj. to je proces prebrajanja (1.23) sa $T_n = Y_1 + \cdots + Y_n$, $n \geq 1$ i (Y_n) (vremena između pristizanja zahteva) su eksponencijalne iid slučajne promenljive sa očekivanjem $1/\lambda$. Svaki proces prebrajanja je generisan iid procesom suma (T_n) i još se naziva obnavljajućim procesom prebrajanja.

Drugačije, (homogeni) Poissonov proces se definiše sledećim osobinama:

(a) Počinje od nule: $N(0) = 0$
(b) Ima nezavisne, stacionarne priraštaje
(c) Za svako $t > 0$, $N(t)$ je Poissonva slučajna promenljiva sa parametrom λt

$$P(N(t) = n) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}, n = 0,1,2 \ldots$$

Poissonov proces $\left(N(t)\right)$ je proces sa trajektorijama na $\mathbb{D}[0,\infty)$ koji ide u ∞ kad $t \to \infty$ i ima skokove visine 1 u slučajnim momentima T_n. To je takođe Levyjev proces (Definicija 1.4.2).

Ako su $(N(t))$ i (X_n) nezavisni, onda se proces $(S(t))_{t\geq 0}$ naziva mešoviti Poissonov proces.
Fluktuacija slučajnih suma $(S(t))$ za veliko t se može opisati i preko graničnih teorema. Dalje ćemo dati osnovne teoreme koje pokazuju da su asimptotska ponašanja (S_n) i $(S(t))$ usko povezana.
U ovom delu, $(Z_n)_{n\geq 0}$ je opšti niz slučajnih promenljivih, a $(N(t))_{t\geq 0}$ je proces nenegativnih celobrojnih slučajnih promenljivih $N(t)$.

Lema 1.5.3
Pretpostavimo da $Z_n \xrightarrow{s.s.} Z$ kada $n \to \infty$ i $N(t) \xrightarrow{s.s.} \infty$ $(N(t) \xrightarrow{P} \infty)$ kad $t \to \infty$. Tada

$$Z_{N(t)} \xrightarrow{s.s.} Z,\ (Z_{N(t)} \xrightarrow{P} Z),\ t \to \infty.$$

Dokaz
Pretpostavimo da $N(t) \xrightarrow{s.s.} \infty$. Neka je

$$A_1 = \{\omega: N(t)(\omega) \to \infty, t \to \infty\},$$

$$A_2 = \{\omega: Z_n(\omega) \to Z(\omega), n \to \infty\}$$

i primetimo da je $P(A_1) = P(A_2) = 1$, pa
$P(\{\omega: Z_{N(t)(\omega)}(\omega) \to Z(\omega), t \to \infty\}) \geq P(A_1 \cap A_2) = 1$ tj. $Z_{N(t)} \xrightarrow{s.s.} Z$
Sada pretpostavimo da $N(t) \xrightarrow{P} \infty$ kad $t \to \infty$. Za svaki niz $t_k \to \infty, N(t)_k \xrightarrow{P} \infty$ kad $k \to \infty$ i postoji podniz $t_{k_j} \uparrow \infty$ tako da $N\left(t_{k_j}\right) \xrightarrow{s.s.} \infty$ kad $j \to \infty$. Iz prvog dela dokaza $Z_{N(t_{k_j})} \xrightarrow{s.s.} Z$, pa $Z_{N(t_{k_j})} \xrightarrow{P} Z$.
Prema tome, svaki niz $(Z_{n(t_k)})$ sadrži podniz koji konvergira u verovatnoći ka Z. Pošto je konvergencija u verovatnoći metrizabilna, to znači da $Z_{N(t)} \xrightarrow{P} Z$. □

Teorema 1.5.4 (Marcinkiewicz-Zygmundov JZVB za slučajne sume)
Pretpostavimo da $E|X|^p < \infty$ za neko $p \in (0,2)$ i $N(t) \xrightarrow{s.s.} \infty$. Tada

$$\left(N(t)\right)^{-1/p}\left(S(t) - aN(t)\right) \xrightarrow{s.s.} 0, \tag{1.24}$$

gde je $a = \begin{cases} 0, & p < 1 \\ \mu = E(X), & p \in [1,2) \end{cases}$ □

Sada posmatramo slučaj slabe konvergencije. Želimo da izvedemo CGT za slučajne sume. Naredna lema pokriva mnoge slučajeve od praktičnog interesa, npr. mešoviti Poissonov slučaj.

Lema 1.5.5
Pretpostavimo da su (Z_n) i $(N(t))$ nezavisni i $N(t) \xrightarrow{P} \infty$ kad $t \to \infty$. Ako $Z_n \xrightarrow{d} Z$ kad $n \to \infty$, onda $Z_{N(t)} \xrightarrow{d} Z$ kad $t \to \infty$.

Teorema 1.5.6 (CGT za slučajne sume)
Pretpostavimo da su $(X_n)\; i\; \big(N(t)\big)$ nezavisne i da $N(t) \xrightarrow{P} \infty$ i da $F \in DA(\alpha)$ za neko $\alpha \in (0,2]$. Tada Teorema 1.2.11 i dalje važi ukoliko se n svugde zameni sa $N(t)$, tj. postoje odgovarajuće centrirajuće konstante a_n i slabo varirajuća funkcija L tako da $\Big(\big(N(t)\big)^{1/\alpha} L\big(N(t)\big)\Big)^{-1} \big(S(t) - a_{N(t)}\big) \xrightarrow{d} G_\alpha,\ t \to \infty$, za neku α-stabilnu raspodelu G_α. □

Uslov da su procesi (Z_n) i $\big(N(t)\big)$ nezavisni može biti znatno oslabljen:

Lema 1.5.7 (Anscombeova teorema)
Pretpostavimo da postoji celobrojna funkcija $b(t) \uparrow \infty$ tako da

$$\frac{N(t)}{b(t)} \xrightarrow{P} 1, t \to \infty \tag{1.25}$$

i da naredni, Anscombeov uslov važi:

$\forall \varepsilon > 0, \forall \eta > 0\ \exists \delta > 0\ \exists n_0$ tako da

$$P\big(\max_{m:|m-n|<n\delta} |Z_m - Z_n| > \varepsilon\big) < \eta,\ n > n_0. \tag{1.26}$$

Ako $Z_n \xrightarrow{d} Z$ kad $n \to \infty$, tada $Z_{N(t)} \xrightarrow{d} Z$ kad $t \to \infty$ □

Uslov (1.25) osigurava da slučajan indeks N_t može biti zamenjen determinističkom funkcijom $b(t)$ i ako uradimo to sa $Z_{N(t)}$, tj. ako zamenimo $Z_{N(t)}$ sa $Z_{b(t)}$, onda (1.26) garantuje da je greška $\big|Z_{N(t)} - Z_{b(t)}\big|$ zanemarljiva. Drugim rečima, Anscombeov uslov je specifična stohastička osobina neprekidnosti niza (Z_n).
Primetimo da je (1.25) zadovoljeno za široku klasu obnavljajućih procesa prebrajanja, uključujući homogeni Poissonov proces. Štaviše, (1.26) važi za (normalizovane i centralizovane) sume S_n.

Teorema 1.5.8 (Anscombeov tip CGT za slučajne sume)
Pretpostavimo da

$$\frac{N(t)}{t} \xrightarrow{P} \lambda,\ t \to \infty, \tag{1.27}$$

za neko pozitivno λ i da $F \in DA(\alpha)$ za neko $\alpha \in (0,2]$ sa

$$\big(n^{1/\alpha} L(n)\big)^{-1} (S_n - \tilde{a}n) \xrightarrow{d} G_\alpha, \tag{1.28}$$

za α-stabilnu raspodelu G_α i sporo varirajuću funkciju L. Ovde je

$$\tilde{a} = \begin{cases} 0, & \alpha \leq 1 \\ \mu, & \alpha \in (1,2]. \end{cases}$$

Tada,

$$\Big(\big(N(t)\big)^{1/\alpha} L\big(N(t)\big)\Big)^{-1} \big(S(t) - \tilde{a}N(t)\big) \xrightarrow{d} G_\alpha, \tag{1.29}$$

$$\big((\lambda t)^{1/\alpha} L(t)\big)^{-1} \big(S(t) - \tilde{a}N(t)\big) \xrightarrow{d} G_\alpha. \tag{1.30}$$

Specijalno, ako $\sigma^2 < \infty$ onda
$(\lambda\sigma^2 t)^{-1/2}\big(S(t) - \mu N(t)\big) \xrightarrow{d} \phi$, gde je ϕ standardna normalna raspodela. □

U pogledu Teoreme 1.2.11, (1.19) je samo restrikcija za raspodelu od X u slučaju $\alpha = 1$.

1.5.2 Obnavljajući proces prebrajanja

Razmatramo obnavljajući proces prebrajanja $\big(N(t)\big)_{t\geq 0}$ tj.
$\mathrm{N(t)} = \sup\{\mathrm{n} \geq 1 : T_n \leq t\}, \mathrm{t} \geq 0$ (1.31)
i $T_n = Y_1 + \cdots + Y_n$, $n \geq 1$ za iid nenegativne (nenula) slučajne promenljive Y, Y_1, Y_2 ... Homogeni Poissonov process (Primer 1.5.2) je takav obnavljajući proces prebrajanja gde je Y eksponencijalno sa očekivanjem $1/\lambda$.
Ovde odgovaramo na pitanje: Koji je red veličine $N(t)$ kad $t \to \infty$?
Važi da je $\{T_n \leq t\} = \{N(t) \geq n\}$. Kolmogorov JZVB implicira da $T_n \xrightarrow{s.s.} \infty$, stoga
$N(t) \xrightarrow{s.s.} \infty$.

Teorema 1.5.9 (Marcinkiewicz-Zygmund jaki zakoni velikih brojeva / Zakon iteriranog logaritma za obnavljajući proces prebrajanja)
Pretpostavimo da je $E(Y) = 1/\lambda \leq \infty$ (ako $E(Y) = \infty$, onda $\lambda = 0$). Tada je
$t^{-1}N(t) \xrightarrow{s.s.} \lambda.$ (1.32)
Ako $E(Y^p) < \infty$ za neko $p \in (1{,}2)$, onda
$t^{-1/p}(N(t) - \lambda t) \xrightarrow{s.s.} 0.$ (1.33)
Ako $\sigma_Y^2 = Var(Y) < \infty$, onda
$\lim sup_{t\to\infty}(2t \ln\ln t)^{-1/2}\,(N(t) - \lambda t) = -\lim inf_{t\to\infty}(2t \ln\ln t)^{-1/2}\,(N(t) - \lambda t) = \sigma_Y \lambda^{3/2}$
skoro sigurno.

Skica dokaza Ograničavamo se na to da pokažemo JZVB (1.32) i (1.33). Kolmogorov JZVB pokazuje da
$\frac{T_{N(t)}}{N(t)} \xrightarrow{s.s.} \frac{1}{\lambda}$.

Ovim i lemom o uklještenim nizovima primenjenim na $\frac{T_{N(t)}}{N(t)} \leq \frac{t}{N(t)} \leq \frac{T_{N(t)+1}}{N(t)+1}\frac{N(t)+1}{N(t)}$ dokazujemo (1.32).
Sad pretpostavimo da je $E(Y^p) < \infty$ za neko $p \in (1{,}2)$. Primetimo da je

$$n^{-1/p}(T_{n+1} - T_n) = n^{-1/p}\, Y_{n+1} \xrightarrow{s.s.} 0.$$

Ovo, Marcinkiewicz-Zygmund jaki zakoni velikih brojeva iz Teoreme 1.5.4 i (1.32) impliciraju da

$t^{-1/p}\big(T_{N(t)+1} - T_{N(t)}\big) = t^{-1/p}\, Y_{N(t)+1} \xrightarrow{s.s.} 0.$

Ovo,Teorema 1.5.4 i lema o uklještenim nizovima primenjeni na

$$\frac{\lambda t-N(t)}{t^{1/p}} \leq \frac{\lambda T_{N(t)+1}-N(t)}{t^{1/p}} \leq \frac{\lambda t-N(t)}{t^{1/p}} + \frac{\lambda T_{N(t)+1}}{t^{1/p}} \quad (1.34)$$

daju (1.33). □

Iz ove teoreme zaključujemo da je $E\big(N(t)\big) \sim \lambda t \; i \; Var\big(N(t)\big) \sim \sigma_Y^2 \lambda^3 t.$

Tvrđenje 1.5.10 (Momenti obnavljajućeg procesa prebrajanja)
Sledeće relacije važe:

(a) $E\big(N(t)\big) = \big(\lambda + o(1)\big)t,$ kad $t \to \infty$.
(b) Pretpostavimo da $\sigma_Y^2 = Var(Y) < \infty$. Tada
$E\big(N(t)\big) = \lambda t + O(1),\, t \to \infty.$

$Var\big(N(t)\big) = \sigma_Y^2 \lambda^3 t + o(t),\, t \to \infty.$ □

Teorema 1.5.11 (CGT za obnavljajuće procese prebrajanja)
Pretpostavimo da je $\sigma_Y^2 < \infty$. Tada

$$(\sigma_Y^2 \lambda^3 t)^{-1/2}(N(t) - \lambda t) \xrightarrow{\mathrm{d}} \phi, \quad (1.35)$$

gde je Φ standardna normalna raspodela.

Dokaz Nastavljamo kao u (1.34):

$$\frac{\lambda t - N(t)}{(\sigma_Y^2 \lambda^3 t)^{1/2}} \leq \frac{\lambda T_{N(t)+1} - N(t)}{(\sigma_Y^2 \lambda^3 t)^{1/2}} \leq \frac{\lambda t - N(t)}{(\sigma_Y^2 \lambda^3 t)^{1/2}} + \frac{\lambda Y_{N(t)+1}}{(\sigma_Y^2 \lambda^3 t)^{1/2}}$$

Zbog nezavisnosti imamo:

$$P\big(Y_{N(t)+1} > \varepsilon t^{1/2}\big) = E\Big(P\Big(Y_{N(t)+1} > \varepsilon t^{1/2} | N(t)\Big)\Big) = P\big(Y > \varepsilon t^{1/2}\big), \forall \varepsilon > 0$$

Stoga, $\frac{\lambda t-N(t)}{(\sigma_Y^2 \lambda^3 t)^{1/2}} = \frac{\lambda T_{N(t)}-N(t)}{(\sigma_Y^2 \lambda^3 t)^{1/2}} + o_p(1).$ (1.36)

U pogledu Teoreme 1.58 i uz pomoć teoreme neprekidnog preslikavanja, desna strana slabo konvergira ka Φ. Ovo dokazuje teoremu. □

1.5.3 Slučajne sume koje zavise od obnavljajućeg procesa prebrajanja

U ovom poglavlju razmatramo jedan od najvažnijih modela u matematici osiguranja.
Tu pretpostavljamo da slučajne sume $S(t) = S_{N(t)}$ zavise od obnavljajućeg procesa prebrajanja, kao što je definisano u (1.31). Proces $\big(S(t)\big)$ je model za ukupan iznos šteta u jednom portfoliju.
Obnavljajući i Cramer-Lundbergov model (Definicija D6) su uključeni kao posebni slučajevi kada su $\big(N(t)\big)$ i (X_n) nezavisni. Dalje nas zanimaju asimptotske osobine procesa $\big(S(t)\big)$.

Podsetimo se iz Poglavlja 1.5.2 da je $N(t) \xrightarrow{s.s.} \infty$.
Dalje možemo da primenimo Marcinkiewicz-Zygmund JZVB za slučajne sume (Teorema 1.5.4)

$$\big(N(t)\big)^{-1/p}\big(S(t) - aN(t)\big) \xrightarrow{s.s.} 0, \tag{1.37}$$

ako je $E|X|^p < \infty$ za neko $p \in (0,2)$ i

$$a = \begin{cases} 0, & p < 1 \\ \mu = E(X), & p \in [1,2). \end{cases} \tag{1.38}$$

Nameće se pitanje:
Da li možemo da zamenimo $N(t)$ u (1.37) determinističkom funkcijom, npr. λt ?
Generalno ne. Međutim, Teorema 1.5.9 kaže da je $N(t)/t \xrightarrow{s.s.} \lambda$, ako je $E(Y) < \infty$. Dalje, možemo da zamenimo normalizirajući proces $\big(N(t)\big)^{1/p}$ sa $(\lambda t)^{1/p}$. Proces centriranja izaziva određene probleme. U nastavku, pretpostavljamo da je $E|X|^p < \infty$ za neko $p \in [1,2)$. Pišemo

$$t^{-1/p}(S(t) - \lambda\mu t) = t^{-1/p}\big(S(t) - \mu N(t)\big) + \mu t^{-1/p}(N(t) - \lambda t).$$

Zbog (1.37), prvi izraz s desne strane jednakosti konvergira u nulu skoro sigurno pod uslovom da je prvi momenat od Y konačan. S druge strane

$$\mu t^{-1/p}(N(t) - \lambda t) \xrightarrow{s.s.} 0 \tag{1.39}$$

generalno ne važi. Međutim, ako je $E(Y^p) < \infty$, zaključujemo iz Teoreme 1.5.9 da je (1.39) zadovoljeno.

Teorema 1.5.12 (Marcinkiewicz-Zygmund JZVB za slučajne sume)
Pretpostavimo da $E|X|^p < \infty$ za neko $p \in (0,2)$.

(a) Ako $E(Y) < \infty$, onda $t^{-1/p}\big(S(t) - aN(t)\big) \xrightarrow{s.s.} 0$, gde je a definisano sa (1.38)
(b) Ako je $p \geq 1$ i $E(Y^p) < \infty$ onda $t^{-1/p}(S(t) - \mu\lambda t) \xrightarrow{s.s.} 0$. □

U slučaju slabe konvergencije možemo da nastavimo analogno. Pošto $N(t) \to \infty$ skoro sigurno za obnavljajući proces prebrajanja, CGT za slučajne sume važi pod slabim uslovima.

Teorema 1.5.13 (Anscombeov tip CGT za slučajne sume)
Pretpostavimo da $F \in DA(\alpha)$ za neko $\alpha \in (0,2]$ i da je

$$\left(n^{1/\alpha}L(n)\right)^{-1}(S_n - \tilde{a}n) \xrightarrow{d} G_\alpha$$

za neku α-stabilnu raspodelu G_α i sporo varirajuću funkciju L. Ovde je

$$\tilde{a}_n = \begin{cases} 0, & \alpha \leq 1 \\ \mu, & \alpha \in (1,2]. \end{cases}$$

(a) Ako $E(Y) < \infty$, onda

$$\left((\lambda t)^{1/\alpha}L(t)\right)^{-1}\big(S(t) - \tilde{a}N(t)\big) \xrightarrow{d} G_\alpha. \tag{1.40}$$

Specijalno, ako je $\sigma^2 = Var(X) < \infty$, onda je

$$(\lambda\sigma^2 t)^{-1/2}\big(S(t)-\mu N(t)\big) \xrightarrow{\mathrm{d}} \Phi, \tag{1.41}$$

gde je Φ standardna normalna raspodela.

(b) Ako je $\alpha \in (1,2)$ i $E(Y)^p < \infty$ za neko $p > \alpha$, onda je

$$\left((\lambda t)^{1/\alpha}L(t)\right)^{-1}(S(t)-\lambda\mu t) \xrightarrow{d} G_\alpha. \tag{1.42}$$

□

Dokaz:

Ako je $E(Y) < \infty$, onda je na osnovu Teoreme 1.5.9, $N(t)/t \xrightarrow{s.s.} \lambda$ i Teorema 1.58 odmah važi. Odatle slede (1.40) i (1.41).
Ako je $E(Y^p) < \infty$ i $p \in [1,2)$, onda je,onda je, na osnovu Teoreme 1.5.9

$$t^{-1/p}(N(t)-\lambda t) \xrightarrow{s.s.} 0. \tag{1.43}$$

Odatle,

$$\left((\lambda t)^{1/\alpha}L(t)\right)^{-1}(S(t)-\mu Nt) = \left((\lambda t)^{1/\alpha}L(t)\right)^{-1}\big(S(t)-\mu N(t)\big)+\mu\big(N(t)-\lambda(t)\big) =$$

$$\left((\lambda t)^{1/\alpha}L(t)\right)^{-1}\big(S(t)-\mu N(t)\big)+o(1) \text{ skoro sigurno.} \tag{1.44}$$

Ovde smo koristili (1.43) i činjenicu da je za $p > \alpha$

$$\lim_{t\to\infty}\frac{t^{1/\alpha}L(t)}{t^{1/p}} = \infty,$$

što je posledica spore varijacije L. Relacija (1.42) odmah sledi iz (1.40) i (1.44) zbog teoreme o neprekidnom preslikavanju (D3). □

Isključili smo slučaj $\alpha = 2$ iz (1.42) jer tada ovaj metod dokazivanja ne važi. (1.44) više nije primenljivo ako je $\sigma^2 < \infty$, što sledi iz CGT u Teoremi 1.5.11.
Sada ćemo da kombinujemo CGT za $\big(S(t)\big)$ i za $\big(N(t)\big)$. Pretpostavimo da je $\sigma^2 < \infty$ i $\sigma_Y^2 < \infty$. Koristeći (1.36), dobijamo

$$\begin{aligned} \mathrm{t}^{-1/2}(\mathrm{S(t)}-\lambda\mu\mathrm{t}) &= \mathrm{t}^{-1/2}\Big(\big(\mathrm{S(t)}-\mu\mathrm{N(t)}\big)+\mu(\mathrm{N(t)}-\lambda\mathrm{t})\Big) \\ &= \mathrm{t}^{-1/2}\Big(\big(\mathrm{S(t)}-\mu\mathrm{N(t)}\big)+\mu(\mathrm{N(t)}-\lambda\mathrm{t})\Big)+\mathrm{o_P}(1) \\ &= \mathrm{t}^{-1/2}\textstyle\sum_{\mathrm{i=1}}^{\mathrm{N(t)}}(\mathrm{X_i}-\mu\lambda\mathrm{Y_i})+\mathrm{o_P}(1) \end{aligned} \tag{1.45}$$

Slučajne promenljive $X_i^{;} = X_i - \mu\lambda Y_i$ imaju očekivanje nula. Štaviše, niz $\big(X_i^{;}\big)$ je iid ako je $\big((X_n, Y_n)\big)$ iid. Pod ovim uslovom, Teorema 1.58 se odmah primenjuje na (1.45), odakle sledi:

Teorema 1.5.14

Pretpostavimo da je $\big((X_n, Y_n)\big)$ niz iid slučajnih vektora i da je $\sigma^2 < \infty$ i $\sigma_Y^2 < \infty$. Tada

$$(Var(X-\mu\lambda Y)\lambda t)^{-1/2}(S(t)-\mu\lambda t) \xrightarrow{\mathrm{d}} \phi$$

gde Φ označava standardnu normalnu raspodelu.

Specijalno, ako su (X_n) i (Y_n) nezavisni, onda
$\left((\sigma^2 + (\mu\lambda\sigma_Y)^2)\lambda t\right)^{-1/2}(S(t) - \mu\lambda t) \overset{\mathrm{d}}{\to} \phi$. □

2. Fluktuacije maksimuma

U ovom poglavlju ćemo se baviti klasičnom teorijom ekstremnih vrednosti. Glavni rezultat jeste Fisher-Tippetova teorema koja određuje oblik granične raspodele za centrirane i normalizovane maksimume.
Osnovni alat za proučavanje retkih događaja je Poissonova aproksimacija, koja je takođe od ključnog značaja za slabu graničnu teoriju, statistiku gornjeg reda i za slabu konvergenciju procesa tačaka.
Asimptotske teorije za maksimume i sume su komplementarne, ali su u isto vreme i u suprotnosti. Odgovarajući rezultati postoje za linearno transformisane sume i maksimume: stabilne raspodele odgovaraju max-stabilnim raspodelama, domeni atrakcije maksimalnim domenima atrakcije. Granične teoreme za maksimume i sume zahtevaju odgovarajuće normalizujuće i centrirajuće konstante. U poglavlju 2.4 uvodimo funkciju očekivanog viška.
Maksimalni domen atrakcije raspodele ekstremne vrednosti može biti opisan regularnom disperzijom i njenim ekstenzijama (poglavlje 2.3).
Teorija u poglavlju 2.4 nam omogućava da predstavimo različite rezultate prethodnih poglavlja na kompaktan način. Ključ je uopštena raspodela ekstremnih vrednosti što nas dovodi do generalizovane Pareto raspodele.
Skoro sigurno ponašanje maksimuma razmatramo u poglavlju 2.5.

2.1 Granične verovatnoće maksimuma

$X, X_{1,} X_{2,} \dots$ predstavlja niz iid nedegenerativnih slučajnih promenljivih sa zajedničkom funkcijom raspodele F. U ovom poglavlju ćemo razmatrati funkcije uzoračkih maksimuma

$$M_1 = X_1, \;\; M_n = \max(X_1, \dots X_n)\,, n \geq 2.$$

Odgovarajući rezultati za minimume se mogu dobiti uz pomoć ovih maksimuma koristeći identitet

$$min(X_1, \dots X_n) = -max(-X_1, \dots - X_n).$$

Funkcija raspodele maksimuma M_n je:

$$P(M_n \leq x) = P(X_1 \leq x, \dots, X_n \leq x) = F^n(x), \qquad x \in \mathbb{R}, n \in \mathbb{N}.$$

Ekstremi se nalaze u okolini gornjeg kraja nosača funkcije raspodele, stoga se intuitivno, asimptotsko ponašanje M_n povezuje sa funkcijom raspodele F na njenom desnom repu u blizini krajnje desne tačke. Sa

$$x_F = sup\{x \in \mathbb{R}\colon\; F(x) < 1\}$$

označavamo krajnju desnu tačku od F. Za $x < x_F$ dobijamo da je

$$P(M_n \leq x) = F^n(x) \to 0,\ n \to \infty,$$

a u slučaju $x_F < \infty$, za $x \geq x_F$ dobijamo da je

$$P(M_n \leq x) = F^n(x) = 1.$$

Prema tome, $M_n \xrightarrow{P} x_F$ kad $n \to \infty$, gde je $x_F \leq \infty$.
Pošto je niz neopadajući po n, konvergira skoro sigurno i odatle zaključujemo

$$M_n \xrightarrow{s.s} x_F, n \to \infty. \tag{2.1}$$

Ova činjenica ne obezbeđuje mnogo informacija. Bolji uvid u red veličine maksimuma je dat rezultatima slabe konvergencije za centrirane i normalizovane maksimume. Ovo je jedna od glavnih tema klasične teorije o ekstremnim vrednostima. Na primer, fundamentalna Fisher-Tippetova teorema (Teorema 2.2.3) se sastoji od sledećeg:

Ako postoje konstante $c_n > 0\ i\ d_n \in \mathbb{R}$ tako da

$$c_n^{-1}(M_n - d_n) \xrightarrow{d} H,\ \ n \to \infty \tag{2.2}$$

za neku nedegenerativnu raspodelu H, onda H mora biti jedna od tri distribucije ekstremne vrednosti. Ovo je slično centralnoj graničnoj teoremi, gde su stabilne raspodele jedini mogući nedegenerativni granični zakoni.
Kao posledicu toga, treba da razmatramo verovatnoće oblika:

$$P(c_n^{-1}(M_n - d_n) \leq x),$$

što se još može zapisati kao

$$P(M_n \leq u_n), \tag{2.3}$$

gde je $u_n = u_n(x) = c_n x + d_n$. Prvo razmatramo (2.3) za opšte nizove (u_n), a posle se vraćamo na linearne transformacije kao u (2.2).
Zanima nas koji uslovi za F garantuju postojanje od $P(M_n \leq u_n)$ kad $n \to \infty$ za odgovarajuće konstante u_n.
Ispostavlja se da su potrebni određeni uslovi neprekidnosti za F u njenoj krajnjoj desnoj tački, što isključuje mnoge važne raspodele. Na primer, ako F ima Poissonovu raspodelu, onda $P(M_n \leq u_n)$ nema granicu na $(0,1)$ ni za jedan niz (u_n). Ovo implicira da normalizovani maksimumi iid slučajnih promenljivih sa Poissonovom raspodelom nemaju nedegenerativne granične raspodele. Odavde se vidi krucijalna razlika između suma i maksimuma. U prethodnom slučaju, CGT daje normalnu raspodelu kao granicu pod opštim uslovom momenta $E(X^2) < \infty$. Ako je $E(X^2) = \infty$ uključuje se relativno mala klasa α-stabilnih graničnih raspodela. Samo u tom slučaju debelog repa, uslovi na repu $\bar{F} = 1 - F$ garantuju postojanje granične raspodele. Za razliku od suma, uvek su nam potrebni delikatni uslovi na repu $\bar{F}$ da bi se osigurala konvergencija $P(M_n \leq u_n)$ ka netrivijalnoj granici, tj. broja iz$(0,1)$.

Tvrđenje 2.1.1 (Poissonova aproksimacija)
Za dato $\tau \in [0,\infty)$ i niz realnih brojeva (u_n), sledeće relacije su ekvivalentne

$$n\bar{F}(u_n) \to \tau, \tag{2.4}$$

$$P(M_n \leq u_n) \to e^{-\tau}. \tag{2.5}$$

□

Dokaz. Prvo razmatramo slučaj $0 \le \tau < \infty$. Ako (2.4) važi, onda je

$$P(M_n \le u_n) = F^n(u_n) = \left(1 - \bar{F}(u_n)\right)^n = \left(1 - \frac{\tau}{n} + o\left(\frac{1}{n}\right)\right)^n,$$

što implicira (2.5).
Obrnuto, ako (2.5) važi, onda $\bar{F}(u_n) \to 0$. Inače bi $\bar{F}\left(u_{n_k}\right)$ bilo ograničeno izvan okoline tačke 0 za neki podniz (n_k). Tada bi $P\left(M_{n_k} \le u_{n_k}\right) = \left(1 - \bar{F}\left(u_{n_k}\right)\right)^{n_k}$ impliciralo da $P\left(M_{n_k} \le u_{n_k}\right) \to 0$. Zbog (2.5) sledi

$$-n \ln\left(1 - \bar{F}(u_n)\right) \to \tau.$$

Pošto je $-\ln(1-x) \sim x$ kad $x \to 0$, odatle sledi $n\bar{F}(u_n) = \tau + o(1)$ što daje (2.4).
Pretpostavimo da je $\tau = \infty$ i da (2.4) važi, a (2.5) ne. Tada mora da postoji podniz (n_k) tako da $P\left(M_{n_k} \le u_{n_k}\right) \to exp\{-\tau'\}$ kad $k \to \infty$ za neko $\tau' < \infty$. Međutim, onda iz (2.5) sledi (2.4) tako da $n_k\bar{F}\left(u_{n_k}\right) \to \tau' < \infty$, što je kontradiktorno sa (2.4) gde je $\tau = \infty$. Analogno, (2.5) implicira (2.4) za $\tau = \infty$. □

Posledica 2.1.2 Pretpostavimo da je $x_F < \infty$ i

$$\bar{F}(x_F-) = F(x_F) - F(x_F-) > 0.$$

Onda je za svaki niz (u_n) takav da je

$$P(M_n \le u_n) \to \rho,$$

važe ili $\rho = 0$ ili $\rho = 1$. □

Dokaz Pošto je $0 \le \rho \le 1$, možemo zapisati $\rho = exp\{-\tau\}$ gde je $0 \le \tau \le \infty$. Na osnovu Tvrđenja 2.1.1 znamo $n\bar{F}(u_n) \to \tau$, kad $n \to \infty$. Ako je $u_n < x_F$ za beskonačno mnogo n-ova dobija se $\bar{F}(u_n) \ge \bar{F}(x_F-) > 0$ za takve n, stoga je $\tau = \infty$.
Druga mogućnost je da je $u_n \ge x_F$ za sve dovoljno velike n-ove, odakle je $n\bar{F}(u_n) = 0$, stoga je $\tau = 0$. Prema tome, $\tau = \infty$ ili 0, odakle sledi da je $\rho = 0$ ili 1. □

Ova posledica pokazuje da za funkciju raspodele sa skokom u krajnjoj desnoj tački, ne postoji nedegenerativna granična raspodela za M_n, bez obzira na normalizaciju.

Teorema 2.1.3 Neka je F funkcija raspodele sa krajnjom desnom tačkom $x_F \le \infty$ i neka je $\tau \in (0, \infty)$. Postoji niz (u_n) koji zadovoljava $n\bar{F}(u_n) \to \tau$ ako i samo ako je

$$\lim_{x \uparrow x_F} \frac{\bar{F}(x)}{\bar{F}(x^-)} = 1 \qquad (2.6)$$

i $F(x_F-) = 1$. □

Ova teorema se primenjuje na diskretne raspodele sa beskonačnim krajnjim desnim tačkama.

Ako visine skokova funkcije raspodele opadaju dovoljno sporo, onda nedegenerativna granična raspodela za maksimume ne postoji.
Na primer, ako X ima celobrojnu vrednost, onda (2.6) postaje $\bar{F}(n)/\bar{F}(n-1) \to 1$ kad $n \to \infty$.
Ova razmatranja pokazuju da postoji neka vrsta komplikovanog asimptotskog ponašanja (M_n). Diskretnost raspodele može da spreči konvergenciju maksimuma i podstakne oscilatorno ponašanje. Uprkos tome, u ovakvoj situaciji je često moguće naći niz celih brojeva (c_n) tako da svaki podniz $(M_n - c_n)$ sadrži slabo konvergentan podniz.

Primer 2.1.4 (Poissonova raspodela)

$$P(X = k) = e^{-\lambda}\lambda^k/k! \quad k \in \mathbb{N}_0, \qquad \lambda > 0$$

Tada

$$\frac{\bar{F}(k)}{\bar{F}(k-1)} = 1 - \frac{F(k) - F(k-1)}{\bar{F}(k-1)} = 1 - \frac{\lambda^k}{k!}\left(\sum_{r=k}^{\infty}\frac{\lambda^r}{r!}\right)^{-1} = 1 - \left(1 + \sum_{r=k+1}^{\infty}\frac{k!}{r!}\lambda^{r-k}\right)^{-1}$$

Poslednja suma se može oceniti sa

$$\sum_{s=1}^{\infty}\frac{\lambda^s}{(k+1)(k+2)\dots(k+s)} \leq \sum_{s=1}^{\infty}\left(\frac{\lambda}{k}\right)^s = \frac{\lambda/k}{1-\lambda/k}, \quad k > \lambda,$$

što teži ka nuli, kad $k \to \infty$, tako da $\bar{F}(k)/\bar{F}(k-1) \to 0$. Teorema 2.1.3 pokazuje da ne postoji nedegenerativna granična raspodela za maksimume i da ne postoji granica oblika $P(M_n \leq u_n) \to \rho \in (0,1)$ ni za jedan niz konstanti (u_n).

Primer 2.1.5 (Geometrijska raspodela)

$$P(X = k) = p(1-p)^k, \qquad k \in \mathbb{N}_0, \qquad 0 < p < 1$$

S obzirom na to da je $\bar{F}(k) = (1-p)^{k+1}$, znamo da je $\frac{\bar{F}(k)}{\bar{F}(k-1)} = 1 - p \in (0,1)$.
Stoga ne postoji granica $P(M_n \leq u_n) \to \rho$ osim za $\rho = 0$ ili 1.
Maksimumi iid slučajnih promenljivih sa geometrijskom raspodelom igraju veoma važnu ulogu u proučavanju dužine najdužeg niza uspeha u slučajnom hodu. □

Primer 2.1.6 (Negativna binomna raspodela)

$$P(X = k) = \binom{v+k-1}{k} p^v (1-p)^k, \qquad k \in \mathbb{N}_0, \qquad 0 < p < 1, v > 0.$$

Za $v \in \mathbb{N}$ negativna binomna raspodela generalizuje geometrijsku raspodelu u smislu: geometrijska raspodela modelira vreme čekanja za prvi uspeh u nizu nezavisnih pokušaja, a negativna binomna raspodela modelira vreme čekanja da se dogodi v-ti uspeh.
Koristeći Stirlingovu formulu dobijamo

$$\mathrm{Lim}_{k\to\infty}\frac{\bar{F}(k)}{\bar{F}(k-1)} = 1 - p \in (0,1),$$

to jest, ne postoji granica $P(M_n \leq u_n) \to \rho$ osim za $\rho = 0$ ili 1.

2.2 Slaba konvergencija maksimuma prilikom linearnih transformacija

Ovde izučavamo karakteristike mogućih graničnih zakona za maksimume M_n iid niza (X_n) prilikom pozitivnih linearnih transformacija (2.2). Ovaj problem ekstremne vrednosti se može razmatrati analogno centralnom graničnom problemu. Shodno tome, glavni delovi Poglavlja 2.2 i 2.3 su slični Poglavlju 1.2, te je neophodno uporediti odgovarajuće rezultate.
U ovom poglavlju dajemo odgovor na pitanje o mogućim (nedegenerativnim) graničnim zakonima za maksimume M_n kada su oni normalizovani i centrirani.
Ovo pitanje je usko povezano sa sledećim:
Koje raspodele zadovoljavaju identitet

$$max(X_1, \dots X_n) \stackrel{d}{=} c_n X + d_n \tag{2.7}$$

za odgovarajuće konstante $c_n > 0,\ d_n \in \mathbb{R}$ za svako $n \geq 2$?
Drugim rečima, koje klase raspodela F su zatvorene za maksimume?

Definicija 2.2.1 (Max-stabilne raspodele)

Nedegenerativna slučajna promenljiva X (odgovarajuća raspodela ili funkcija gustine) se naziva max-stabilnom ukoliko zadovoljava (2.7) za iid $X, X_1, \dots, X_n$ za odgovarajuće konstante $c_n > 0, d_n \in \mathbb{R}$ i za svako $n \geq 2$. □

Ako pretpostavimo da je (X_n) niz iid max-stabilnih slučajnih promenljivih, onda se (2.7) može zapisati kao

$$c_n^{-1}(M_n - d_n) \stackrel{d}{=} X. \tag{2.8}$$

Zaključujemo da je svaka max-stabilna raspodela granična raspodela za maksimume iid slučajnih promenljivih. Štaviše, max-stabilne raspodele su jedini granični zakoni za normalizovane maksimume.

Teorema 2.2.2 (Granične osobine max-stabilnih zakona)
Klasa max-stabilnih raspodela se podudara sa klasom svih mogućih (nedegenerativnih) graničnih zakona za (normalizovane) maksimume iid slučajnih promenljivih. □

Dokaz Ostaje da dokažemo da je granična raspodela linearno transformisanih maksimuma max-stabilna. Pretpostavimo da je za odgovarajuće normirajuće konstante,

$$\lim_{n\to\infty} F^n(c_n x + d_n) = H(x), \qquad x \in \mathbb{R}$$

za neku nedegenerativnu funkciju raspodele H.
Ovde uzimamo da su moguće granične funkcije raspodele H neprekidne funkcije na $\mathbb{R}$.
Onda za svako $k \in \mathbb{N}$

$$\lim_{n\to\infty} F^{nk}(c_n x + d_n) = (\lim_{n\to\infty} F^n(c_n x + d_n))^k = H^k(x), \quad x \in \mathbb{R}.$$

Dalje,

$$\lim_{n\to\infty} F^{nk}(c_{nk}x + d_{nk}) = H(x), \qquad x \in \mathbb{R}$$

Zbog konvergencije (Teorema D1) postoje konstante $\tilde{c}_k > 0$ i $\tilde{d}_k \in \mathbb{R}$ tako da

$$\lim_{n\to\infty} \frac{c_{nk}}{c_n} = \tilde{c}_k \text{ i } \lim_{n\to\infty} \frac{d_{nk} - d_n}{c_n} = \tilde{d}_k,$$

i za iid slučajne promenljive $Y, Y_1, \dots, Y_k$ sa funkcijom raspodele H,

$max(Y_1, \dots, Y_k) \overset{d}{=} \tilde{c}_k Y_1 + \tilde{d}_k.$ □

Teorema 2.2.3 (Fisher-Tippetova teorema, granični zakoni za maksimume)

Neka je (X_n) niz iid slučajnih promenljivih. Ako postoje normirajuće konstante $c_n > 0$, $d_n \in \mathbb{R}$ i neka nedegenerativna funkcija raspodele H tako da

$$c_n^{-1}(M_n - d_n) \overset{d}{\to} H, \tag{2.9}$$

onda H pripada jednom od sledeća tri tipa funkcije raspodele:

Frechet: $\phi_\alpha(x) = \begin{cases} 0, & x \leq 0 \\ exp\{-x^{-\alpha}\}, & x > 0 \end{cases} \quad \alpha > 0.$

Weibull: $\psi_\alpha(x) = \begin{cases} exp\{-(-x)^\alpha\}, & x \leq 0 \\ 1, & x > 0 \end{cases} \quad \alpha > 0.$

Gumbel: $\Lambda(x) = exp\{-e^{-x}\}, \ x \in \mathbb{R}.$

Skica dokaza Iako je potpun dokaz tehničke prirode, mi ćemo pokazati kako se pojavljuju tri granična tipa. Zaista, (2.9) implicira da za svako $t > 0$

$$F^{[nt]}\left(c_{[nt]}x + d_{[nt]}\right) \to H(x), \qquad x \in \mathbb{R},$$

gde $[\cdot]$ označava ceo deo. Međutim,

$$F^{[nt]}(c_n x + d_n) = \left(F^n(c_n x + d_n)\right)^{[nt]/n} \to H^t(x),$$

pa je prema Teoremi D1 postoje funkcije $\gamma(t) > 0, \delta(t) \in \mathbb{R}$ koje zadovoljavaju

$$\lim_{n\to\infty} \frac{c_n}{c_{[nt]}} = \gamma(t), \lim_{n\to\infty} \frac{d_n - d_{[nt]}}{c_{[nt]}} = \delta(t), \ \ t > 0 \ \text{ i}$$

$$H^t(x) = H\left(\gamma(t)x + \delta(t)\right). \tag{2.10}$$

Iz (2.10) zaključujemo da je za $s, t > 0$

$$\gamma(st) = \gamma(s)\gamma(t), \ \ \delta(st) = \gamma(t)\delta(s) + \delta(t) \tag{2.11}$$

Rešenja funkcionalnih jednačina (2.10) i (2.11) vode do tri tipa $\Lambda, \phi_\alpha, \psi_\alpha$. □

Definicija 2.2.6 (Raspodela ekstremne vrednosti i ekstremne slučajne promenljive) Funkcije raspodele $\phi_\alpha, \psi_\alpha \ i \ \Lambda$ se nazivaju standardnim raspodelama ekstremnih vrednosti, a odgovarajuće slučajne promenljive standardnim ekstremim slučajnim promenljivima. Funkcije raspodele tipova ϕ_α, ψ_α i Λ su raspodele ekstremnih vrednosti, a odgovarajuće slučajne promenljive ekstremne slučajne promenljive. □

Prema Teoremi 2.2.2, raspodele ekstremnih vrednosti su baš max-stabilne raspodele. Prema tome, ako je X ekstremna slučajna promenljiva, ona zadovoljava (2.8). Specijalno, tri slučaja iz Teoreme 2.2.3:

$$\text{Frechet: } M_n \stackrel{d}{=} n^{1/\alpha} X$$

$$\text{Weibull: } M_n \stackrel{d}{=} n^{-1/\alpha} X$$

$$\text{Gumbel: } M_n \stackrel{d}{=} X + \ln n$$

Primer 2.2.7 (Maksimumi eksponencijalnih slučajnih promenljivih)
Neka je (X_i) niz iid standardnih eksponencijalnih slučajnih promenljivih. Tada
$P(M_n - \ln n \le x) = \left(P(X \le x + \ln n)\right)^n = (1 - n^{-1}e^{-x})^n \to exp\{-e^{-x}\} = \Lambda(x), x \in \mathbb{R}$.
Poređenja radi, za Gumbelove slučajne promenljive X_i je:
$P(M_n - \ln n \le x) = \Lambda(x), x \in \mathbb{R}$ □

2.3 Maksimalni domeni atrakcije i normirajuće konstante

U ovom poglavlju ispitujemo koji uslovi za funkciju raspodele F impliciraju da normalizovani maksimumi M_n slabo konvergiraju ka H za datu raspodelu ekstremnih vrednosti H. Ova problematika je usko povezana sa sledećom:
Kako možemo da izaberemo normirajuće konstante $c_n > 0$ i $d_n \in \mathbb{R}$ tako da

$$c_n^{-1}(M_n - d_n) \stackrel{d}{\to} H? \quad (2.12)$$

Definicija 2.3.1 (Maksimalni domen atrakcije)
Kažemo da slučajna promenljiva X (funkcija raspodele F za X, raspodela za X) pripada maksimalnom domenu atrakcije raspodele ekstremne vrednosti H ako postoje konstante $c_n > 0, d_n \in \mathbb{R}$ tako da važi (2.12). Pišemo $X \in MDA(H) \big(F \in MDA(H)\big)$. □

Tvrđenje 2.3.2 (Karakterizacija $MDA(H)$)
Funkcija raspodele F pripada maksimalnom domenu atrakcije raspodele ekstremne vrednosti H sa normirajućim konstantama $c_n > 0, d_n \in \mathbb{R}$ ako i samo ako

$$\lim_{n\to\infty} n\, \bar{F}(c_n x + d_n) = -\ln H(x), x \in \mathbb{R}.$$

Kada je $H(x) = 0$, granica se tumači kao ∞. □

Definicija 2.3.3 (Ekvivalencija repa)

Dve funkcije raspodele $F\ i\ G$ se nazivaju ekvivalentnima u repu ako imaju istu krajnju desnu tačku, to jest ako je $x_F = x_G$ i

$$\lim_{x\uparrow x_F} \bar{F}(x)/\bar{G}(x) = c$$

za neku konstantu $0 < c < \infty$. □

Definicija 2.3.4 (Uopštena inverzna monotona funkcija)
Pretpostavimo da je h neopadajuća funkcija na $\mathbb{R}$.
Uopštena inverzna funkcija h je definisana kao $h^{\leftarrow}(t) = inf\{x \in \mathbb{R}: h(x) \geq t\}$
(uzimamo da je infimum praznog skupa ∞). □

Definicija 2.3.5 (Funkcija kvantila)
Uopštena inverzna funkcija raspodele F

$$F^{\leftarrow}(t) = inf\{x \in \mathbb{R}: F(x) \geq t\},\ \ 0 < t < 1$$

se naziva funkcijom kvantila funkcije raspodele F. Veličina $x_t = F^{\leftarrow}(t)$ se naziva t-kvantil od F. □

2.3.1 Maksimalni domen atrakcije Freschetove raspodele

$$\boldsymbol{\phi_\alpha(x) = exp\{-x^{-\alpha}\}}$$

U ovom delu ćemo razmatrati slučaj $\alpha > 0$.
Iz Tejlorove ekspanzije

$$1 - \phi_\alpha(x) = 1 - exp\{-x^{-\alpha}\} \sim x^{-\alpha}, x \to \infty$$

vidimo da rep od ϕ_α opada polinomijalnom brzinom. Dalje nas zanima koliko daleko možemo da se udaljimo od stepenog repa, a da i dalje ostanemo u $MDA(\phi_\alpha)$.
Pokazaćemo da se maksimalni domen atrakcije od ϕ_α sastoji od funkcija raspodele F čiji desni rep regularno varira sa indeksom $-\alpha$. Za $F \in MDA(\phi_\alpha)$ konstante d_n mogu biti izabrane da budu nule, a za c_n funkcije kvantila, to jest

$$\begin{aligned} c_n = F^{\leftarrow}(1 - n^{-1}) &= inf\{x \in \mathbb{R}: F(x) \geq 1 - n^{-1}\} \\ &= inf\{x \in \mathbb{R}: (1/\bar{F})(x) \geq n\} \\ &= (1/\bar{F})^{\leftarrow}(n). \end{aligned} \tag{2.13}$$

Teorema 2.3.6 (Maksimalni domen atrakcije od ϕ_α)
Funkcija raspodele F pripada maksimalnom domenu atrakcije od ϕ_α, $\alpha > 0$, ako i samo ako $\bar{F}(x) = x^{-\alpha}L(X)$ za neku sporo varirajuću funkciju L.
Ako $F \in MDA(\phi_\alpha)$ onda

$$c_n^{-1} M_n \xrightarrow{d} \phi_\alpha, \tag{2.14}$$

gde normirajuće konstante c_n mogu biti izabrane kao u (2.13).

Odavde sledi da svaka $F \in MDA(\phi_\alpha)$ ima beskonačnu krajnju desnu tačku $x_F = \infty$. Dalje, normirajuće konstante c_n formiraju regularno varirajući niz, tj. $c_n = \mathrm{n}^{1/\alpha}\,\mathrm{L(n)}$ za neku sporo varirajuću funkciju L.

Dokaz

Neka je $\bar{F} \in \mathcal{R}_{-\alpha}$ za $\alpha > 0$. Zbog izbora za c_n i regularne varijacije znamo,

$$\bar{F}(c_n) \sim n^{-1}, \;\; n \to \infty, \tag{2.15}$$

pa $\bar{F}(c_n) \to 0$, što daje da $c_n \to \infty$. Za $x > 0$,

$$n\bar{F}(c_n x) \sim \frac{\bar{F}(c_n x)}{\bar{F}(c_n)} \to x^{-\alpha}, n \to \infty.$$

Za $x < 0$ odmah znamo $F^n(c_n x) \leq F^n(0) \to 0$, pošto regularna varijacija zahteva $F(0) < 1$, a na osnovu Tvrđenja 2.3.2 znamo da $F \in MDA(\phi_\alpha)$.

Obrnuto, pretpostavimo da $\lim_{n\to\infty} F^n(c_n x + d_n) = \phi_\alpha(x)$ za svako $x > 0$ i odgovarajuće $c_n > 0$, $d_n \in \mathbb{R}$. To nas dovodi do

$$\lim_{n\to\infty} F^n\big(c_{[ns]}x + d_{[ns]}\big) = \phi_\alpha^{1/s}(x) = \phi_\alpha\big(s^{1/\alpha}\, x\big), \;\; s > 0, x > 0.$$

Zbog konvergencije (Teorema D1)

$$c_{[ns]}/c_n \to s^{1/\alpha} \;\text{ i }\; \big(d_{[ns]} - d_n\big)/c_n \to 0.$$

Prema tome, (c_n) je regularno varirajući niz u smislu Definicije D5, posebno $c_n \to \infty$. Pretpostavimo prvo da je $d_n = 0$, onda $n\bar{F}(c_n x) \to x^{-\alpha}$ tako da $\bar{F} \in \mathcal{R}_{-\alpha}$ zbog Tvrđenja D2(a). Slučaj $d_n \neq 0$ je teži jer se mora pokazati $d_n/c_n \to 0$. Ako to važi, može se ponoviti gore navedeni argument kad se zameni d_n sa 0. □

Dobili smo odgovor na naše pitanje:

$$F \in MDA(\phi_\alpha) \Leftrightarrow \bar{F} \in \mathcal{R}_{-\alpha}.$$

Posledica 2.3.7 (Von Miesesov uslov)

Neka je F apsolutno neprekidna funkcija raspodele sa gustinom f koja zadovoljava

$$\lim_{x\to\infty} \frac{xf(x)}{\bar{F}(x)} = \alpha > 0. \tag{2.16}$$

Tada $F \in MDA(\phi_\alpha)$. □

Tvrđenje 2.3.8 (Osobina zatvorenosti za $MDA(\phi_\alpha)$)

Neka su $F\ i\ G$ funkcije raspodele i pretpostavimo da $F \in MDA(\phi_\alpha)$ sa normirajućim konstantama $c_n > 0$, to jest

$$\lim_{n\to\infty} F^n(c_n x) = \phi_\alpha(x), x > 0. \tag{2.17}$$

Tada

$$\lim_{n\to\infty} G^n(c_n x) = \phi_\alpha(cx), x > 0,$$

za neko $c > 0$ ako i samo ako su $F\ i\ G$ ekvivalentni u repu sa $\lim_{n\to\infty} \bar{F}(x)/\bar{G}(x) = c^\alpha$. □

Dokaz dovoljnog uslova
Pretpostavimo da je $\bar{F}(x) \sim q\bar{G}(x)$ kad $x \to \infty$ za neko $q > 0$. Zbog Tvrđenja 2.3.2, granična relacija (2.15) je ekvivalentna sa

$$\lim_{n\to\infty} n\bar{F}(c_n x) = x^{-\alpha}$$

za svako $x > 0$. Za takvo x, $c_n x \to \infty$ kad $n \to \infty$ pa je, zbog ekvivalencije repa,

$$n\bar{G}(c_n x) \sim nq^{-1}\bar{F}(c_n x) \to q^{-1}x^{-\alpha} \ ,$$

to jest, opet zbog Tvrđenja 2.3.2,

$$\lim_{n\to\infty} G^n(c_n x) = exp\left\{-\left(q^{1/\alpha}x\right)^{-\alpha}\right\} = \phi_\alpha\left(q^{1/\alpha}x\right).$$

Uzmimo da je $c = q^{1/\alpha}$. □

Dakle, $MDA(\phi_\alpha)$ se sastoji od funkcija raspodele koje zadovoljavaju Von Misesov uslov (2.14) i od njihovih rep-ekvivalentnih funkcija raspodele.

U Tabeli 1 navodimo primere raspodela koje pripadaju maksimalnom domenu atrakcije Freschetove raspodele.

Raspodela	**Rep $\bar{F}$ ili gustina f**	**Parametar**	**Normirajuće konstante**
Pareto	$\bar{F}(x) = (\frac{k}{k+x})^\alpha$	$\alpha, k > 0$	$c_n = (Kn)^{1/\alpha}$
Cauchy	$f(x) = \left(\pi(1+x^2)\right)^{-1}$	$x \in \mathbb{R}$	$c_n = n/\pi$
Burr	$\bar{F}(x) = (\frac{k}{k+x^\tau})^\alpha$	$\alpha, k, \tau > 0$	$c_n = (Kn)^{1/\alpha}$

Tabela 1 Maksimalni domen atrakcije Frechetove raspodele

2.3.2 Maksimalni domen atrakcije Weibullove raspodele

$$\boldsymbol{\psi_\alpha(x) = exp\{-(-x)^\alpha\}}$$

Ovde ćemo okarakterisati maksimalan domen atrakcije od ψ_α za $\alpha > 0$. Važna, ali ne tako očigledna činjenica jeste da sve funkcije raspodele $F \in MDA(\psi_\alpha)$ imaju konačnu krajnju desnu tačku x_F.
$\psi_\alpha \ i \ \phi_\alpha$ su usko povezani, to jest

$$\psi_\alpha(-x^{-1}) = \phi_\alpha(x), \ x > 0.$$

Prema tome, za očekivati je da i $MDA(\psi_\alpha)$ i $MDA(\phi_\alpha)$ budu usko povezani.

Teorema 2.3.9 (Maksimalan domen atrakcije od (ψ_α))
Funkcija raspodele F pripada maksimalnom domenu atrakcije od $\psi_\alpha, \alpha > 0$, ako i samo ako $x_F < \infty$ i $\bar{F}(x_F - x^{-1}) = x^{-\alpha}L(X)$ za neku sporo varirajuću funkciju L.
Ako $F \in MDA(\psi_\alpha)$, onda

$$c_n^{-1}(M_n - x_F) \xrightarrow{d} \psi_\alpha, \tag{2.18}$$

gde normirajuće konstante c_n mogu biti izabrane kao $c_n = x_F - F^{\leftarrow}(1 - n^{-1})$ i $d_n = x_F$.

Skica dokaza Dokazujemo dovoljan uslov tako što ćemo koristiti vezu između ϕ_α i ψ_α.
Pretpostavimo da $x_F < \infty$ i $\bar{F}(x_F - x^{-1}) = x^{-\alpha}L(x)$ i definišemo

$$F_*(x) = F(x_F - x^{-1}), x > 0. \tag{2.19}$$

Tada $F_* \in \mathcal{R}_{-\alpha}$, pa prema Teoremi 2.3.6, $F_* \in MDA(\phi_\alpha)$ sa normirajućim konstantama $c_n^* = F_*^{\leftarrow}(1 - n^{-1})$ i $d_n^* = 0$.

Tada $F_* \in MDA(\phi_\alpha)$ implicira da za $x > 0$

$$F_*^n(c_n^* x) \to \phi_\alpha(x).$$

To jest, $F^n(x_F - (c_n^* x)^{-1}) \to exp\{-x^{-\alpha}\}$.

Uvodimo smenu $x = -y^{-1}$, pa

$$F^n(x_F + y/c_n^*) \to exp\{-(-y)^\alpha\},\ \ y < 0. \tag{2.20}$$

Konačno,

$c_n^* = F_*^{\leftarrow}(1 - n^{-1})$=

$$= inf\{x \in \mathbb{R}: F(x_F - x^{-1}) \geq 1 - n^{-1}\} = inf\{(x_F - u)^{-1}: F(u) \geq 1 - n^{-1}\} =$$

$$= (x_F - inf\{u: F(u) \geq 1 - n^{-1}\})^{-1} = \left(x_F - F^{\leftarrow}(1 - n^{-1})\right)^{-1}$$

što je kraj dokaza zbog (2.18) □
Posledica prethodno navedenog je

$$F \in MDA(\psi_\alpha) \Leftrightarrow x_F < \infty,\ \ \bar{F}(x_F - x^{-1}) \in \mathcal{R}_{-\alpha}.$$

$MDA(\psi_\alpha)$ se sastoji od funkcija raspodele F koje imaju nosač ograničen s desne strane. Ove funkcije raspodele nisu nužno najbolji izbor za modeliranje ekstremnih događaja u osiguranju i finansijama, upravo zbog $x_F < \infty$. Iako u praksi postoji gornja granica, nećemo uvek želeti da uključimo ovaj dodatni parametar x_F u model. Često preferiramo raspodele sa $x_F = \infty$, s obzirom na to da one dozvoljavaju proizvoljno velike vrednosti u uzorku. Takve raspodele uglavnom pripadaju $MDA(\phi_\alpha)$ ili $MDA(\Lambda)$.

Posledica 2.3.10 (Von Misesov uslov)
Neka je F apsolutno neprekidna funkcija raspodele sa gustinom f koja je pozitivna na nekom konačnom intervalu (z, x_F). Ako

$$\lim_{x \uparrow x_F} \frac{(x_F - x)f(x)}{\bar{F}(x)} = \alpha > 0, \tag{2.21}$$

onda $F \in MDA(\psi_\alpha)$. □

Ako primenimo transformaciju (2.17) i Tvrđenje 2.3.8 može biti modifikovano na sledeći način:

Tvrđenje 2.3.11 (Osobina zatvorenosti za $MDA(\psi_\alpha)$)
Neka su F i G funkcije raspodele sa krajnjim desnim tačkama $x_F = x_G < \infty$ i pretpostavimo da $F \in MDA(\psi_\alpha)$ sa normirajućim konstantama $c_n > 0$, to jest

$$\lim_{n\to\infty} F^n(c_n x + x_F) = \psi_\alpha(x), \quad x < 0.$$

Tada

$$\lim_{n\to\infty} G^n(c_n x + x_F) = \psi_\alpha(cx), \quad x < 0,$$

za neko $c > 0$ ako i samo ako za F i G važi ekvivalencija repa sa

$$\lim_{x\uparrow x_F} \bar{F}(x)/\bar{G}(x) = c^{-\alpha}.$$ □

Zaključujemo, $MDA(\psi_\alpha)$ se sastoji od funkcija raspodele koje zadovoljavaju Von Misesov uslov (2.19) i od rep-ekvivalentnih funkcija raspodele.

U Tabeli 2 navodimo primere raspodela koje pripadaju maksimalnom domenu atrakcije Weibullove raspodele.

Raspodela	**Rep $\bar{F}$ ili gustina f**	**Parametar**	**Normirajuće konstante**
Unifirmna	$f(x) = 1$	$x \in (0,1)$	$c_n = n^{-1}, d_n = 1$
Stepene funkcije po x_F	$\bar{F}(x) = K(x_F - x)^\alpha$, $x_F - K^{-1/\alpha} \le x \le x_F$	$K, \alpha > 0$	$c_n = (Kn)^{-1/\alpha}$, $d_n = x_F$

Tabela 2 Maksimalni domen atrakcije Weibullove raspodele

2.3.3 Maksimalni domen atrakcije Gumbelove raspodele

$$\Lambda(\boldsymbol{x}) = \boldsymbol{exp}\{-\boldsymbol{exp}\{-\boldsymbol{x}\}\}$$

Von Misesove funkcije
Maksimalan domen atrakcije Gumbelove raspodele Λ pokriva širok spektar funkcija raspodele F. Iako ne postoji direktna veza sa regularnim varijacijama kao što je to slučaj kod maksimalnog domena atrakcije Frechetove i Weibullove raspodele, daćemo ekstenzije regularne varijacije koje omogućavaju potpunu karakterizaciju $MDA(\Lambda)$.
Tejlorova ekspanzije daje

$$1 - \Lambda(x) \sim e^{-x}, \quad x \to \infty,$$

pa $\overline{\Lambda}(x)$ eksponencijalno opada ka nuli. Ono što nas dalje interesuje je koliko daleko možemo da se udaljimo od eksponencijalnog repa, a da ostanemo u $MDA(\Lambda)$.
$MDA(\Lambda)$ obuhvata funkcije raspodele sa različitim repovima, koji idu od „umereno teških" (lognormalna raspodela), do „lakih" (normalna raspodela). Takođe, oba slučaja $x_F < \infty$ i $x_F = \infty$ su moguća.
Pre nego što damo odgovor na postavljeno pitanje, ograničavamo se na apsolutno neprekidne $F \in MDA(\Lambda)$.

Definicija 2.3.12 (Von Misesova funkcija)
Neka je F funkcija raspodele koja je neprekidna u krajnjoj desnoj tački $x_F \leq \infty$. Pretpostavimo da postoji neko $z < x_F$ tako da F ima reprezentaciju

$$\bar{F}(x) = c \cdot exp\left\{-\int_z^x \frac{1}{a(t)} dt\right\}, \quad z < x < x_F \tag{2.22}$$

gde je c neka pozitivna konstanta, $a(\cdot)$ je pozitivna i apsolutno neprekidna funkcija (u odnosu na Lebegovu meru) sa gustinom a' i $\lim\limits_{x \uparrow x_F} a'(x) = 0$.

Onda se F naziva von Misesova funkcija, a funkcija $a(\cdot)$ pomoćnom funkcijom od F. □

Tvrđenje 2.3.13 (Osobine von Misesovih funkcija)
Svaka von Misesova funkcija je apsolutno neprekidna na (x, x_F) sa pozitivnom gustinom f. Možemo izabrati pomoćnu funkciju $a(x) = \bar{F}(x)/f(x)$.
Štaviše, važe osobine:

(a) Ako $x_F = \infty$, onda $\bar{F} \in \mathcal{R}_{-\infty}$ i

$$\lim_{x\to\infty} \frac{xf(x)}{\bar{F}(x)} = \infty. \tag{2.23}$$

(b) Ako $x_F < \infty$, onda $\bar{F}(x_F - x^{-1}) \in \mathcal{R}_{-\infty}$ i

$$\lim_{x \uparrow x_F} \frac{(x_F - x)f(x)}{\bar{F}(x)} = \infty. \tag{2.24}$$

□

Primetimo da iz (2.23) sledi da $\lim\limits_{x\to\infty} x^{-1}a(x) = 0$, a iz (2.24) da je

$a(x) = o(x_F - x) = o(1)$ kad $x \uparrow x_F$.

Takođe, $a^{-1}(x) = f(x)/\bar{F}(x)$ je stopa rizika od F.

Dokaz
Iz (2.22) dobijamo

$$\frac{d}{dx}(-\ln \bar{F}(x)) = \frac{f(x)}{\bar{F}(x)} = \frac{1}{a(x)}, \qquad z < x < x_F$$

(a) Pošto $a'(x) \to 0$ kad $x \to \infty$ Cesaro sredina a' takođe konvergira:

$$\lim_{x\to\infty} \frac{a(x)}{x} = \lim_{x\to\infty} \frac{1}{x} \int_z^x a'(t)\, dt = 0. \tag{2.25}$$

Odavde sledi (2.23). $\bar{F} \in \mathcal{R}_{-\infty}$ sledi iz primene Teoreme D4 (b).

(b) Znamo

$$\lim_{x\uparrow x_F}\frac{a(x)}{x_F-x}=\lim_{x\uparrow x_F}-\int_x^{x_F}\frac{a'(t)}{x_F-x}dt=\lim_{s\downarrow 0}\frac{1}{s}\int_0^s a'(x_F-t)dt$$

zbog smene promenljivih. S obzirom na to da je $a'(x_F-t)\to 0$, kad $t\downarrow 0$, poslednji limes teži nuli, odakle sledi (2.24). $\bar{F}(x_F-x^{-1})\in\mathcal{R}_{-\infty}$ sledi kao malopre. □

Tvrđenje 2.3.14 (Von Misesove funkcije i $MDA(\Lambda)$)
Pretpostavimo da je funkcija raspodele F von Misesova funkcija. Tada $F\in MDA(\Lambda)$. Moguć izbor normirajućih konstanti je

$$d_n=F^{\leftarrow}(1-n^{-1}) \quad \text{i} \quad c_n=a(d_n), \tag{2.26}$$

gde je a pomoćna funkcija od F. □

Dokaz
(2.22) implicira da je za $t\in\mathbb{R}\ i\ x$ koje je dovoljno blizu x_F

$$\frac{\bar{F}(x+ta(x))}{\bar{F}(x)}=exp\left\{-\int_x^{x+ta(x)}\frac{1}{a(u)}du\right\}.$$

Uvodimo smenu $v=(u-x)/a(x)$ i dobijamo

$$\frac{\bar{F}(x+ta(x))}{\bar{F}(x)}=exp\left\{-\int_0^t\frac{a(x)}{a(x+va(x))}dv\right\}. \tag{2.27}$$

Pokazaćemo da podintegralna veličina konvergira uniformno ka 1. Za dato $\varepsilon>0\ i\ x\geq x_0(\varepsilon)$,

$$\left|a(x+va(x))-a(x)\right|=\left|\int_x^{x+va(x)}a'(s)ds\right|\leq\varepsilon|v|a(x)\leq\varepsilon|t|a(x)$$

gde $a'(x)\to 0$ kad $x\uparrow x_F$, što za $x\geq x_0(\varepsilon)$ implicira

$$\left|\frac{a(x+va(x))}{a(x)}-1\right|\leq\varepsilon|t|\ .$$

Desna strana nejednakosti može da bude zanemarljivo mala, pa

$$\lim_{x\uparrow x_F}\frac{a(x)}{a(x+va(x))}=1, \tag{2.28}$$

uniformno na ograničenim v intervalima. Ovo, zajedno sa (2.27) daje

$$\lim_{x\uparrow x_F}\frac{\bar{F}(x+ta(x))}{\bar{F}(x)}=e^{-t}, \tag{2.29}$$

uniformno na ograničenim t intervalima. Sada za normirajuće konstante d_n biramo $d_n=(1/\bar{F})^{\leftarrow}(n)$ i $c_n=a(d_n)$. Tada (2.29) implicira

$$\lim_{n\to\infty}n\bar{F}(d_n+tc_n)=e^{-t}=-\ln\Lambda(t)\,,t\in\mathbb{R}.$$

Sada primenimo Tvrđenje 2.3.2 i dobijamo da $F\in MDA(\Lambda)$. □

Raspodela	Rep $\bar{F}$ ili gustina f	Parametar	Normirajuće konstante
Gama	$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$	$x > 0$, $\alpha, \beta > 0$	$c_n = \beta^{-1}$, $d_n = \beta^{-1}(\ln n + (\alpha - 1) ln ln n - ln\Gamma(\alpha))$
Normalna	$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$	$x \in \mathbb{R}$	$c_n = (2ln n)^{-1/2}$, $d_n = \sqrt{2ln n} - \frac{\ln(4\pi) + ln ln n}{2(2ln n)^{1/2}}$
Lognormalna	$f(x) = \frac{1}{\sqrt{2\pi}\sigma x} e^{-(ln x - \mu)^2/2\sigma^2}$	$x > 0$, $\mu \in \mathbb{R}$, $\sigma > 0$	$c_n = \sigma(2ln n)^{-1/2} d_n$, $d_n = exp\left\{\mu + \sigma(\sqrt{2ln n} - \frac{\ln(4\pi) + ln ln n}{2(2ln n)^{1/2}})\right\}$

Tabela 3 Maksimalni domen atrakcije Gumbelove raspodele

Karakterizacija $MDA(\Lambda)$

Von Misesove funkcije ne daju kompletnu karakterizaciju maksimalnog domena atrakcije od Λ. Međutim, male modifikacije relacije (2.22) von Misesove funkcije obezbeđuju potpunu karakterizaciju $MDA(\Lambda)$.

Teorema 2.3.15 (Karakterizacija $MDA(\Lambda)$)

Funkcija raspodele F sa krajnjom desnom tačkom $x_F \leq \infty$ pripada maksimalnom domenu atrakcije od Λ ako i samo ako postoji neko $z < x_F$ tako da za F važi

$$\bar{F}(x) = c(x) exp\left\{-\int_z^x \frac{g(t)}{a(t)} dt\right\}, \quad z < x < x_F \tag{2.30}$$

gde su c i g merljive funkcije koje zadovoljavaju $c(x) \to c > 0$, $g(x) \to 1$ kad $x \uparrow x_F$ i $a(x)$ je pozitivna, apsolutno neprekidna funkcija (u odnosu na Lebesgueovu meru) sa gustinom $a'(x)$ za koju važi $\lim_{x \uparrow x_F} a'(x) = 0$.
Za F za koje važi (2.28) možemo izabrati

$$d_n = F^{\leftarrow}(1 - n^{-1}) \text{ i } c_n = a(d_n)$$

kao normirajuće konstante.
Moguć izbor funkcije a je

$$a(x) = \int_x^{x_F} \frac{\bar{F}(t)}{\bar{F}(x)} dt, \quad x < x_F. \tag{2.31}$$

□

Tvrđenje 2.3.16 (Osobina zatvaranja $MDA(\Lambda)$ u uslovima ekvivalencije repa)
Neka su F i G funkcije raspodele sa istom desnom krajnjom tačkom $x_F = x_G$ i pretpostavimo da $F \in MDA(\Lambda)$ za normirajuće konstante $c_n > 0$ i $d_n \in \mathbb{R}$, to jest

$$\lim_{n\to\infty} F^n(c_n x + d_n) = \Lambda(x), \; x \in \mathbb{R}. \tag{2.32}$$

Tada

$$\lim_{n\to\infty} G^n\,(c_n x + d_n) = \Lambda(x+b),\ x \in \mathbb{R}$$

ako i samo ako za $F\ i\ G$ važi ekvivalencija repa sa

$$\lim_{x\uparrow x_F} \bar{F}(x)/\bar{G}(x) = e^b.$$ □

2.4 Uopštene raspodele ekstremnih vrednosti i uopštena Pareto raspodela

U poglavlju 2.2 smo pokazali da standardne raspodele ekstremnih vrednosti obezbeđuju samo nedegeneratine granične zakone za linearno transformisane maksimume iid slučajnih promenljivih.

Jednoparametarske reprezentacije tri standardna slučaja jedne familije funkcija raspodele mogu biti predstavljene uvođenjem parametra ξ tako da

$\xi = \alpha^{-1} > 0$ za Frechetovu raspodelu ϕ_α

$\xi = 0$ za Gumbelovu raspodelu Λ

$\xi = -\alpha^{-1} < 0$ za Weibullovu raspodelu ψ_α

Definicija 2.4.1 (Jenkinson-Von Misesova reprezentacija raspodela ekstremnih vrednosti: uopštena raspodela ekstremnih vrednosti)
Definišemo funkciju raspodele H_ξ sa

$$H_\xi(x) = \begin{cases} exp\left\{-(1+\xi x)^{-1/\xi}\right\}, & ako\ je\ \xi \neq 0 \\ exp\{-exp\{-x\}\}, & ako\ je\ \xi = 0, \end{cases}$$

gde je $1 + \xi x > 0$.
Prema tome, nosač funkcijeH_ξ je

$$x > -\xi^{-1} \text{ za } \xi > 0,$$

$$x < -\xi^{-1} \text{ za } \xi < 0,$$

$$x \in \mathbb{R} \text{ za } \xi = 0.$$

H_ξ se naziva standardnom raspodelom ekstremnih vrednosti. Možemo uvesti familiju slučajnih promenljivih $H_{\xi;\mu,\psi}$ invarijantnu u odnosu na linearne transfomacije smenom argumenata x sa $(x-\mu)/\psi$ za $\mu \in \mathbb{R}, \psi > 0$.
Nosač funkcije treba da se modifikuje u skladu sa tim. Za $H_{\xi;\mu,\psi}$ takođe kažemo da je uopštena raspodela ekstremnih vrednosti. □

Smatramo da su funkcije raspodele H_0 granice od H_ξ za $\xi \to 0$. Ova reprezentacija

$$H_\xi(x) = exp\left\{-(1+\xi x)^{-1/\xi}\right\},\ \ 1 + \xi x > 0$$

se koristi kao reprezentacija za sve $\xi \in \mathbb{R}$.

Teorema 2.4.2 (Karakterizacija $MDA(H_\xi)$)
Za $\xi \in \mathbb{R}$ sledeća tvrđenja su ekvivalentna:

(a) $F \in MDA(H_\xi)$
(b) Postoji pozitivna, merljiva funkcija $a(\cdot)$ tako da za $1 + \xi x > 0$,

$$\lim_{u\uparrow x_F} \frac{\bar{F}(u+xa(u))}{\bar{F}(u)} = \begin{cases}(1+\xi x)^{-1/\xi}, & ako\ \xi \neq 0 \\ e^{-x}, & ako\ \xi = 0.\end{cases} \tag{2.33}$$

(c) Za $x, y > 0, y \neq 1$

$$\lim_{s\to\infty} \frac{U(sx)-U(s)}{U(sy)-U(s)} = \begin{cases}\frac{x^\xi - 1}{y^\xi - 1}, & ako\ \xi \neq 0 \\ \frac{\ln x}{\ln y}, & ako\ \xi = 0,\end{cases} \tag{2.34}$$

gde je $U(t) = F^{\leftarrow}(1 - t^{-1})$, $t > 0$.

□

Skica dokaza
$(a) \Leftrightarrow (b)\ za\ \xi = 0$, to je u stvari Teorema 2.3.16.
Za $\xi > 0$ imamo $H_\xi(x) = \phi_\alpha\big(\alpha^{-1}(x+\alpha)\big)$ za $\alpha = 1/\xi$
Na osnovu Teoreme 2.3.6, (a) je ekvivalentno sa $\bar{F} \in \mathcal{R}_{-\alpha}$. Dalje, na osnovu Teoreme A3.3, za neko $z > 0$

$$\bar{F}(x) = c(x)exp\left\{-\int_z^x \frac{1}{a(t)}dt\right\}, \quad z < x < \infty,$$

gde je $c(x) \to c > 0$ i $\frac{a(x)}{x} \to \alpha^{-1}$ kad $x \to \infty$ lokalno uniformno. Prema tome,

$$\lim_{u\to\infty} \frac{\bar{F}\big(u + xa(u)\big)}{\bar{F}(u)} = \left(1 + \frac{x}{\alpha}\right)^{-\alpha}$$

što je (2.33). Ako važi (b), biramo $d_n = (1/\bar{F})^{\leftarrow}(n) = U(n)$, onda

$$1/\bar{F}(d_n) \sim n,$$

i sa $u = d_n$ u (2.33),

$$\left(1 + \frac{x}{\alpha}\right)^{-\alpha} = \lim_{n\to\infty} \frac{\bar{F}(d_n + xa(d_n))}{\bar{F}(d_n)} = \lim_{n\to\infty} n\bar{F}\big(d_n + xa(d_n)\big),$$

odakle na osnovu Tvrđenja 2.3.2, $F \in MDA(H_\xi)\ za\ \xi = \alpha^{-1}$
Slučaj $\xi < 0$ se pokazuje analogno.

$(b) \Leftrightarrow (c)$ Ograničavamo sa na slučaj $\xi \neq 0$, a $\xi = 0$ se dokazuje analogno.

Zbog jednostavnosti, pretpostavljamo da je F neprekidna i rastuća na $(-\infty, x_F)$. Neka je $s = 1/\bar{F}(u)$, tada (2.33) pišemo kao

$$A_s(x) = \left(s\bar{F}\left(U(s) + xa\big(U(s)\big)\right)\right)^{-1} \to (1+\xi x)^{1/\xi}, \quad s \to \infty.$$

Sada, za svako $s > 0, A_s(x)$ je opadajuća i za $s \to \infty$ konvergira neprekidnoj funkciji. Tada na osnovu Tvrđenja D7 i $A_s^{\leftarrow}(t)$ tačkasto konvergira odgovarajućoj inverznoj graničnoj funkciji, to jest

$$\lim_{s\to\infty} \frac{U(st)-U(s)}{a(U(s))} = \frac{t^{\xi}-1}{\xi} \tag{2.35}$$

(2.34) se dobija primenom (2.35) za $t = x \; i \; t = y$ i njihovim količnikom. □

Definicija 2.4.3 (Funkcija raspodele viška, srednja vrednost raspodele viška)
Neka je X slučajna promenljiva sa funkcijom raspodele F i krajnjom desnom tačkom x_F. Za fiksirano $u < x_F$

$$F_u(x) = P(X - u \leq x | X > u), x \geq 0 \tag{2.36}$$

je funkcija raspodele viška za slučajnu promenljivu X iznad praga u. Funkcija

$$e(u) = E(X - u | X > u)$$

se naziva srednja vrednost raspodele viška od X. □

U osiguranju se F_n uglavnom odnosi na funkciju raspodele gubitka iznad nekog praga.

Definicija 2.4.4 (Uopštena Pareto raspodela (UPR))
Defnišemo funkciju raspodele G_ξ sa

$$G_\xi(x) = \begin{cases} 1 - (1 + \xi x)^{-1/\xi}, & \xi \neq 0, \\ 1 - e^{-x}, & \xi \neq 0, \end{cases} \tag{2.37}$$

gde je

$$x \geq 0 \text{ ako } \xi \geq 0,$$
$$0 \leq x \leq -1/\xi \text{ ako } \xi < 0.$$

G_ξ se naziva standardnom uopštenom Pareto raspodelom. Može se uvesti familija slučajnih promenljivih $G_{\xi;\nu,\beta}$ invarijantna u odnosu na linearne transformacije smenom argumenta $x \; sa \; (x - \nu)/\beta$ za $\nu \in \mathbb{R}, \beta > 0$.
Nosač funkcije treba da se modifikuje u skladu sa tim. $G_{\xi;\nu,\beta}$ se takođe naziva uopštenom Pareto raspodelom. □

G_0 je granica za G_ξ kad $\xi \to 0$.
Uvodimo notaciju

$$G_{\xi,\beta}(x) = 1 - \left(1 + \xi \frac{x}{\beta}\right)^{-1/\xi}, \quad x \in D(\xi, \beta), \tag{2.38}$$

gde je

$$x \in D(\xi, \beta) = \begin{cases} [0, \infty), & \xi \geq 0, \\ [0, -\beta/\xi], & \xi < 0. \end{cases}$$

Kada kažemo da X ima uopštenu Pareto raspodelu sa parametrima $\xi\ i\ \beta$, podrazumevamo da X ima funkciju raspodele $G_{\xi,\beta}$.

Teorema 2.4.5 (Osobine uopštene Pareto raspodele)

(a) Pretpostavimo da X ima UPR sa parametrima $\xi\ i\ \beta$
$E(X) < \infty$ ako i samo ako $\xi < 1$. U tom slučaju

$$E\left(1+\frac{\xi}{\beta}X\right)^{-r} = \frac{1}{1+\xi r},\ \ r > -1/\xi,$$

$$E\left(\ln\left(1+\frac{\xi}{y}X\right)\right)^{k} = \xi^k k!,\ \ k \in N,$$

$$EX\left(\bar{G}_{\xi,\beta}(X)\right)^{r} = \frac{\beta}{(r+1-\xi)(r+1)},\ \ (r+1)/|\varepsilon| > 0.$$

Ako je $\xi < 1/r,\ r \in \mathbb{N}$, tada

$$E(X^r) = \frac{\beta^r}{\xi^{r+1}}\frac{\Gamma(\xi^{-1}-r)}{\Gamma(1+\xi^{-1})}r!$$

(b) Za svako $\xi \in \mathbb{R}, F \in MDA(H_\xi)$ ako i samo ako

$$\lim_{u\uparrow x_F}\sup_{0<x<x_F-u}\left|F_u(x) - G_{\xi,\beta(u)}(x)\right| = 0 \qquad (2.39)$$

za neku pozitivnu funkciju β.

(c) Pretpostavimo $x_i \in D(\xi,\beta)\ i = 1{,}2$, tada

$$\frac{\bar{G}_{\xi,\beta}(x_1+x_2)}{\bar{G}_{\xi,\beta}(x_1)} = \bar{G}_{\xi,\beta+\xi x_1}(x_2). \qquad (2.40)$$

(d) Pretpostavimo da N ima Poissonovu raspodelu sa parametrom λ i da je nezavisno od iid niza (X_n) koji ima UPR sa parametrima $\xi\ i\ \beta$. Pišemo $M_N = max(X_1, \dots, X_N)$. Tada

$$P(M_N \le x) = exp\left\{-\lambda\left(1+\xi\frac{x}{\beta}\right)^{-1/\xi}\right\} = H_{\xi;\mu,\psi}(x),$$

gde je $\mu = \beta\xi^{-1}(\lambda^3-1)\ i\ \psi = \beta\lambda^3$

(e) Pretpostavimo da X ima UPR sa parametrima $\xi < 1$ i β. Tada za $u < x_F$

$$e(u) = E(X-u|X>u) = \frac{\beta+\xi u}{1-\xi}, \beta + u\xi > 0$$

□

2.5 Skoro sigurno ponašanje maksimuma

U ovom poglavlju proučavamo skoro sigurno ponašanje maksimuma

$$M_1 = X_1, M_n = max(X_1, \dots, X_n),\ n \ge 2,$$

za iid niz $X, X_1, X_2 \dots$ sa zajedničkom nedegenerativnom funkcijom raspodele F. Prvo proučavamo verovatnoće.
$P(M_n > u_n\, i.o)\ i\ P(M_n \le u_n\, i.o)$ za neopadajući niz (u_n) realnih brojeva. Ove verovatnoće ćemo potpuno okarakterisati uz pomoć repa $\bar{F}(u_n)$.

Događaj $\{A_n \ i.o\}$ koji opisuje da se realizuje beskonačno mnogo događaja iz niza (A_n) se može zapisati kao

$$\limsup_{n\to\infty} A_n = \bigcap_{k=1}^{\infty} \bigcup_{n\geq k} A_n.$$

Standardna Borel-Cantellieva lema kaže da

$$\sum_{n=1}^{\infty} P(A_n) < \infty \text{ implicira } P(A_n \ i.o) = 0.$$

Borel–Cantellijeva lema 2 za nezavisno A_n nam govori da

$$\sum_{n=1}^{\infty} P(A_n) = \infty \text{ implicira } P(A_n \ i.o) = 1.$$

Teorema 2.5.1 (Karakterizacija maksimalnog skoro sigurnog rasta parcijalnih maksimuma)
Pretpostavimo da je (u_n) neopadajuće. Tada

$$P(M_n > u_n \ i.o) = P(X_n > u_n \ i.o). \qquad (2.41)$$

Specijalno,

$$P(M_n > u_n \ i.o) = 0 \text{ ako i samo ako } \sum_{n=1}^{\infty} P(X > u_n) < \infty$$

i (2.42)

$$P(M_n > u_n \ i.o) = 1 \text{ ako i samo ako } \sum_{n=1}^{\infty} P(X > u_n) = \infty.$$

□

Primetimo da drugo tvrđenje teoreme odmah sledi iz (2.41). Zaista, na osnovu Borel-Cantellieve leme 1 i 2,
$P(X_n > u_n \ i.o.) = 0$ ili 1 ako i samo ako (2.42) važi.
Dokaz
Dovoljno je pokazati da (2.41) važi. Pošto je $M_n \geq X_n$ za svako n, treba samo pokazati da

$$P(M_n > u_n \ i.o) \leq P(X_n > u_n \ i.o). \qquad (2.43)$$

Neka x_F označava krajnju desnu tačku raspodele F i pretpostavimo da $u_n \geq x_F$ za neko n. Tada

$$P(M_n > u_n) = P(X_n > u_n) = 0$$

za svako veliko n, pa je (2.43) zadovoljeno. Stoga pretpostavimo da je $u_n < x_F$ za svako n. Tada

$$F(u_n) < 1 \text{ za svako } n.$$

Ako $u_n \uparrow x_F$ i $M_n > u_n$ za beskonačno mnogo n, tada postoji beskonačno mnogo n za koje je $X_n > u_n$.
Sada pretpostavimo da $u_n \uparrow b < x_F$. Međutim, tada

$$\bar{F}(u_n) \geq \bar{F}(b) > 0.$$

Na osnovu Borel-Cantellieve leme 2, $P(X_n > u_n \ i.o) = 1$ i tada (2.41) sigurno važi. □

Teorema 2.5.2 (Karakterizacija minimalnog skoro sigurnog rasta parcijalnih maksimuma)
Pretpostavimo da je (u_n) neopadajuće i da važe sledeći uslovi:

$$\bar{F}(u_n) \to 0, \tag{2.44}$$

$$n\bar{F}(u_n) \to \infty. \tag{2.45}$$

Tada

$$P(M_n \leq u_n \; i.o) = 0 \text{ ili } = 1,$$

u zavisnosti od toga da li je

$$\sum_{n=1}^{\infty} \bar{F}(u_n) exp\{-n\bar{F}(u_n)\} < \infty \; ili = \infty. \tag{2.46}$$

Štaviše, ako

$$\bar{F}(u_n) \to c > 0 \text{ onda } P(M_n \leq u_n \; i.o) = 0,$$

dok ako

$$liminf_{n\to\infty} n\bar{F}(u_n) < \infty, \text{onda } P(M_n \leq u_n \;\; i.o) = 1.$$

□

Važi da je $limsup_{n\to\infty} c_n^{-1}(M_n - d_n) = 1$ skoro sigurno za $c_n > 0$ i $d_n \in \mathbb{R}$, pa je

$$P(M_n > c_n(1+\varepsilon) + d_n \; i.o) = 0 \;\; \text{ili} = 1,$$

kad $\varepsilon > 0$ ili $\varepsilon < 0$ za malo $|\varepsilon|$. Slično,

$$liminf_{n\to\infty} c_n^{-1}(M_n - d_n) = 1 \text{ skoro sigurno},$$

važi ako i samo ako

$$P(M_n \leq c_n(1+\varepsilon) + d_n \; i.o) = 0 \text{ ili} = 1,$$

kad $\varepsilon < 0$ ili $\varepsilon > 0$ za malo $|\varepsilon|$. Tada sledeća posledica odmah sledi iz Teoreme 2.5.1 i 2.5.2.

Posledica 2.5.3 (Karakterizacija gornje i donje granice maksimuma)

(a) Pretpostavimo da su nizovi $u_n(\varepsilon) = c_n(1+\varepsilon) + d_n,\; n \in \mathbb{N}$ neopadajući za svako $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$. Tada relacija
$$\sum_{n=1}^{\infty} \bar{F}\big(u_n(\varepsilon)\big) < \infty \;\text{ ili }\; = \infty,$$
kad $\varepsilon \in (0, \varepsilon_0) \; ili \; \varepsilon \in (-\varepsilon_0, 0)$ implicira da
$$limsup_{n\to\infty} c_n^{-1}(M_n - d_n) = 1 \text{ skoro sigurno.}$$

(b) Pretpostavimo da su nizovi $u_n(\varepsilon) = c_n(1+\varepsilon) + d_n, n \in \mathbb{N}$ neopadajući i da zadovoljavaju (2.44) i (2.45) za svako $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$. Tada relacija
$$\sum_{n=1}^{\infty} \bar{F}\big(u_n(\varepsilon)\big)\, exp\big\{-n\, \bar{F}\big(u_n(\varepsilon)\big)\big\} < \infty \;\text{ ili }\; = \infty,$$
kad $\varepsilon \in (-\varepsilon_0, 0)$ ili $\varepsilon \in (0, \varepsilon_0)$ implicira da
.
$$liminf_{n\to\infty} c_n^{-1}\,(M_n - d_n) = 1 \text{ skoro sigurno.}$$
□

3. Fluktuacije statistika višeg reda

3.1 Statistike reda

Neka $X, X_1, X_2 \dots$ označava niz iid nedegenerativnih slučajnih promenljivih sa zajedničkom funkcijom raspodele F. U ovom delu ćemo navesti neke važne osobine statistika višeg reda konačnog uzorka $X_1, X_2 \dots X_n$.
Definišemo uređeni uzorak (varijacioni niz)

$$X_{n,n} \leq \cdots \leq X_{1,n}.$$

Prema tome, $X_{n,n} = min(X_1, \dots, X_n)$ i $X_{1,n} = M_n = max(X_1, \dots, X_n)$. Slučajna promenljiva $X_{k,n}$ se naziva k-ta statistika višeg reda.

Veza između statistike reda i empirijske funkcije raspodele uzorka je odmah vidljiva: za $x \in \mathbb{R}$ uvodimo empirijsku funkciju raspodele ili uzoračku funkciju raspodele

$$F_n(x) = \frac{1}{n} card\{i: 1 \leq i \leq n,\ X_i \leq x\} = \frac{1}{n} \sum_{i=1}^{n} I_{\{X_i \leq x\}},\ x \in \mathbb{R},$$

gde I_A označava indikator funkciju skupa A. Sada

$$X_{k,n} \leq x \text{ ako i samo ako } \sum_{i=1}^{n} I_{\{x_i > x\}} < k, \tag{3.1}$$

što implicira da

$$P(X_{k,n} \leq x) = P\left(F_n(x) > 1 - \frac{k}{n}\right).$$

Statistike višeg reda procenjuju repove i kvantile, kao i verovatnoće prelaska preko određenog praga. Podsećamo se definicije funkcije kvantila za funkciju raspodele F,

$$F^{\leftarrow}(t) = inf\{x \in \mathbb{R}: F(x) \geq t\},\ \ 0 < t < 1.$$

Za uzorak $X_1, \dots, X_n$ empirijsku funkciju kvantila označavamo sa $F_n^{\leftarrow}$. Ako je F neprekidna funkcija, onda se ponavljanja u uzorku javljaju jedino sa verovatnoćom 0, te mogu biti zanemarene, to jest možemo da pretpostavimo da $X_{n,n} < \cdots < X_{1,n}$. U tom slučaju je $F_n^{\leftarrow}$ funkcija statistike reda, to jest

$$F_n^{\leftarrow}(t) = X_{k,n} \text{ za } 1 - \frac{k}{n} < t \leq 1 - \frac{k-1}{n}, \tag{3.2}$$

za $k = 1, \dots, n$.

Tvrđenje 3.1.1 (Funkcija raspodele k-te statistike višeg reda)
Za $k = 1, \dots, n$ neka $F_{k,n}$ označava funkciju raspodele za $X_{k,n}$. Tada

(a) $F_{k,n}(x) = \sum_{r=0}^{k-1} \binom{n}{r} \bar{F}^r(x) F^{n-r}(x)$

(b) Ako je F neprekidna, tada

$$F_{k,n}(x) = \int_{-\infty}^{x} f_{k,n}(z) dF(z),$$

gde je

$$f_{k,n}(x) = \frac{n!}{(k-1)!(n-k)!} F^{n-k}(x)\bar{F}^{k-1}(x);$$

to jest $f_{k,n}$ je gustina za $F_{k,n}$.

Dokaz

(a) Za $n \in \mathbb{N}$ definišemo

$$B_n = \sum_{i=1}^{n} I_{\{x_i > x\}}.$$

Tada je B_n suma n iid Bernoullievih promenljivih sa verovatnoćom uspeha

$$E(I_{\{X>x\}}) = P(X > x) = \bar{F}(x).$$

Prema tome, B_n je binomna slučajna promenljiva sa parametrima n i $\bar{F}(x)$. Primenom (3.1) dobijamo za $x \in \mathbb{R}$

$$F_{k,n}(x) = P(B_n < k) = \sum_{r=0}^{k-1} P(B_n = r) = \sum_{r=0}^{k-1} \binom{n}{r} \bar{F}^r(x)\, F^{n-r}(x).$$

(b) Uzimajući u obzir neprekidnost od F dobijamo

$$\frac{n!}{(k-1)!(n-k)!} \int_{-\infty}^{x} F^{n-k}(z)\bar{F}^{k-1}(z)dF(z) = \frac{n!}{(k-1)!(n-k)!} \int_{\bar{F}(x)}^{1} (1-t)^{n-k} t^{k-1} dt =$$
$$\sum_{r=0}^{k-1} \binom{n}{r} \bar{F}^r(x) F^{n-r}(x) = F_{k,n}(x).$$

Poslednje sledi iz reprezentacije nepotpune beta funkcije, može se pokazati višestrukom parcijalnom integracijom. □

Slični argumenti dovode do zajedničke raspodele konačnog broja različitih statstika reda. Ako je na primer F apsolutno neprekidna sa gustinom f, onda je zajednička gustina od $(X_1, \dots, X_n)$

$$f_{X_1,\dots,X_n}(x_1, \dots, x_n) = \prod_{i=1}^{n} f(x_i), \quad (x_1, \dots, x_n) \in \mathbb{R}^n.$$

Pošto se n vrednosti od $(X_1, \dots, X_n)$ mogu zapisati na $n!$ načina, svaka kolekcija reda $(X_{k,n})_{k=1,\dots,n}$ je mogla biti uzeta iz $n!$ različitih uzoraka. Zajednička raspodela uzoraka reda postaje:

$$f_{X_{1,n},\dots,X_{n,n}}(x_1, \dots, x_n) = n! \prod_{i=1}^{n} f(x_i), \quad x_n < \cdots < x_1. \tag{3.3}$$

Sledeća teorema o marginalnim raspodelama je posledica jednakosti (3.3).

Teorema 3.1.2 (Zajednička gustina k statistika višeg reda)
Ako je F apsolutno neprekidna sa gustinom f, tada

$$f_{X_{1,n},\dots,X_{k,n}}(x_1, \dots, x_k) = \frac{n!}{(n-k)!} F^{n-k}(x_k) \prod_{i=1}^{k} f(x_i), \quad x_k < \cdots < x_1.$$ □

Definicija 3.1.3 (Razmaci uzorka)
Za uzorak $X_1, \dots, X_n$ razmake definišemo sa

$$X_{k,n} - X_{k+1,n}, \quad k = 1, \dots, n-1.$$

Za slučajnu promenljivu sa konačnim levim (desnim) krajnjim tačkama $\tilde{x}_F$ (x_F) definišemo n-ti (0-ti) razmak sa $X_{n,n} - X_{n+1,n} = X_{n,n} - \tilde{x}_F$ $(X_{0,n} - X_{1,n} = x_F - X_{1,n})$ □

Lema 3.1.4 (Transformacija kvantila)
Neka je $X_1, \dots, X_n$ sa funkcijom raspodele F. Dalje, neka su $U_1, \dots, U_n$ iid slučajne promenljive sa uniformnom raspodelom na $(0,1)$ i sa $U_{n,n} < \dots < U_{1,n}$ označimo odgovarajuću statistiku reda. Tada važe sledeći rezultati:

(a) $F_n^{\leftarrow}(U_1) \stackrel{d}{=} X_1$.
(b) za svako $n \in \mathbb{N}$,

$$\left(X_{1,n}, \dots, X_{n,n}\right) \stackrel{d}{=} \left(F_n^{\leftarrow}(U_{1,n}), \dots, F_n^{\leftarrow}(U_{n,n})\right).$$

(c) Slučajna promenljiva $F(X_1)$ ima uniformnu raspodelu na $(0,1)$ ako i samo ako je F neprekidna funkcija. □

Tvrđenje 3.1.5 (Skoro sigurna konvergencija statistika reda)
Neka je F funkcija raspodele sa krajnjom desnom (levom) tačkom $x_F < \infty$ $(\tilde{x}_F \geq -\infty)$ i $(k(n))$ neopadajući celobrojni niz tako da

$$\lim_{n\to\infty} n^{-1}k(n) = c \in [0,1].$$

(a) Ako je $x_F < \infty$, onda $X_{k(n),n} \stackrel{s.s}{\rightarrow} x_F$ kad $c = 0$
i
ako je $\tilde{x}_F \geq -\infty$, onda je $X_{k(n),n} \stackrel{s.s}{\rightarrow} (\tilde{x}_F)$ kad je $c = 1$.
(b) Pretpostavimo da je $c \in (0,1)$ takvo da postoji jedinstveno rešenje $x(c)$ jednačine $\bar{F}(x) = c$. Tada

$$X_{k(n),n} \stackrel{s.s}{\rightarrow} x(c).$$

Dokaz Pokazaćemo (b). Dokaz za (a) se dobija na isti način.
Na osnovu (3.1) i JZVB,

$$P\left(X_{k(n),n} \leq x \quad i.o\right) = P\left(\frac{n}{k(n)} \frac{1}{n}\sum_{i=1}^{n} I_{\{X_i > x\}} < 1 \quad i.o\right) = P\left(\bar{F}(x)\left(1 + o(1)\right) < c \quad i.o\right)$$

Poslednja verovatnoća je 0 ili 1 kad $x < x(c)$ ili $x > x(c)$.Prema tome, $\liminf_{n\to\infty} X_{k(n),n} = x(c)$ skoro sigurno. Analogno možemo pokazati da važi relacija $\limsup_{n\to\infty} X_{k(n),n} = x(c)$ skoro sigurno. □

3.2 Granična raspodela statistika višeg reda

Neka je $X_1, \dots, X_n$ iid sa funkcijom raspodele F. Tvrđenjem 2.1.1 se podsećamo da za niz (u_n) pragova i $0 \leq \tau \leq \infty$,

$$\lim_{n\to\infty} P\left(X_{1,n} \leq u_n\right) = e^{-\tau} \Leftrightarrow \lim_{n\to\infty} n\bar{F}(u_n) = \tau. \tag{3.4}$$

Interesuje nas da li možemo da proširimo relaciju (3.4) do neke statistike višeg reda $X_{k,n}$ za fiksirano $k \in \mathbb{N}$, ili čak dobijemo zajedničke granične verovatnoće za fiksiran broj k statistika višeg reda $X_{k,n}, \dots, X_{1,n}$.
Razmotrićemo za $n \in \mathbb{N}$ broj prelazaka preko praga u_n za $X_1, \dots, X_n$:

$$B_n = \sum_{i=1}^{n} I_{\{X_i > u_n\}}.$$

Tada je B_n binomna slučajna promenljiva sa parametrima n i $\bar{F}(u_n)$. U tvrđenju 3.1.1 smo koristili ovaj kvantitet za konačno n da bismo izračunali funkciju raspodele k-te statistike višeg reda. Kada podignemo prag, prelasci $\{X_i > u_n\}$ postaju ređi. S druge strane, povećamo veličinu uzorka. Kada uzmemo oba u obzir, $E(B_n) = n\bar{F}(u_n) \to \tau$ kad $n \to \infty$ pa zbog toga možemo da primenimo klasičnu Poissonovu teoremu:

$$B_n \overset{d}{\to} \mathcal{P}(\tau).$$

Pragovi u_n su birani tako da očekivani broj prelazaka konvergira.

Teorema 3.2.1 (Granični zakon za broj prelazaka preko praga)
Pretpostavimo da je (u_n) niz na $\mathbb{R}$ tako da $n\bar{F}(u_n) \to \tau$ za neko $\tau \in [0,\infty)$ kad $n \to \infty$. Tada

$$\lim_{n\to\infty} P(B_n \le k) = e^{-\tau} \sum_{r=0}^{k} \frac{\tau^r}{r!}, \quad k \in \mathbb{N}_0. \tag{3.5}$$

Kad je $\tau = 0$, desna strana jednakosti je 1, a kad je $\tau = \infty$, onda je 0.
Ako (3.5) važi za neko $k \in \mathbb{N}_0$, onda $n\bar{F}(u_n) \to \tau$ kad $n \to \infty$, pa (3.5) važi za svako $k \in \mathbb{N}_0$.

Dokaz: Za $\tau \in (0,\infty)$ dokaz za dovoljan uslov je Poissonova granična teorema. Za $\tau = 0$,

$$P(B_n \le k) \ge P(B_n = 0) = \left(1 - \bar{F}(u_n)\right)^n = \left(1 + o\left(\frac{1}{n}\right)\right)^n \to 1.$$

Za $\tau = \infty$ i proizvoljno $\theta > 0$ imamo da $n\bar{F}(u_n) \ge \theta$ za veliko n. Pošto binomna funkcija raspodele opada po θ, dobijamo

$$P(B_n \le k) \le \sum_{r=0}^{k} \binom{n}{r} \left(\frac{\theta}{n}\right)^r \left(1 - \frac{\theta}{n}\right)^{n-r}.$$

Prema tome, za fiksirano k

$$\limsup_{n\to\infty} P(B_n \le k) \le e^{-\theta} \sum_{r=0}^{k} \frac{\theta^r}{r!} \to 0, \quad \theta \to \infty,$$

pa je $P(B_n \le k) \to 0$ kad $n \to \infty$.
Obrnuto, pretpostavimo da (4.13) važi za neko $k \in \mathbb{N}_0$, ali $n\bar{F}(u_n) \nrightarrow \tau$. Tada postoji neko $\tau' \neq \tau$ na $[0,\infty]$ i podniz (n_k) tako da $n_k \bar{F}(u_{n_k}) \to \tau'$ kad $k \to \infty$, pa B_{n_k} slabo konvergira Poissonovoj slučajnoj promenljivoj sa parametrom τ', što je kontradikcija sa (3.5). □

Iz Poissonove aproksimacije (3.5) dobijamo asimptote za k-tu statistiku reda. Definicija za B_n i (4.1) impliciraju da

$$P(B_n < k) = P\big(X_{k,n} \le u_n\big), \quad 1 \le k \le n, \tag{3.6}$$

odakle, zbog (3.5) dobijamo sledeću teoremu:

Teorema 3.2.2 (Granične verovatnoće za statistike višeg reda)
Pretpostavimo da je (u_n) niz na $\mathbb{R}$ tako da $n\bar{F}(u_n) \to \tau \in [0,\infty)$ kad $n \to \infty$. Tada

$$\lim_{n\to\infty} P\big(X_{k,n} \le u_n\big) = e^{-\tau} \textstyle\sum_{r=0}^{k-1} \frac{\tau^r}{r!}, \; k \in \mathbb{N}. \tag{3.7}$$

Za $\tau = 0$ je desna strana jednakosti 1, a za $\tau = \infty$ je 0. Ako važi (3.7) za neko $k \in \mathbb{N}$, onda $n\bar{F}(u_n) \to \tau$ kad $n \to \infty$, pa (3.7) važi za svako $k \in \mathbb{N}$. □

Za $u_n = c_n x + d_n$ i $\tau = \tau(x) = -\ln H(x)$ kao u Tvrđenju 2.3.2, dobijamo sledeću posledicu.

Posledica 3.2.3 (Granična raspodela statistike višeg reda)
Pretpostavimo da $F \in MDA(H)$ sa normirajućim konstantama $c_n > 0 \; i \; d_n \in \mathbb{R}$.
Definišemo

$$H^{(k)}(x) = H(x) \sum_{r=0}^{k-1} \frac{(-\ln H(x))^r}{r!}, x \in \mathbb{R}.$$

Za x za koje je $H(x) = 0$, kažemo da je $H^{(k)}(x) = 0$. Tada za svako $k \in \mathbb{N}$,

$$\lim_{n\to\infty} P\big(c_n^{-1}\big(X_{k,n} - d_n\big) \le x\big) = H^{(k)}(x). \tag{3.8}$$

S druge strane, ako je za neko $k \in \mathbb{N}$

$$\lim_{n\to\infty} P\big(c_n^{-1}\big(X_{k,n} - d_n\big) \le x\big) = G(x), \quad x \in \mathbb{R},$$

za nedegenerativnu funkciju raspodele G, tada je $G = H^{(k)}$ za neku raspodelu ekstremne vrednosti H i (3.8) važi za svako $k \in \mathbb{N}$. □

Teorema 3.2.4 (Multivarijantni granični zakon za broj prelazaka preko praga)
Pretpostavimo da niz $\left(u_n^{(j)}\right)$ zadovoljava jednakost

$$\lim_{n\to\infty} n\overline{F}\left(u_n^{(j)}\right) = \tau_j, \; j = 1, \dots, k, \tag{3.9}$$

gde je $0 \le \tau_1 \le \cdots \le \tau_k \le \infty$. Tada

$$\lim_{n\to\infty} P\Big(B_n^{(1)} = l_1, B_n^{(2)} = l_1 + l_2, \dots, B_n^{(k)} = l_1 + \cdots + l_k\Big) =$$

$$= \frac{\tau_1^{l_1}}{l_1!} \frac{(\tau_2-\tau_1)^{l_2}}{l_2!} \cdots \frac{(\tau_k-\tau_{k-1})^{l_k}}{l_k!} e^{-\tau_k}, \tag{3.10}$$

gde je $B_n^{(j)} = \sum_{i=1}^{n} I_{\left\{X_i > u_n^{(j)}\right\}}, \; j = 1, \dots, k$, tj. $B_n^{(j)}$ je broj prelazaka preko praga $u_n^{(j)}$ sa $X_1, \dots, X_n$.

Desna strana jednakosti je nula ako je $\tau_k = \infty$. □

Dokaz Pišemo $p_{n,j} = \bar{F}\left(u_n^{(j)}\right)$. Koristeći osobine multinomne raspodele, vidimo da je verovatnoća s leve strane jednakosti (3.10) jednaka

$$\binom{n}{l_1} p_{n,1}^{l_1} \binom{n-l_1}{l_2} \left(p_{n,2} - p_{n,1}\right)^{l_2} \dots \binom{n-l_1-\dots-l_{k-1}}{l_k} \left(p_{n,k} - p_{n,k-1}\right)^{l_k} \left(1 - p_{n,k}\right)^{n-l_1-\dots-l_k}.$$

Ako je $\tau_k < \infty$ iz (3.9) dobijamo

$$\binom{n}{l_1} p_{n,1}^{l_1} \sim \frac{\left(np_{n,1}\right)^{l_1}}{l_1!} \to \frac{\tau_1^{l_1}}{l_1!},$$

$$\binom{n-l_1-\dots-l_{i-1}}{l_i} \left(p_{n,1} - p_{n,i-1}\right)^{l_i} \sim \frac{\left(np_{n,i} - np_{n,i-1}\right)^{l_i}}{l_i!} \to \frac{(\tau_i - \tau_{i-1})^{l_1}}{l_i!}, \text{ za } 2 \le i \le k,$$

$$\left(1 - p_{n,k}\right)^{n-l_1-\dots-l_k} \sim \left(1 - \frac{np_{n,k}}{n}\right)^n \to e^{-\tau_k}$$

što daje (3.10).

Ako je $\tau_k = \infty$, verovatnoća u (3.10) ne prelazi $P\left(B_n^{(k)} = \sum_{i=1}^{k} l_i\right)$. Na osnovu Teoreme 3.2.1, poslednje konvergira ka nuli. □

Očigledno je, kao u (3.6)

$$P\left(X_{1,n} \le u_n^{(1)}, \dots, X_{k,n} \le u_n^{(k)}\right) = P\left(B_n^{(1)} = 0,\ B_n^{(2)} \le 1, B_n^{(k)} \le k-1\right) \quad (3.11)$$

pa zajednička asimptotska raspodela k statistika višeg reda može da se dobije direktno iz Teoreme 3.2.4. Specijalno, ako $c_n^{-1}\left(X_{1,n} - d_n\right)$ slabo konvergira, tada slabo konvergira i vektor

$$\left(c_n^{-1}\left(X_{1,n} - d_n\right), \dots, c_n^{-1}\left(X_{k,n} - d_n\right)\right).$$

Iako se za malo k zajedničke granične raspodele k statistika višeg reda mogu lako dobiti iz (3.11) i Teoreme 3.2.4, opšti slučaj je veoma komplikovan. Ako je funkcija raspodele F apsolutno neprekidna sa gustinom f koja zadovoljava određene uslove regularnosti, sledeći argument se može precizirati:

Pretpostavimo da $F \in MDA(H)$ sa gustinom f, tada funkcija raspodele maksimuma $F^n(c_n x + d_n) = P\left(c_n^{-1}\left(X_{1,n} - d_n\right) \le x\right)$ takođe ima gustinu takvu da za skoro sve $x \in \mathbb{R}$,

$$nc_n f(c_n x + d_n)\ F^{n-1}(c_n x + d_n) \sim nc_n f(c_n x + d_n) H(x) \to h(x),$$

gde je h gustina raspodele ekstremne vrednosti H. Dalje, za $k \in \mathbb{N}$ slaba granica slučajnog vektora $\left(c_n^{-1}\left(X_{j,n} - d_n\right)\right)_{j=1,\dots,k}$ na osnovu Teoreme 3.1.2, ima gustinu

$$\lim_{n\to\infty} F^{n-k}(c_n x_k + d_n) \prod_{j=1}^{k} \left((n-j+1) c_n f\left(c_n x_j + d_n\right)\right)$$

$$= H(x_k) \prod_{j=1}^{k} \frac{h(x_j)}{H(x_j)}, x_k < \dots < x_1. \quad (3.12)$$

Definicija 3.2.5 (k-dimenzionalna H-ekstremna obeležja)
Za bilo koju raspodelu ekstremnih vrednosti H sa gustinom h za $x_k < \cdots < x_1$ iz nosača funkcije H

$$h^{(k)}(x_1, \dots, x_k) = H(x_k) \prod_{j=1}^{k} \frac{h(x_j)}{H(x_j)}.$$

Vektor $\left(Y^{(1)}, \dots, Y^{(k)}\right)$ slučajnih promenljivih sa zajedničkom gustinom $h^{(k)}$ se naziva k-dimenzionalno H-ekstremno obeležje.

Teorema 3.2.6 (Zajednička granična raspodela k statistika višeg reda)
Pretpostavimo da $F \in MDA(H)$ sa normirajućim konstantama $c_n > 0$ i $d_n \in \mathbb{R}$. Tada, za svako fiksirano $k \in \mathbb{N}$,

$$\left(c_n^{-1}(X_{i,n} - d_n)\right)_{i=1,\dots,k} \xrightarrow{d} \left(Y^{(i)}\right)_{i=1,\dots,k} \quad n \to \infty,$$

gde je $\left(Y^{(1)}, \dots, Y^{(k)}\right)$ k-dimenzionalno H-ekstremno obeležje. □

Posledica 3.2.7 (Zajednička granična raspodela viših razmaka u $MDA(\Lambda)$)
Pretpostavimo da $F \in MDA(\Lambda)$ sa normirajućim konstantama $c_n > 0$. Tada

(a) $\left(c_n^{-1}(X_{i,n} - X_{i+1,n})\right)_{i=1,\dots,k} \xrightarrow{d} (i^{-1}E_i)_{i=1,\dots,k}$ za $k \geq 1$

(b) $c_n^{-1}\left(\sum_{i=1}^{k} X_{i,n} - kX_{k+1,n}\right) \xrightarrow{d} \sum_{i=1}^{k} E_i$ za $k \geq 2$,

gde su $E_1, \dots, E_k$ iid standardne eksponencijalne slučajne promenljive. □

Posledica 3.2.8 (Zajednička granična raspodela viših razmaka u $MDA(\phi_\alpha)$)
Pretpostavimo da $F \in MDA(\phi_\alpha)$ sa normirajućim konstantama $c_n > 0$. Neka je $\left(Y^{(1)}, \dots, Y^{(k+1)}\right)$ $(k+1)$-dimenzionalno Frechetovo obeležje. Tada

(a) $\left(c_n^{-1}(X_{i,n} - X_{i+1,n})\right)_{i=1,\dots,k} \xrightarrow{d} \left(Y^{(i)} - Y^{(i+1)}\right)_{i=1,\dots,k}$ $k \geq 1$,

(b) $c_n^{-1}\left(\sum_{i=1}^{k} X_{i,n} - kX_{k+1,n}\right) \xrightarrow{d} \sum_{i=1}^{k} i\left(Y^{(i)} - Y^{(i+1)}\right)$, $k \geq 2$

Granične promenljive u (a) i u (b) su definisane razmacima $Y^{(1)} - Y^{(2)}, \dots, Y^{(k)} - Y^{(k+1)}$ koji imaju zajedničku gustinu

$$g_{Y^{(1)}-Y^{(2)},\dots,Y^{(k)}-Y^{(k+1)}}(x_1, \dots, x_1)$$
$$= \alpha^{k+1} \int_0^\infty exp\{-y^{-\alpha}\} \left(y(y + x_k) \dots (y + x_k + \cdots + x_1)\right)^{-\alpha-1} dy$$

za $x_1, \dots, x_k > 0$. □

3.3 Granična raspodela slučajno indeksiranih statistika višeg reda

U ovom poglavlju upoređujemo slabo granično ponašanje statistika višeg reda i slučajno indeksiranih maksimuma za iid niz (X_n) slučajnih promenljivih sa zajedničkom funkcijom raspodele F. Neka je $\big(N(t)\big)_{t\geq 0}$ proces celobrojnih slučajnih promenljivih za koje pretpostavljamo da su nezavisne od (X_n). Pišemo

$$X_{n,n} \leq \cdots \leq X_{1,n} \text{ i } X_{N(t),N(t)} \leq \cdots \leq X_{1,N(t)}$$

za statistike reda uzoraka $X_1, \dots, X_n$ i $X_1, \dots, X_{N(t)}$, respektivno, koristimo i

$$M_n = X_{1,n} \text{ i } M_{N(t)} = X_{1,N(t)}$$

za odgovarajuće uzoračke maksimume.
Ako F pripada maksimalnom domenu atrakcije raspodele ekstremne vrednosti H $\big(F \in MDA(H)\big)$, postoje konstante $c_n > 0$ i $d_n \in \mathbb{R}$ tako da

$$c_n^{-1}(M_n - d_n) \xrightarrow{d} H. \tag{3.13}$$

Prirodno se postavlja pitanje:
Da li relacija (3.13) ostaje važeća za slučajan skup indeksa $\big(N(t)\big)$?
Iz Leme 1.5.5 znamo da (3.13) implicira da

$$c_{N(t)}^{-1}\big(M_{N(t)} - d_{N(t)}\big) \xrightarrow{d} H$$

pod pretpostavkom da $N(t) \xrightarrow{P} \infty$, ali želimo da zadržimo stare normirajuće nizove (c_n) i (d_n) umesto slučajnih procesa $\big(c_{N(t)}\big), \big(d_{N(t)}\big)$. Ovo može da se postigne pod opštim uslovima. Međutim, granične raspodele će se takođe promeniti.
Uvodimo promenljive

$$B_t^{(i)} = \sum_{j=1}^{N(t)} I_{\left\{X_j > u_t^{(i)}\right\}}, \qquad i = 1, \dots, k$$

koje računaju broj prelazaka preko pragova

$$u_t^{(k)} \leq \cdots \leq u_t^{(1)}, \quad t \geq 0 \tag{3.14}$$

za $X_1, \dots, X_{N(t)}$. Takođe pretpostavimo da postoje brojevi

$$0 \leq \tau_1 \leq \cdots \leq \tau_k \leq \infty,$$

tako da za $i = 1, \dots, k$,

$$tp_{t,i} = t\bar{F}\left(u_t^{(i)}\right) \to \tau_i, \quad t \to \infty. \tag{3.15}$$

Sledeći rezultat je analogan Teoremi 3.2.4:

Teorema 3.3.1 (Multivarijantni granični zakon za broj prelazaka preko praga)

Pretpostavimo da $\left(u_t^{(i)}\right)_{t\geq 0}$, $i = 1, \dots, k$ zadovoljava (3.14) i (3.15). Pretpostavimo takođe da postoji nenegativna slučajna promenljiva Z tako da

$$\frac{N(t)}{t} \xrightarrow{P} Z, t \to \infty. \tag{3.16}$$

Tada, za sve cele brojeve $l_i \geq 0, i = 1, \dots, k$

$$\lim_{t\to\infty} P\left(B_t^{(1)} = l_1,\ B_t^{(2)} = l_1 + l_2, \dots,\ B_t^{(k)} = l_1 + \dots + l_k\right) =$$
$$= E\left[\frac{(Z\tau_1)^{l_1}}{l_1!}\frac{\left(Z(\tau_2-\tau_1)\right)^{l_2}}{l_2!}\dots\frac{\left(Z(\tau_k-\tau_{k-1})\right)^{l_k}}{l_k!}\ e^{-Z\tau_k}\right].$$

Desna strana jednakosti je 0 kad je $\tau_k = \infty$.

Dokaz Zbog jednostavnosti se ograničavamo na slučaj $k = 2$. Koristimo nezavisnost $\left(N(t)\right)\ i\ (X_n)$ i primenjujemo (3.15) na (3.16):

$$P\left(B_t^{(1)} = l_1,, B_t^{(2)} = l_1 + l_2 | N(t)\right) \tag{3.17}$$
$$= \binom{N(t)}{l_1} p_{t,1}^{l_1} \binom{N(t) - l_1}{l_2} \left(p_{t,2} - p_{t,1}\right)^{l_2} \left(1 - p_{t,2}\right)^{N(t)-l_1-l_2} =$$
$$= \left(1 + o_P(1)\right) \frac{\left(N(t)p_{t,1}\right)^{l_1}}{l_1!} \frac{\left(N(t)\left(p_{t,2} - p_{t,1}\right)^{l_2}\right)}{l_2!} \left(1 - p_{t,2}\right)^{N(t)} =$$
$$= \left(1 + o_P(1)\right) \frac{\left(N(t)p_{t,1}\right)^{l_1}}{l_1!} \frac{\left(\frac{N(t)(tp_{t,1})}{t}\right)^{l_1}}{l_1!} \frac{\left(\frac{N(t)\left(t(p_{t,2}-p_{t,1})\right)}{t}\right)^{l_2}}{l_2!} exp\left\{\frac{N(t)\left(t\ln(1-p_{t,2})\right)}{t}\right\} \xrightarrow{P}$$
$$\xrightarrow{P} \frac{(Z\tau_1)^{l_1}}{l_1!} \frac{\left(Z(\tau_2-\tau_1)\right)^{l_2}}{l_2!} e^{-Z},\ t \to \infty. \tag{3.18}$$

Primetimo da su izrazi u (3.17) uniformno integrabilni i da je (3.18) integrabilno. Prema tome, možemo da zaključimo da

$$P\left(B_t^{(1)} = l_1,\ B_t^{(2)} = l_1 + l_2\right) = E\left[P\left(B_t^{(1)} = l_1,\ B_t^{(2)} = l_1 + l_2 | N(t)\right)\right] \to$$
$$\to E\left[\frac{(Z\tau_1)^{l_1}}{l_1!} \frac{\left(Z(\tau_2-\tau_1)\right)^{l_2}}{l_2!} e^{-Z\tau_2}\right] \text{kad } t \to \infty. \qquad \square$$

Možemo da iskoristimo identitet:

$$P\left(X_{1,N(t)} \leq u_t^{(1)}, \dots, X_{k,N(t)} \leq u_t^{(k)}\right) = P\left(B_t^{(1)} = 0,\ B_t^{(2)} \leq 1, \dots, B_t^{(k)} \leq k - 1\right),$$

i Teoremu 3.3.1 da izvedemo graničnu raspodelu za vektor statistika višeg reda $X_{1,N(t)}, \dots, X_{k,N(t)}$.
Zbog jednostavnosti ćemo se ograničiti na neke određene slučajeve.
Prvo razmatramo graničnu raspodelu jedne statistike reda $X_{k,N(t)}$ za fiksirano $k \in \mathbb{N}$. Pretpostavljamo da $F \in MDA(H)$, to jest (3.13) je zadovoljeno za odgovarajuće konstante $c_n > 0$ i $d_n \in \mathbb{R}$. Iz Tvrđenja 2.3.2 znamo da je (3.13) ekvivalentno sa

$$\lim_{n\to\infty} n\bar{F}(c_n x + d_n) = -\ln H(x), x \in \mathbb{R}. \tag{3.19}$$

Pod uslovom (3.19) sledi da za svako $k \in \mathbb{N}$ važi relacija

$$\lim_{n\to\infty} P\left(c_n^{-1}\left(X_{k,n} - d_n\right) \le x\right) = \Gamma_{\mathrm{k}}(-\ln \mathrm{H(x)}), x \in \mathbb{R},$$

gde Γ_{k} označava nepotpunu gama funkciju (videti Posledicu 3.2.3).

Teorema 3.3.2 (Granična raspodela k-te statistike višeg reda u slučajno indeksiranom uzorku)

Pretpostavimo da $N(t)/t \xrightarrow{P} Z$ važi za nenegativnu slučajnu promenljivu Z sa funkcijom raspodele F_Z i da je (3.19) zadovoljeno. Tada

$$\lim_{n\to\infty} P\left(c_n^{-1}\left(X_{k,N(n)} - d_n\right) \le x\right) = \int_0^\infty \Gamma_{\mathrm{k}}(-\mathrm{z}\ln \mathrm{H(x)})\mathrm{dF_Z(z)} \tag{3.20}$$

$$= E[\Gamma_{\mathrm{k}}(-\ln \mathrm{H^Z(x)})], \quad x \in \mathbb{R}.$$

□

Teorema 3.3.3 (Granična raspodela za vektor slučajno indeksirane statistike višeg reda)

Pretpostavimo da $\frac{N(t)}{t} \xrightarrow{P} \lambda$ važi za pozitivnu konstantu λ i da $F \in MDA(H)$ za raspodelu ekstremne vrednosti H tako da je (3.13) zadovoljeno. Tada

$$\left(c_n^{-1}\left(X_{i,N(n)} - d_n\right)\right)_{i=1,\dots,k} \xrightarrow{d} \left(Y_\lambda^{(i)}\right)_{i=1,\dots,k}$$

gde $\left(Y_\lambda^{(1)}, \dots, Y_\lambda^{(k)}\right)$ označava k -dimenzionalno obeležje koje odgovara raspodeli ekstremne vrednosti H^λ. Specijalno,

$$\lim_{n\to\infty} P\left(c_n^{-1}\left(X_{k,n} - d_n\right) \le x\right) = \Gamma_{\mathrm{k}}\left(-\ln \mathrm{H}^\lambda(x)\right), x \in \mathbb{R}.$$ □

3.4 Teorija ekstremne vrednosti za stacionarne nizove

Jedno od prirodnih uopštenja iid niza je strogo stacionaran proces. Kažemo da je niz slučajnih promenljivih (X_n) strogo stacionaran ako su njegove konačno dimenzionalne raspodele invarijantne u odnosu na promenu vremena, to jest

$$\left(X_{t_1}, \dots, X_{t_m}\right) \stackrel{\mathrm{def}}{=} \left(X_{t_1+h}, \dots, X_{t_m+h}\right)$$

za bilo koji izbor indeksa $t_1 < \cdots < t_m$ i celih brojeva h.

Uobičajeno je da se (X_n) definiše skupom indeksa $\mathbb{Z}$. (X_n) možemo da posmatramo kao vremenski niz posmatranja diskretnih ekvidistantnih vremenskih trenutaka gde je raspodela bloka $(X_t, X_{t+1}, \dots, X_{t+h})$ dužine h jednaka za sve cele brojeve t.

Dalje ćemo koristiti izraz stacionaran niz umesto strogo stacionaran niz. Strogo stacionaran niz je takođe stacionaran u širem smislu, što znači da je drugi moment od

$X = X_0$ konačan, to jest $E(X_n) = E(X)$ za svako n i $Cov(X_n, X_m) = Cov\left(X_0, X_{|n-m|}\right)$ za svako $n\ i\ m$.
Nemoguće je postaviti opštu teoriju ekstremnih vrednosti za klasu svih stacionarnih nizova. Zaista, treba da se odredi struktura zavisnosti od (X_n). Na primer, pretpostavimo da je $X_n = X$ za svako n. Ova relacija definiše stacionaran niz i

$$P(M_n \leq x) = P(X \leq x) = F(x), \quad x \in \mathbb{R}.$$

Prema tome, raspodela maksimuma uzorka može biti bilo koja raspodela F, što nije realna osnova za opštu teoriju.

Drugi ekstrem stacionarnog niza se dešava kad su X_n međusobno nezavisni, to jest (X_n) je iid niz. Pod tim uslovom smo razmatrali slabo granično ponašanje statistika višeg reda u poglavlju 3.2. Specijalno, znamo da postoje samo tri tipa različitih graničnih zakona: Frechetova raspodela Φ_α, Weibullova raspodela ψ_α i Gumbelova raspodela Λ (Fisher-Tipettova Teorema 2.2.7). Funkcije raspodele tipa $\Phi_\alpha, \psi_\alpha,\ \Lambda$ se nazivaju raspodelama ekstremnih vrednosti. U ovom poglavlju dajemo uslove za stacionaran niz (X_n) koji osigurava da njegove maksimume uzorka (M_n) i odgovarajuće maksimume $\left(\widetilde{M}_n\right)$ iid niza $\left(\tilde{X}_n\right)$ sa zajedničkom funkcijom raspodele $F(x) = P\left(\tilde{X} \leq x\right)$ pokazuju slično granično ponašanje. $\left(\tilde{X}_n\right)$ nazivamo iid nizom koji je asociran sa (X_n) ili jednostavno asociran iid niz. Kao ranije, pišemo $F \in MDA(H)$ za bilo koju od raspodela ekstremnih vrednosti H ako postoje konstante $c_n > 0$ i $d_n \in \mathbb{R}$ takve da

$$c_n^{-1}\left(\widetilde{M}_n - d_n\right) \xrightarrow{d} H.$$

Za izvođenje granične verovatnoće koristimo da je

$$P\left(\widetilde{M}_n \leq u_n\right) = P^n\left(\tilde{X} \leq u_n\right) = \tag{3.21}$$
$$= exp\left\{n \ln\left(1 - P\left(\tilde{X} \geq u_n\right)\right)\right\} \approx exp\{-n\bar{F}(u_n)\}.$$

Specijalno, u Tvrđenju 2.1.1 smo zaključili da za bilo koje $\tau \in [0, \infty)$
$P\left(\widetilde{M}_n \leq u_n\right) \to exp\{-\tau\}$ ako i samo ako $n\bar{F}(u_n) \to \tau \in [0, \infty)$. Jasno je da ne možemo direktno da primenimo (3.21) na maksimume zavisnih stacionarnih nizova. Da bismo prevazišli ovaj problem, pretpostavićemo da postoji određeni tip asimptotske nezavisnosti.

Uslov $\boldsymbol{D(u_n)}$: Za bilo koje cele brojeve $p, q\ i\ n$

$$1 \leq i_1 < \cdots < i_p < j_1 < \cdots < j_q \leq n$$

takve da $j_1 - i_p \geq l$ važi

$$\left|P\left(\max_{i \in A_1 \cup A_1} X_i \leq u_n\right) - P\left(\max_{i \in A_1} X_i \leq u_n\right) P\left(\max_{i \in A_2} X_i \leq u_n\right)\right| \leq \alpha_{n,l}$$

gde je $A_1 = \{i_1, \dots, i_p\}$, $A_2 = \{j_1, \dots, j_q\}\ i\ \alpha_{n,l} \to 0$ kad $n \to \infty$ za neki niz $l = l_n = o(n)$.

Uslov $D(u_n)$ je uslov mešanja raspodela i implicira da

$$P(M_n \leq u_n) = P^k\left(M_{[n/k]} \leq u_n\right) + o(1) \tag{3.22}$$

za konstantno ili sporo rastuće k. Ova relacija implicira da su granično ponašanje nizova (M_n) i $(\widetilde{M}_n)$ u uskoj vezi. Sledeća teorema pokazuje da klase mogućih graničnih zakona za normalizovane i centrirane nizove (M_n) i $\widetilde{M}_n$ koincidiraju.

Teorema 3.4.1 (Granični zakoni za maksimume stacionarnog niza)

Pretpostavimo da $c_n^{-1}(M_n - d_n) \xrightarrow{d} G$ za neku raspodelu G i odgovarajuće konstante $c_n > 0$, $d_n \in \mathbb{R}$. Ako uslov $D(c_n x + d_n)$ važi za svako realno x, tada je G raspodela ekstremne vrednosti. □

Dokaz: Iz Teoreme 2.2.2, 2.2.3 i Definicije 2.2.6 znamo da je G raspodela ekstremne vrednosti ako i samo ako je G max-stabilna.

Na osnovu (3.22)

$$P(M_{nk} \le c_n x + d_n) = P^k(M_{nk} \le c_n x + d_n) + o(1) \to G^k(x),$$

za sve cele brojeve $k \ge 1$ i svaku tačku x u kojoj je G neprekidno. S druge strane,

$$P(M_{nk} \le c_{nk} x + d_{nk}) \to G(x).$$

Dalje možemo da nastavimo kao u Teoremi 2.2.2 da bismo došli do zaključka da je G max-stabilna. □

Prema tome, max-stabilnost graničnih raspodela je potrebna pod uslovima $D(c_n x + d_n)$, $x \in \mathbb{R}$. Dalje želimo da nađemo dovoljne uslove za konvergenciju verovatnoća $P(M_n \le u_n)$ za dat niz pragova (u_n) koji zadovoljava

$$n\bar{F}(u_n) \to \tau \tag{3.23}$$

za neko $\tau \in [0, \infty)$. Iz Tvđenja 2.1.1 znamo da su (3.23) i $P(\widetilde{M}_n \le u_n) \to exp\{-\tau\}$ ekvivalentni. Međutim, da li možemo da zamenimo $(\widetilde{M}_n)$ sa (M_n) pod uslovom $D(u_n)$? Odgovor je ne. Jedino što možemo da izvedemo je

$$liminf_{n\to\infty} P(M_n \le u_n) \ge e^{-\tau}.$$

Uslov $\boldsymbol{D'(u_n)}$: $\lim_{k\to\infty} \limsup_{n\to\infty} n \sum_{j=2}^{[n/k]} P\left(X_1 > u_n,\ \ X_j > u_n\right) = 0.$

Primedba: $D'(u_n)$ je „anti-klastering uslov“ za stacionarne nizove (X_n). On implicira da

$$E\left(\sum_{1\le i<j\le [n/k]} I_{\{X_i > u_n,\, X_j > u_n\}}\right) \le [n/k] \sum_{j=2}^{[n/k]} E\left(I_{\{X_i > u_n,\, X_j > u_n\}}\right)\Big) \to 0,$$

tako da, u očekivanju, zajednički prelasci preko praga u_n parovima (X_i, X_j) postanu malo verovatni za veliko n.

Tvrđenje 3.4.2 (Granične verovatnoće za uzoračke maksimume)

Pretpostavimo da stacionarni niz (X_n) i niz pragova (u_n) zadovoljava uslove $D(u_n)$ i $D'(u_n)$. Pretpostavimo $t \in [0, \infty)$. Tada uslov (3.23) važi ako i samo ako

$$\lim_{n\to\infty} P(M_n \le u_n) = e^{-\tau}. \tag{3.24}$$

□

Dokaz Ograničićemo se na dokazivanje dovoljnog uslova da bismo pokazali primenu uslova $D(u_n) i\, D'(u_n)$. Potreban uslov se dokazuje na sličan način.
Za bilo koje $l \ge 1$,

$$\sum_{i=1}^{l} P(X_i > u_n) - \sum_{1\le i<j\le l} P(X_i > u_n, X_j > u_n) \le \tag{3.25}$$
$$\le P(M_l > u_n) \le \sum_{i=1}^{l} P(X_i > u_n).$$

Koristeći stacionarnost od (X_n) vidimo:

$$\sum_{i=1}^{l} P(X_i > u_n) = l\bar{F}(u_n),$$
$$\sum_{1\le i<j\le l} P(X_i > u_n,\ X_j > u_n) \le l \sum_{j=2}^{l} P(X_1 > u_n,\ X_j > u_n\ ).$$

Kombinovanjem ovoga i (3.25) za $l = [n/k]$ ($[x]$ označava ceo deo od x) i za fiksirano k dobijamo gornje i donje ocene za $P\big(M_{[n/k]} \le u_n\big)$:

$$1 - [n/k]\bar{F}(u_n) \le P\big(M_{[n/k]} \le u_n\big) \le 1 - [n/k]\bar{F}(u_n) + [n/k] \sum_{j=2}^{[n/k]} P\big(X_1 > u_n,\ X_j > u_n\big).$$

Iz (3.23) odmah dobijamo

$$[n/k]\bar{F}(u_n) \to \tau/_k,\ n \to \infty,$$

i na osnovu uslova $D'(u_n)$

$$\lim\ \sup_{n\to\infty} [n/k] \sum_{j=2}^{[n/k]} P\big(X_i > u_n,\ X_j > u_n\big) = o(1/k), \quad k \to \infty\ ,$$

pa dobijemo granice

$$1 - \frac{\tau}{k} \le \liminf_{n\to\infty} P\big(M_{[n/k]} \le u_n\big) \le \lim\ \sup_{n\to\infty} P\big(M_{[n/k]} \le u_n\big) \le 1 - \frac{\tau}{k} + o(1/k).$$

Ovo i relacija (3.22) impliciraju da

$$\left(1 - \frac{\tau}{k}\right)^k \liminf_{n\to\infty} P(M_n \le u_n) \le \lim\ \sup_{n\to\infty} P(M_n \le u_n) \le \left(1 - \frac{\tau}{k} + o(1/k)\right)^k.$$

Kada pustimo $k \to \infty$, vidimo da

$$\lim_{n\to\infty} P(M_n \le u_n) = e^{-\tau}.$$

□

Teorema 3.4.3 (Granična raspodela maksimuma stacionarnog niza)
Neka je (X_n) stacionaran niz sa zajedničkom funkcijom raspodele $F \in MDA(H)$ za neku raspodelu ekstremnih vrednosti H, to jest neka postoje konstane $c_n > 0, d_n \in \mathbb{R}$ tako da

$$\lim_{n\to\infty} n\bar{F}(c_n x + d_n) = -\ln H(x), x \in \mathbb{R}. \tag{3.26}$$

Pretpostavimo da za $x \in \mathbb{R}$ nizovi $(u_n) = (c_n x + d_n)$ zadovoljavaju uslove $D(u_n)\ i\ D'(u_n)$

Tada je (3.26) ekvivalentno sa bilo kojom od naredne dve relacije

$$c_n^{-1}(M_n - d_n) \xrightarrow{d} H \tag{3.27}$$

$$c_n^{-1}\left(\widetilde{M}_n - d_n\right) \xrightarrow{d} H \tag{3.28}$$

□

Dokaz Ekvivalencija (3.26) i (3.28) sledi iz Tvrđenja 2.3.2, a ekvivalencija (3.26) i (3.27) sledi iz Tvrđenja 3.4.2 □

4. Najduži niz uspeha

Da bismo što bolje objasnili pojam najdužeg niza uspeha, pre nego što damo matematičku analizu problema, daćemo dva jednostavna primera.

Prvi primer se tiče najdužeg niza uspeha prilikom bacanja novčića. Učenici jednog razreda su podeljeni u dve grupe. Svakom detetu iz prve grupe je dat novčić koji treba da baci dve stotine puta i da beleži rezultate nizova pismo i glava na papir. U drugoj grupi deca ne dobijaju novčić, umesto toga im je rečeno da pokušaju da napišu slučajan niz pismo i glava dužine dve stotine. Nakon prikupljanja ovih papira, trebalo je da ih podelimo u originalne grupe. U većini slučajeva je u tome postignut uspeh. Primećeno je da se nakon bacanja novčića u slučajnom nizu dužine dve stotine pojavljuju, npr. nizovi glava dužine sedam. S druge strane, primećeno je da većina dece koja je imala zadatak da zapiše slučajan niz, obično ima strah da napiše niz duži od četiri. Prema tome, da bismo pronašli izuzetke kod grupe koja je bacala novčić, jednostavno biramo one kod kojih su nizovi dužine veće od pet.
Iz ovog eksperimenta proizilazi pitanje - koja je dužina najdužeg niza glava u n Bernoullijevih pokušaja?

Drugi primer proizilazi iz razgovora sa mladim kockarskim amaterom na temu ruleta. On tvrdi da je zajedno sa prijateljima zabeležio rezultat svih rulet igara u svakom kazinu u periodu od tri godine. Očigledno ih je interesovala šema koju bi kasnije mogli da koriste kako bi pobedili sistem. Rečeno je da su jedno veče zabeležili šesnaest puta crveno za redom, što je naizgled suprotno u odnosu na pretpostavku o slučajnosti u ruletu.

Razmatramo veoma jednostavan model: Slučajne promenljive $X_1, X_2, \ldots$ su nezavisne Bernoullijeve promenljive sa istom raspodelom i verovatnoćom uspeha p, tj.

$$P(X=1)=p, \quad P(X=0)=q=1-p,$$

za neko $p\epsilon(0,1)$. Niz uspeha jedinica dužine j u $X_1, \ldots, X_n$ je definisan kao podniz $(X_{i+1}, \ldots, X_{i+j})$ niza $(X_1, \ldots, X_n)$ tako da je

$X_i = 0,\ X_{i+1} = \cdots = X_{i+j} = 1,\ X_{i+j+1} = 0,$

gde uzimamo da je $X_0 = X_{n+1} = 0$.

Zanima nas koliko je dugačak najduži niz uspeha jedinica u $X_1, \ldots, X_n$.

Alternativna formulacija je data slučajnim hodom (S_n) generisanim sa (X_n):

$$S_0 = 0,\ S_n = X_1 + \cdots + X_n.$$

Neka je

$$I_n(j) = \max_{0 \le i \le n-j}(S_{i+1} - S_i), \quad 1 \le j \le n,$$

i Z_n najveći ceo broj takav da je $I_n(Z_n) = Z_n$. Tada je Z_n dužina najdužeg niza uspeha jedinica u $X_1, \ldots, X_n$.

0	0	0	1	1	0	1	1	0	1	1	1	1	0	0	1	0	1	0	1
0	1	1	1	1	1	0	1	0	1	0	0	0	0	1	1	0	1	1	1
0	1	1	1	1	0	1	0	1	1	1	0	1	0	0	0	1	1	1	1
0	1	1	1	1	1	0	0	0	0	1	0	0	1	1	0	1	0	0	0
1	1	0	1	1	1	0	1	1	0	1	1	1	1	0	0	1	1	1	0
0	1	0	0	0	0	0	0	1	0	1	0	1	1	1	0	0	1	0	0
0	1	1	1	0	0	1	1	0	0	1	1	1	0	1	1	0	0	1	1
0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0
1	1	0	1	0	1	1	1	1	1	0	0	0	0	1	1	0	1	1	0
1	0	1	1	1	1	0	0	1	1	0	0	1	1	1	1	1	1	1	0
0	1	1	1	0	0	1	0	1	0	0	0	0	0	1	1	0	1	1	1
0	0	1	0	1	1	1	1	0	1	1	1	0	0	0	0	1	0	1	0
0	1	1	0	0	1	1	1	0	0	1	1	1	1	0	1	0	1	0	0
0	0	0	0	1	0	1	1	0	0	0	1	1	1	1	1	0	0	0	1
1	1	1	1	1	1	0	1	1	0	0	0	0	0	1	0	0	0	1	1
1	1	0	0	1	0	1	0	0	0	0	0	1	1	0	0	1	1	0	0
0	0	1	0	0	1	1	1	1	0	0	0	1	0	1	1	0	1	1	0
0	0	1	0	0	0	0	1	0	1	0	1	1	1	1	0	1	0	0	1
0	1	0	0	1	0	1	1	0	1	1	1	0	1	0	1	0	1	1	1
1	1	0	1	0	1	1	1	1	1	0	1	0	0	0	1	0	1	0	0
1	0	1	0	1	1	0	0	0	1	0	1	0	1	0	1	0	0	0	0
1	1	0	1	1	0	0	1	0	1	0	0	0	0	0	1	0	1	1	0
0	0	0	1	1	1	0	1	1	1	0	1	1	1	1	0	1	0	0	0
1	1	1	0	0	1	1	0	0	1	0	0	0	1	1	0	1	1	0	0
1	0	1	0	1	0	0	0	0	1	1	1	1	0	0	1	1	0	0	0
0	0	0	1	0	1	0	0	1	0	1	1	1	0	1	1	0	0	1	1
0	0	0	1	1	0	1	0	0	0	0	0	1	0	1	1	0	0	0	1
1	0	1	1	0	0	0	0	0	0	1	0	0	1	1	1	0	0	0	1
1	1	0	1	1	1	0	0	1	0	0	1	1	1	0	0	1	1	1	1
1	1	1	0	1	1	1	1	0	0	0	1	0	1	0	0	0	1	0	1

Tabela 4. Slučajan niz 600 vrednosti 0 i 1. Obe vrednosti se pojavljuju sa jednakom verovatnoćom. Najduži niz jedinica je dužine 7.

Primer 4.1.1 (Najduži niz uspeha u osiguranju)
Neka su Y_i nezavisne slučajne promenljive sa istom raspodelom koje označavaju broj zahteva u nekom portfoliju i neka je $u \geq 0$ dati prag. Uvodimo Bernoullijeve slučajne promenljive

$$X_i = I_{\{Y_i > u\}}, \qquad i \geq 1,$$

sa verovatnoćom uspeha $p = P\{Y_i > u\}$. Najduži niz uspeha jedinica u Bernoullijevom nizu $\{X_i\}$ odgovara najdužem uzastopnom nizu prekoračenja praga u. □

Primer 4.1.2 (Najduži niz uspeha u finansijama)
Standardni Cox-Ross-Rubinstein model u finansijama pretpostavlja da vrednost rizične aktive ili raste sa verovatnoćom p, ili opada sa verovatnoćom $1 - p$. Rezultirajući model

binomnog stabla služi kao skelet za mnoge naprednije modele, uključujući Black-Scholes model. Niz uspeha jedinica u ovoj postavci odgovara uzastopnom porastu cena rizične aktive. □

Primer 4.1.3 (Tačna raspodela Z_n za simetričan slučajan hod)
Simetričan slučajan hod odgovara verovatnoći $p = 0.5$. Poznato je da je precizna raspodela za najveći ceo broj Z_n takav da važi $Z_n = I_n(Z_n)$ za simetričan slučajan hod. Za svako $j = 1, \dots, n$,

$$P(Z_n < j) = \frac{1}{2^n} \sum_{i=1}^{n+1} \sum_{k=0}^{i} (-1)^k \binom{i}{k} \binom{n-kj}{i-1}, \mathrm{j} = 1, \dots, \mathrm{n} \qquad (4.1)$$

Formula (4.1) je ograničene vrednosti, s obzirom na to da može biti primenjena pravilno samo za male n. Vreme koje potrebno da kompjuter oceni formulu (4.1) izuzetno brzo raste sa n. Numeričke aproksimacije binomnih izraza Stirlingovom formulom ili drugim metodama ne daju zadovoljavajuće odgovore. □

4.1 Totalna udaljenost varijacije u odnosu na Poissonovu raspodelu

S obzirom na to da se slučajna šetnja S_n sastoji od binomnih slučajnih promenljivih, Poissonova aproksimacija je opravdana. Primenjuje se Stein-Chen metod da bi se izvele granice totalne udaljenosti varijacije između zakona raspodele F_{W_j} od

$$W_j = \sum_{i=1}^{n} I_{\{X_i=0, X_{i+1}=\cdots=X_{i+j}=1, X_{i+j+1}=0\}}$$

i od Poissonove raspodele sa parametrom

$$\lambda = nq^2p^j = E(W_j).$$

W_j uključuje sve nizove uspeha jedinica dužine j. Totalna udaljenost varijacije između dve raspodele F_1 i F_2 za nenegativne cele brojeve je data sa

$$d_{TV}(F_1, F_2) = sup_{A\subseteq N_0}|F_1(A) - F_2(A)| = \frac{1}{2}\sum_{k=0}^{\infty}|F_1(\{k\}) - F_2(\{k\})|.$$

Gornja granica za $d_{TV}(F_{W_j}, \mathcal{P}(\lambda))$ takođe daje ocenu pojedinačnih udaljenosti

$$\left|P(W_j = k) - e^{-\lambda}\frac{\lambda^k}{k!}\right|, k \geq 0,$$

$$d_{TV}\left(F_{W_j}, \mathcal{P}(nq^2p^j)\right) \leq \big((2j-1)q+2\big)qp^j, \; n > 2j+2.$$

Prema tome,

$$sup_{k\geq 0}\left|P(W_j = k) - e^{-\lambda}\frac{\lambda^k}{k!}\right| = d_{TV}(F_1, F_2) =$$

$$= sup_{k\geq 0}\left|P(postoji\ tačno\ k\ nizova\ dužine\ j) - e^{-\lambda}\frac{\lambda^k}{k!}\right| \leq ((2j-1)q+2)qp^j. \qquad (4.2)$$

Ova ocena daje korisne informacije kada je desna strana nejednakosti mala u poređenju sa Poissonovim verovatnoćama na levoj strani nejednakosti.

4.2 Skoro sigurno ponašanje

Teorema 4.2.1 (Skoro siguran rast dužine Z_n najdužeg niza jedinica)
Za svako fiksirano $p \in (0,1)$,

$$\lim_{n\to\infty} \frac{Z_n}{\ln n} = \frac{1}{-\ln p} \quad \text{s.s.} \tag{4.3}$$

□

Pokazaćemo kako se rezultat ove teoreme može pokazati klasičnom teorijom graničnih vrednosti. Najduži niz jedinica je otprilike reda $-\frac{\ln n}{\ln p}$, tj. raste sporo po n. (Tabela 5). Postavlja se pitanje odakle dolazi logaritamska normalizacija u jakom zakonu velikih brojeva. Osnovna ideja je

$L_1 = min\{n \geq 1 : X_n = 0\}$,

$L_k = min\{n : n > L_1 + \cdots + L_{k-1}, X_n = 0\} - (L_1 + \cdots + L_{k-1})$, za $k \geq 2$.

Pošto su X_n nezavisne slučajne promenljive sa jednakom raspodelom, iz osobine procesa Markova S_n, zaključujemo da su L_n nezavisne pozitivne slučajne promenljive sa jednakom raspodelom. Nije teško videti da je

$P(L_1 = k) = qp^{k-1}$, $k \geq 1$.

Prema tome, L_1 ima geometrijsku raspodelu. Po konstrukciji, vrednosti $L_{i-1} \geq 1$ su dužine nizova jedinica i slučajna promenljiva

$$N(n) = card\{m : \textstyle\sum_{i=1}^{m} L_i \leq n\}$$

računa broj nula između $X_1, \ldots, X_n$. Prema tome, ona ima binomnu raspodelu sa parametrima (n, q). Dalje, u cilju određivanja dužine najdužeg niza jedinica u $X_1, \ldots, X_n$, treba da proučimo maksimume geometrijskih slučajnih promenljivih sa identičnim raspodelama duž slučajno indeksiranog niza. Zaista,

$$\max_{i\leq N(n)} L_i - 1 \leq Z_n \leq \max_{i\leq N(n)+1} L_i - 1. \tag{4.4}$$

Za dokaz teoreme, treba nam sledeće pomoćno tvrđenje.

Tvrđenje 4.2.2 (Karakterizacija minimalnog i maksimalnog skoro sigurnog rasta maksimuma geometrijskih slučajnih promenljivih)

Neka je d_n rastući niz pozitivnih celih brojeva takvih da $d_n \longrightarrow \infty$.

a) Relacija $P(\max_{i\leq n} L_i > d_n \quad i.o.) = 0$ važi ako je $\sum_{n=1}^{\infty} p^{d_n} < \infty$,
ili
relacija $P(\max_{i\leq n} L_i > d_n \quad i.o.) = 1$ važi ako je $\sum_{n=1}^{\infty} p^{d_n} = \infty$.

b) Dodatno pretpostavimo da $np^{d_n} \longrightarrow \infty$. Tada je
$P(\max_{i\le n} L_i \le d_n \quad i.o.) = 0$ ako je $\sum_{n=1}^{\infty} p^{d_n} exp\{-np^{d_n}\} < \infty$,
Ili
$P(\max_{i\le n} L_i \le d_n \quad i.o.) = 1$ ako je $\sum_{n=1}^{\infty} p^{d_n} exp\{-np^{d_n}\} = \infty$.

Štaviše, ako je $lim\ inf_{n\to\infty} np^{d_n} < \infty$, onda
$P(\max_{i\le n} L_i \le d_n \quad i.o) = 1$.

c) a) i b) ostaju važeći ako se n svuda zameni sa $[nc]$, gde $[x]$ označava ceo deo od x i c je pozitivna konstanta.

□

Restrikcija na nizove celih brojeva d_n je prirodna, pošto su u pitanju celobrojne slučajne promenljive i događaji oblika
$$\{\max_{i\le n} L_i \le x\} = \{\max_{i\le n} L_i \le [x]\}.$$

Dokaz Teoreme 4.2.1 je sada posledica prethodnog tvrđenja i relacije (4.4). Međutim, i dalje postoji manji problem: iz tvrđenja dobijamo samo
$$\max_{i\le n} \frac{L_i}{\ln n} \xrightarrow{s.s.} -\frac{1}{\ln p}.$$
Prema tome, treba da zamenimo n slučajnim indeksom $N(n)$ ili $N(n)+1$, što se može uraditi na osnovu Leme 1.5.3 i činjenice da $N(n)$ može da se protumači kao obnavljajući proces prebrajanja koji posmatramo u celobrojnim vremenskim trenucima, pa je $N(n)/n \xrightarrow{s.s.} q$ (Teorema 1.5.9). Primenjujući slične tehnike, rezultati se mogu pokazati preciznije nego Teoremom 4.2.1. Postoje i rezultati o dužini nizova jedinica koji su prekinuti datim brojem nula. Koristimo notaciju:

$ln_0 x = x, ln_1 x = \max(0, \ln(ln_0 x))\, , ln_k x = \max\,(0, \ln\,(ln_{k-1} x)),\ x > 0,\ k \ge 2.$

Teorema 4.2.3 (Skoro sigurno ponašanje dužine najdužeg niza jedinica)

a) Za svako $r \in N$ važi sledeća relacija:
$$P\left(Z_n > \left[\frac{ln_1(nq)+\cdots+ln_r(nq)+\varepsilon ln_r(nq)}{-\ln p}\right] \quad i.o.\right) = \begin{cases} 0, & \varepsilon > 0, \\ 1, & \varepsilon < 0. \end{cases}$$

b) Za svako **ε**> 0,
$$P\left(Z_n \le \left[\frac{ln(nq)-ln_3(nq)-\varepsilon}{-\ln p}\right] - 1 \quad i.o.\right) = 0,$$
$$P\left(Z_n \le \left[\frac{ln(nq)-ln_3(nq)}{-\ln p}\right] + 1 \quad i.o.\right) = 1.$$

Dokaz. Za fiksirano $r \ge 1$, $c > 0$ i malo **ε** pišemo
$$b_n(\varepsilon, c) = \left[\frac{ln_1(nc)+\cdots+ln_r(nc)+\varepsilon ln_r(nc)}{-\ln p}\right].$$

Onda se lako proverava da je

$$\sum_{n=1}^{\infty} p^{b_n(\varepsilon,c)} < \infty \text{ ili } = \infty$$

kada je $\varepsilon > 0$ ili $\varepsilon \leq 0$.
Pa, prema Tvrđenju 4.2.2,

$$P\left(\max_{i \leq [nc]} L_i - 1 > b_n(\varepsilon, c) \quad i.o.\right) = 0 \; ili = 1 \tag{4.5}$$

kada je $\varepsilon > 0$ ili $\varepsilon \leq 0$.

Pošto $N(n)/n \xrightarrow{s.s.} q$, sledi da je za svako malo fiksirano $\delta > 0$ i za veliko n, sa verovatnoćom 1,

$$n(1-\delta)q \leq N(n) \leq n(1+\delta)q - 1. \tag{4.6}$$

Takođe primetimo da za veliko n,

$$b_n(\varepsilon/2, q(1+\delta)) \leq b_n(\varepsilon, q). \tag{4.7}$$

Zbog (4.6), (4.7) i (4.4) dobijamo

$$P(\max_{i \leq n(1+\delta)q} L_i - 1 > b_n\left(\frac{\varepsilon}{2}, (1+\delta)q\right) \quad i.o.) \geq P(Z_n > b_n(\varepsilon, q) \quad i.o.).$$

Ovo zajedno sa (4.5) pokazuje prvi deo teoreme za $\varepsilon > 0$. Za $\varepsilon < 0$ se može pokazati na sličan način. Nastavljamo slično u drugom delu. Za $\varepsilon > 0$, stavimo $\varepsilon' = \ln(1+\varepsilon)$. Pišemo, za $\varepsilon > 0$, $c > 0$ i za dovoljno veliko n,

$$b'_n(\varepsilon, c) = \left[\frac{ln(nc) - ln_3(nc) - \varepsilon}{-\ln p}\right] = \left[\frac{ln(nc) - \ln\left((1+\varepsilon')ln_2(nc)\right)}{-\ln p}\right]$$

i razlomački deo

$$z_n = \left\{\frac{\ln(nc) - ln_3(nc) - \varepsilon}{-lnp}\right\}.$$

Tada, prema $P(L_1 > x) = P(L_1 > [x]) = p^{[x]} = p^{x - \{x\}}$, za veliko n,

$$p^{b'_n(\varepsilon,c)} exp\left\{-[nc]p^{b'_n(\varepsilon,c)}\right\} = \frac{ln_2(nc)}{nc}(1+\varepsilon')p^{-Z_n} exp\left\{-\frac{[nc]}{nc} ln_2(nc)(1+\varepsilon')p^{-Z_n}\right\}$$

$$\leq const \frac{ln_2 n}{n(\ln n)^{(1+\varepsilon'/2)}}.$$

Poslednji niz može da se sumira i sledi iz Tvrđenja 4.2.2 da

$$P(\max_{i \leq nc} L_i \leq b'_n(\varepsilon, c) \quad i.o.) = 0. \tag{4.8}$$

Slično, neka je

$$b''_n(c) = \left[\frac{\ln(nc) - ln_3(nc)}{-\ln p}\right] + 1$$

i

$$Z'_n = \left\{\frac{\ln(nc) - ln_3(nc)}{-\ln p}\right\}.$$

Onda

$p^{b''_n(c)} exp\{-[nc]p^{b''_n(c)}\} = \frac{ln_2(nc)}{nc} p^{1-Z'_n} exp\left\{-\frac{[nc]}{nc} ln_2(nc)p^{1-Z'_n}\right\} \geq const \frac{ln_2 n}{n \ln n}$.

Ovaj niz divergira kada se sumira, pa zbog Tvrđenja 4.2.2:

$P(\max_{i \leq nc} L_i \leq b''_n(c) \quad i.o.) = 1.$

Preostaje još da se prebacimo na slučajan niz indeksa $(N(n))$. Nastavljamo kao u prvom delu dokaza: biramo malo $\delta > 0$ i $\varepsilon'' > 0$ tako da za veliko n,

$$\frac{\ln(nq) - ln_3(nq) - \varepsilon}{-\ln p} \leq \frac{\ln\big(nq(1-\delta)\big) - ln_3\big(nq(1-\delta)\big) - \varepsilon''}{-\ln p}.$$

Tada prema (4.8),

$P(Z_n \leq b'_n(\varepsilon, q) - 1 \quad i.o.) \leq P\big(\max_{i \leq n(1-\delta)q} L_i \leq b'_n(\varepsilon'', (1-\delta)q) \quad i.o.\big) = 0.$

Može se slično nastaviti sa nizom $(b''_n(c))$, ali preskačemo detalje. Ovo dokazuje drugi deo teoreme. □

Ovi rezultati pokazuju suptilno skoro sigurno ponašanje dužine najdužeg niza jedinica.

Posledica 4.2.4
Za svako fiksirano $\varepsilon > 0$ i $r \in \mathbb{N}$, sa verovatnoćom 1 dužina najdužeg niza jedinica u $X_1, \ldots, X_n$ upada, za velike n, u interval $[\alpha_n, \beta_n]$, gde je

$$\alpha_n = \left[\frac{\ln(nq) - ln_3(nq) - \varepsilon}{-\ln p}\right] - 1,$$

$$\beta_n = \left[\frac{\ln(nq) + \cdots + ln_r(nq) + \varepsilon ln_r(nq)}{-\ln p}\right]. \qquad (4.9)$$

□

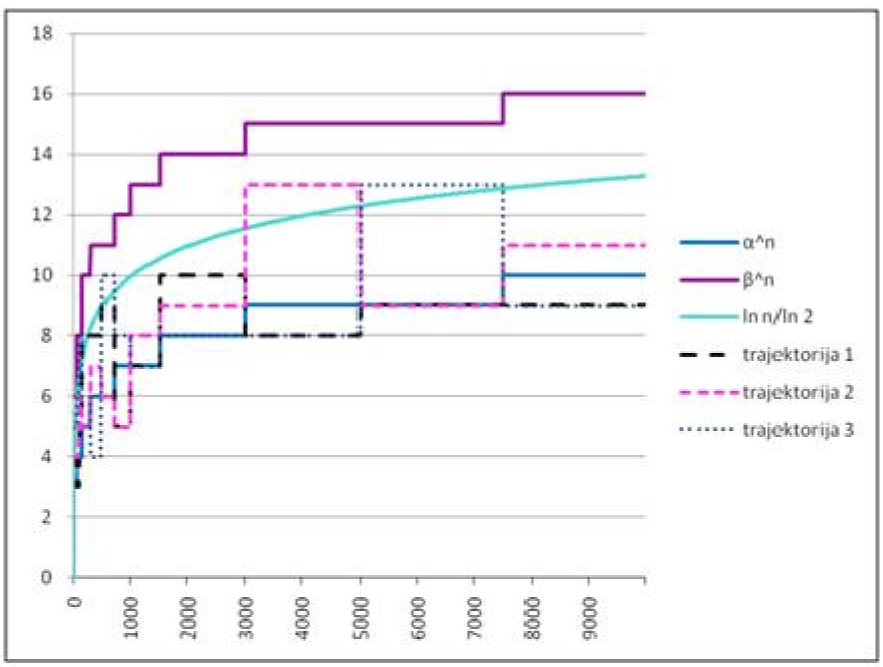
U Tabeli 5 upoređujemo stopu skoro sigurnog rasta za Z_n, tj. $-\ln n / \ln p$ sa gornjim i donjim granicama (4.9). Biramo $p = 1/2$ i $r = 3$, $\varepsilon = 0.001$ i granice

$$\alpha_n = \left[\frac{\ln(n/2) - ln_3(n/2) - 0.001}{\ln 2}\right] - 1,$$

. $$\beta_n = \left[\frac{\ln(n/2) - ln_2(n/2) + 1.001 ln_3(n/2)}{\ln 2}\right].$$

n	$ln\, n/\ln 2$	α_n	β_n
50	5,64	3	6
100	6,64	4	8
150	7,23	4	8
300	8,23	5	10
500	8,97	6	11
700	9,45	6	11
1000	9,97	7	12
1500	10,55	7	13
3000	11,55	8	14
5000	12,29	9	15
7500	12,87	9	15
10000	13,29	10	16

Tabela 5. *Skoro sigurne granice α_n i β_n za najduži niz Z_n jedinica u slučajnom hodu*



Grafik 2. *Tri simulirane trajektorije dužine Z_n najdužeg niza jedinica u slučajnoj šetnji*

4.3 Ponašanje raspodela

Iz primera 2.1.5 smo videli da maksimumi geometrijskih slučajnih promenljivih nemaju graničnu raspodelu, kakve god da su normalizujuće konstante i centriranja.
Međutim, rep geometrijske raspodele je veoma blizu repa eksponencijalne raspodele. Tačnije,

$$L_1 - 1 = [E_1/(-\ln p)] \tag{4.10}$$

za standardnu eksponencijalnu slučajnu promenljivu E_1, $[x]$ označava ceo deo od x.

Lako uočavamo:

$$\begin{aligned} qp^k &= P(L_1 - 1 = k) = p^k - p^{k+1} \\ &= P(E_1 \epsilon [k(-\ln p), (k+1)(-\ln p)]) \\ &= P(\left[\tfrac{E_1}{-\ln p}\right] = k). \end{aligned}$$

Neka su sad (E_n) nezavisne standardne eksponencijalne slučajne promenljive sa identičnim raspodelama. Iz primera 2.2.7 znamo,

$$\max(E_1, \dots, E_n) - \ln n \xrightarrow{d} \Lambda \tag{4.11}$$

gde je $\Lambda(x) = e^{-e^{-x}}$, $x \epsilon \mathbb{R}$, Gumbelova raspodela.
Kada znamo (4.4), možemo da se nadamo da raspodela Z_n nije daleko od Gumbelove raspodele.

Teorema 4.3.1 (Asimptotsko ponašanje raspodele dužine najdužeg niza jedinica)
Neka je Y slučajna promenljiva sa Gumbelovom raspodelom Λ. Onda je

$$sup_{k\epsilon\mathbb{Z}} \left| P\left(Z_n - \left[\frac{\ln(nq)}{-\ln p}\right] \le k \right) - P(\left[\frac{Y}{-\ln p} + \left\{\frac{\ln (nq)}{-\ln p}\right\}\right] \le k) \right| \to 0$$

Ovde $[x]$ označava ceo deo od x, a $\{x\}$ označava razlomački deo od x, tj. $\{x\} = x - [x]$.

Dokaz. Iz (4.11) i Leme 1.5.5 zaključujemo da

$$\frac{\max_{i \le N(n)} E_i}{-\ln p} - \frac{\ln N(n)}{-\ln p} \xrightarrow{d} \frac{Y}{-\ln p}, \tag{4.12}$$

ako je Anscombeov uslov zadovoljen i Y ima raspodelu Λ. Ovo možemo videti i iz sledećeg:

$$P(\max_{n(1-\delta)q < m \le n(1+\delta)q} \left| (\max_{i \le m} E_i - \ln m) - (\max_{i \le nq} E_i - \ln([nq])) \right| > \varepsilon)$$

$$\le P\left(\max_{i \le n(1+\delta)q} E_i - \max_{i \le n(1-\delta)q} E_i > \frac{\varepsilon}{2} \right) + I_{\left(\frac{\varepsilon}{2}, \infty\right)}(\ln (\frac{[n\_\_(1+\delta)q]}{[n(1-\delta)q]})) = p_1 + p_2.$$

Važi da je p_2 jednako 0 za malo δ i veliko n. Dalje,

$$p_1 = P(\max_{n(1-\delta)q < i \le n(1+\delta)q} E_i - \max_{i \le n(1-\delta)q} E_i > \varepsilon/2)$$

$= P(\max_{i \leq [n(1+\delta)q]-[n(1-\delta)q]} E_i^{'} - \ln([n(1+\delta)q] - [n(1-\delta)q]) - \max_{i \leq [n(1-\delta)q]} E_i - \ln([n(1-\delta)q]) > \frac{\varepsilon}{2} - \ln(\frac{[n(1+\delta)q]-[n(1-\delta)q]}{[n(1-\delta)q]})),$

gde je $(E_i^{'})$ nezavisna kopija od (E_i).
Na osnovu (4.11), vidimo da p_1 konvergira ka

$$P(Y_1 - Y_2 > \frac{\varepsilon}{2} - \ln(2\delta/(1-\delta)))$$

za nezavisne slučajne promenljive Y_i koje imaju Gumbelovu raspodelu. Sada, desna strana poslednje granične verovatnoće može da se napravi da bude mala, tako što biramo dovoljno malo δ. Prema tome, Anscombeov uslov je zadovoljen.
Pošto $N(n)/n \xrightarrow{s.s.} q$, možemo sada da zamenimo izraz $\ln N(n)/(-\ln p)$ u (4.12) sa $\ln(nq)/(-\ln p)$. Pošto je Λ neprekidna, odmah iz (4.12) dobijamo da

$sup_t \left| P\left(\frac{\max_{i \leq N(n)} E_i}{-\ln p} - \frac{\ln(nq)}{-\ln p} \leq t\right) - P(\frac{Y}{-\ln p} \leq t) \right| \to 0.$

To implicira da

$$sup_t \left| P\left(\left[\frac{\max\limits_{i \leq N(n)} E_i}{-\ln p}\right] - \left[\frac{\ln(nq)}{-\ln p}\right] \leq t\right) - P\left(\left[\frac{Y}{-\ln p} + \left\{\frac{ln(nq)}{-\ln p}\right\}\right] \leq t\right) \right|$$

$= sup_t \left| P\left(\max_{i \leq N(n)} \left[\frac{E_i}{-\ln p}\right] - \left[\frac{\ln(nq)}{-\ln p}\right] \leq t\right) - P(\left[\frac{Y}{-\ln p} + \left\{\frac{\ln(nq)}{-\ln p}\right\}\right] \leq t) \right| \to 0.$

S obzirom na to da su promenljive u posednjoj graničnoj relaciji celi brojevi, supremum za svako $t \in \mathbb{R}$ se smanjuje na supremum za cele brojeve. □

Iz Teoreme 4.3.1 možemo da zaključimo sledeće o asimptotskom ponašanju raspodela za Z_n:

$$P\left(Z_n \leq k + \left[\frac{\ln(nq)}{-\ln p}\right]\right) = P\left(\left[\frac{Y}{-\ln p} + \left\{\frac{\ln(nq)}{-\ln p}\right\}\right] \leq k\right) + o(1) =$$

$$= P\left(\frac{Y}{-\ln p} + \left\{\frac{\ln(nq)}{-\ln p}\right\} < k+1\right) + o(1) =$$

$$= P\left(Y < (k+1)(-\ln p) - \left\{\frac{\ln nq}{-\ln p}\right\}(-\ln p)\right) + o(1) =$$

$$= exp\left\{-p^{k+1-\{\ln(nq)/(-\ln p)\}}\right\} + o(1),$$

za sve cele brojeve $k \geq 1 - [\ln(nq)/(-\ln p)]$.
Specijalno, za k pozitivno,

$P(Z_n = k + \left[\frac{\ln(nq)}{-\ln p}\right]) \approx qp^{k-\{\ln(nq)/(-\ln p)\}}.$

Pitanja o najdužem nizu jedinica i sličnim problemima su privukla mnogo pažnje u literaturi. Primene leže, ne samo u teoriji ekstremne vrednosti, finansijama, osiguranju, već i u molekularnoj biologiji (najduži nizovi koji se poklapaju u lancima DNK),

prepoznavanju obrasca (najduži obrascu koji se ponavljaju u slučajnim nizovima) i slično.
Problem najdužeg niza jedinica je usko povezan sa pitanjem o redu veličine priraštaja generalnog slučajnog hoda (S_n). Zgodan alat za opis reda priraštaja je dat sa

$I_n(j) = \max_{0 \le i \le n-j}(S_{i+j} - S_i), \quad 1 \le j \le n.$

Za specijalan slučaj Bernoullijevih slučajnih promenljivih sa identičnim raspodelama, sa verovatnoćom uspeha $p \in (0,1)$, možemo da zaključimo iz Poglavlja 4.2 da je

$\lim_{n\to\infty} \frac{I_n\left(\left[\frac{\ln n}{-\ln p}\right]\right)}{\left[\frac{\ln n}{-\ln p}\right]} = 1$ skoro sigurno.

Ovaj rezultat već pokazuje tipičan red priraštaja za slučajan hod. Sada pretpostavimo da $E(X) = 0$ i da X ima funkciju generatrise momenta $M(h) = E(e^{hX})$.
Neka je $h_0 = sup\{h: M(h) < \infty\} \ge 0$ i definišemo broj

$$c = c(\alpha) \text{ sa } e^{-1/c} = inf_{h\in\mathbb{R}} e^{-h\alpha} M(h),\ \alpha > 0. \tag{4.13}$$

Lako je videti da ako $h_0 > 0$, onda infimum leži strogo između 0 i 1, pa je c pozitivno.

Teorema 4.3.2 (Erdos-Renyi jak zakon velikih brojeva za priraštaje slučajnog hoda)
Relacija

$\lim_{n\to\infty} \frac{I_{n([c \ln n])}}{[c \ln n]} = \alpha$ skoro sigurno

važi za svako

$\alpha \in \left\{\frac{M'(h)}{M(h)}: 0 < h < h_0\right\}.$

Ovde je $c = c(\alpha)$ dato jednačinom (4.13). □

Brojne generalizacije ove Teoreme su dokazane. One zavise u mnogom od tehnika velikih devijacija (zbog koje je potrebna pretpostavka o postojanju funkcije generatrise momenta) i od uopštenja teorije procesa obnavljanja.

Zaključak

Zakon velikih brojeva je zakon teorije verovatnoće koji opisuje rezultat vršenja istog eksperimenta veliki broj puta. Prema ovom zakonu, prosek rezultata dobijenih putem velikog broja pokušaja bi trebalo da bude blizu očekivane vrednosti i biće joj sve bliži kako se broj pokušaja bude povećavao.

Iz zakona velikih brojeva sledi da empirijska verovatnoća uspeha u nizu Bernoullijevih pokušaja konvergira teorijskoj verovatnoći. Za Bernoullijevu slučajnu promenljivu, očekivana vrednost je jednaka teorijskoj verovatnoći uspeha, a prosek takvih n promenljivih je baš jednak relativnoj frekvenciji, pod uslovom da su promenljive međusobno nezavisne i da imaju identičnu raspodelu.

Ovaj zakon je od izuzetnog značaja zato što garantuje određenu stabilnost dugoročnih rezultata za slučajne događaje. Na primer, iako kazino može da izgubi novac na jednom spinu ruleta, njegova zarada će težiti predviđenom procentu nakon velikog broja spinova. Svaki niz pobeda će na kraju biti prekinut parametrima igre.
Važno je imati na umu da zakon velikih brojeva važi samo kad se uzima u obzir veliki broj posmatranja. Ne postoji pravilo koje tvrdi da će mali broj posmatranja konvergirati očekivanoj vrednosti, ili da će jedan ishod odmah biti izbalansiran drugim.

Literatura

- S. I. Resnick, Heavy-tail phenomena: Probabilistic and Statistical Modeling, New York: Springer, 2007

- P. Embrechts, C. Klüppelberg, T. Mikosch, Modelling extremal events for Insurance and Finance, Berlin: Springer-Verlag, 2008

- P. Mladenović, Verovatnoća i statistika, Matematički fakultet, Beograd: Vesta 1995

- www.investopedia.com

- www.wikipedia.com

- www.emaths.co.uk

Biografija

Bojana Jovanić je rođena 27.11.1986. godine u Beogradu. Pohađala je osnovnu školu „Miroslav Antić“ na Paliću u periodu od 1993. do 2001. godine i završila je skroz odličnim uspehom, kao dobitnik Vukove diplome. Gimnaziju „Svetozar Marković“ pohađala je u Subotici u periodu od 2001. do 2005. godine i takođe je završila skroz odličnim uspehom, kao dobitnik Vukove diplome. Godine 2005. upisala je Prirodno-matematički fakultet u Novom Sadu, odsek za matematiku i informatiku, smer matematika finansija, gde je i diplomirala u septembru 2009. godine sa prosečnom ocenom 9,09.
Od septembra 2010. godine zaposlena u kompanij Leversys d.o.o. u Beogradu kao analitičar konvertibilnih obveznica.

Novi Sad, januar 2011. Bojana Jovanić

UNIVERZITET U NOVOM SADU

PRIRODNO-MATEMATIČKI FAKULTET,

DEPARTMAN ZA MATEMATIKU I INFORMATIKU

KLJUČNA DOKUMENTACIJSKA INFORMACIJA

Redni broj:
RBR

Identifikacioni broj:
IBR

Tip dokumentacije: Monografska dokumentacija
TD

Tip zapisa: Tekstualni štampani materijal
TZ

Vrsta rada: Master rad
VR

Autor: Bojana Jovanić
AU

Mentor: Docent dr Dora Seleši
MN

Naslov rada: Granične raspodele: zakoni velikih i malih brojeva
NR

Jezik publikacije: Srpski (latinica)
JP

Jezik izvoda: s/e
JI

Zemlja publikovanja: Srbija
ZP

Uže geografsko područje: Vojvodina
UGP

Godina: 2011.
GO

Izdavač: Autorski reprint
IZ

Mesto i adresa: Novi Sad, Departman za matematiku i informatiku, PMF, Trg Dositeja Obradovića 4
MA

Fizički opis rada: (4, 87, 0, 5, 2, 0, 0)
Broj poglavlja, broj strana, br. literarnih citata, br. tabela, br. slika, br. grafika, br priloga.
FO

Naučna oblast: Matematika
NO

Naučna disciplina: Aktuarska matematika
ND

Ključne reči: Zakon velikih brojeva, zakon malih brojeva, granična raspodela, centralna granična teorema, najduži niz uspeha, statistike reda, fluktuacije maksimuma, fluktuacije suma
PO

UDK:

Čuva se: U biblioteci Departmana za matematiku i informatiku
ČU

Važna napomena:
VN

Izvod:
IZ
Zakon velikih brojeva je zakon teorije verovatnoće koji opisuje rezultat vršenja istog eksperimenta veliki broj puta. Prema ovom zakonu, prosek rezultata dobijenih putem velikog broja pokušaja bi trebalo da bude blizu očekivane vrednosti i biće joj sve bliži kako se broj pokušaja bude povećavao.

Datum prihvatanja teme od strane NN veća: 16. 09. 2010.
DP

Datum odbrane:
DO

Članovi komisije:
KO
dr Dora Seleši, docent Prirodno-matematičkog fakulteta u Novom Sadu, mentor
dr Sanja Rapajić, docent Prirodno-matematičkog fakulteta u Novom Sadu
dr Danijela Rajter-Ćirić, vanredni profesor Prirodno-matematičkog fakulteta u Novom Sadu, predsednik komisije

UNIVERSITY OF NOVI SAD

FACULTY OF SCIENCES

DEPARTMENT OF MATHEMATICS AND INFORMATICS

KEY WORDS DOCUMENTATION

Accession number:
ANO

Identification number:
INO

Document type: Monograph type
DT

Type of record: Printed text
TR

Contents Code: Master thesis
CC

Author: Bojana Jovanić
AU

Mentor: Dr. Dora Seleši
MN

Title: Marginal Distributions: The Law of Large Numbers and The Law of Small Numbers
TI

Language of text: Serbian
LT

Language of abstract: English
LA

Country of publication: Serbia
CP

Locality of publication: Vojvodina
LP

Publication year: 2011
PY

Publisher: Author's reprint
PU

Publ. place: University of Novi Sad, Faculty of Sciences, Department of Mathematics and Informatics, Trg Dositeja Obradovića 4
PP

Physical description: (4, 87, 0, 5, 2, 0, 0)
PD

Scientific field: Mathematics
SF

Scientific discipline: Actuarial Mathematics
SD

Key words: Law of large numbers, law of small numbers, marginal distribution, central limit theorem, the longest success run, order statistics, fluctuations of maxima,fluctuations of sums

UC:

Holding data: In library of Department of Mathematics and Informatics
HD

Note:
N

Abstract:
AB
The law of large numbers is a probability theory law that describes the result of conducting the same experiment a large number of times. According to this law, the average of the results obtained from a large number of trials should be close to the expected value and will tend to get closer as more trials are performed.

Accepted by the Scientific Board on: September 16, 2010.
ASB

Defended:
De

Thesis defend board:
DB
Member: Dr. Dora Seleši, Assistant Professor, Faculty of Sciences, University of Novi Sad, mentor
Member: Dr. Sanja Rapajić, Assistant Professor, Faculty of Sciences, University of Novi Sad
Member: Dr. Danijela Rajter-Ćirić, Associate Professor, Faculty of Sciences, University of Novi Sad, Chairman of the Board