Univerzitet u Beogradu

Matematički fakultet

MASTER RAD

Paretova raspodela i njena primena u aktuarskoj i finansijskoj matematici

Mentor:
Prof. Dr. Slobodanka Janković

Student:
Svetlana Đenić 1103/2013

BEOGRAD, 2016

Sadržaj

1 Uvod

Cilj ovog rada je da se približimo Paretovoj raspodeli koja je jedna od raspodela veličine (*size distributions*) i ima veliku primenu u odredjivanju veličine prihoda, veličine firmi, osiguranju.
U prvom poglavlju su navedene definicije različitih tipova Paretove raspodele i kratka istorija razvoja Paretove raspodele.
U drugom poglavlju smo se bavili različitim izvodjenjima Paretove raspodele, iz raspodele talenta, Markovljevog procesa i hijerarhijskog zaradjivanja.
U trećem poglavlju su navedene osobine Paretove raspodele, njeni momenti, karakterizacija, Lorencova kriva i koeficijenti nejednakosti mere koji su izvedeni iz Lorencove krive.
U četvrtom poglavlju su izvedene ocene parametara Paretove raspodele različitim metodama, kao i uporedjivanje dobijenih ocena.
U petom poglavlju su navedeni rezultati koje su istraživači dobili modelirajući podatke Paretovom raspodelom.
U šestom poglavlju su navedene raspodele koje su nastale modifikacijom Paretove raspodele.

1.1 Definicija Paretove raspodele

Paretova raspodela je tipičan predstavnik raspodela veličine, koje se često koriste u aktuarskoj i finansijskoj matematici, za odredjivanje raspodele prihoda, veličine firmi, veličine bogatstva. Otkriće ove raspodele se pripisuje Paretu (*Vilfredo Pareto*, 1895.).

Klasična Paretova raspodela ima funkciju raspodele

$$F(x) = 1 - \left(\frac{x}{x_0}\right)^{-\alpha}, \qquad x \geq x_0 > 0, \tag{1.1}$$

gde je $\alpha > 0$ parametar oblikovanja, koji meri težinu desnog repa, a x_0 je parametar skaliranja. Gustina raspodele je

$$f(x) = \frac{\alpha {x_0}^{\alpha}}{x^{\alpha+1}}, \qquad x \geq x_0 > 0, \tag{1.2}$$

funkcija kvantila je

$$F^{-1}(u) = x_0(1-u)^{-1/\alpha}, \qquad 0 < u < 1. \tag{1.3}$$

Uobičajeno obeležavanje da slučajna veličina ima Paretovu raspodelu je $X \sim Par(x_0, \alpha)$.

U svojim najznačajnijim doprinosima, krajem 19. veka, Pareto je predložio tri tipa ove raspodele. Prvi tip raspodele je definisan u tački (1.1). Drugi tip Paretove raspodele ima funkciju raspodele

$$F(x) = 1 - \left(1 + \frac{x-\mu}{x_0}\right)^{-\alpha}, \qquad x \geq \mu. \tag{1.4}$$

Ovaj tip raspodele se zove troparametarska Paretova raspodela, a postoji i specijalan slučaj kada je $\mu = 0$,

$$F(x) = 1 - \left(1 + \frac{x}{x_0}\right)^{-\alpha}, \quad x \geq 0, \tag{1.5}$$

gde su $x_0, \alpha > 0$, parametri, koja se obično naziva Paretova raspodela II tipa. Postoji veza izmedju Paretove raspodele I tipa i II tipa

$$X \sim Par(II)(x_0, \alpha) \Leftrightarrow X + x_0 \sim Par(x_0, \alpha). \tag{1.6}$$

Treći tip raspodele, Paretova raspodela III tipa, ima funkciju raspodele

$$F(x) = 1 - \frac{Ce^{-\beta x}}{(x-\mu)^{\alpha}}, \;\; x \geq \mu, \tag{1.7}$$

gde je $\mu \in R,\ \beta,\ \alpha > 0$, a C je funkcija tri parametra. Od ova tri tipa raspodele u praksi se najmanje koristi Paretova raspodela III tipa.

1.2 Istorija Paretove raspodele

Paretova raspodela je prikazana u Arnoldovoj (*Arnold*, 1983.) monografiji, ali je istorija Paretove raspodele i dalje predmet modernog istraživanja.

Pareto (1895. i 1896.) je posmatrajući poreske prihode, primetio opadajuću linearnu vezu izmedju logaritma prihoda i logaritma N_x, gde je N_x broj primalaca prihoda sa prihodom većim od x, gde je $x \geq x_0$. Definisao je

$$\log N_x = A - \alpha \log x, \tag{1.8}$$

odnosno,

$$N_x = e^A x^{-\alpha}, \tag{1.9}$$

gde su a i $\alpha > 0$. Deljenjem leve i desne strane sa brojem primalaca $N = N_{x_0}$ dobija se

$$\frac{N_x}{N} = \frac{e^A x^{-\alpha}}{N} = 1 - F(x) = \left(\frac{x}{x_0}\right)^{-\alpha}, \quad x \geq x_0 > 0. \tag{1.10}$$

Ubrzo se probudila pažnja javnosti i ekonomisti su počeli da kritikuju ideju o univerzalnoj formi sa samo jednim parametrom pomeranja, čime se dozvoljava neravnopravan poredak u društvu. Vrlo brzo je postalo jasno da je Paretova raspodela dobra jedino za aproksimaciju velikih prihoda preko nekog praga, pored toga došli su do zaključka i da α nije uvek blizu 1.5, kako je Pareto na početku mislio. Tek 1941. godine Dejvis (*H.T.Davis*) smatra da vrednost

$\alpha = 1.5$ treba da bude granica izmedju ravnopravnog društva ($\alpha > 1.5$) i neravnopravnog društva ($\alpha < 1.5$).

Paretovo otkriće na početku nije naišlo na odobravanje ni od strane engleske, ni od američke škole. Izuzetak su bili Stamp (*Stamp*, 1914.), Bovli (*Bowley*, 1926.) i Širas (*Shirras*, 1935.).

U aktuarskoj literaturi značajan doprinos je dao Norvežanin aktuar Birger Majdel (*Birger Miedell*, 1912.) primenjujući je na odredjivanju maksimalnog rizika u osiguranju života. Njegova hipoteza je da je osigurana suma proporcionalna primanjima nosioca osiguranja. U kasnijim istraživanjima u okviru aktuarstva Paretova raspodela se više koristi u vezi sa neživotnim osiguranjima, pre svega osiguranju automobila i osiguranju od požara.

2 Izvodjenja Paretove raspodele

2.1 Izvodjenje Paretove raspodele prihoda iz raspodele talenta

Više istraživača je razmatralo vezu izmedju raspodele prihoda i raspodele talenata ili raspodele sklonosti. Na primer Amon (*Ammon*, 1895.), Pareto (1897.), Rodes (*Rhodes*, 1944.). Rodes je pretpostavio da je "talenat" neprekidna promenljiva Z, sa gustinom $h(z)$ i prosečnim dobitkom $m(z) = E(X|Z = z)$, gde X označava promenljivu prihoda. Njegova osnovna pretpostavka je da je uslovni koeficijent varijacije $\lambda = CV(X|Z = z)$ konstantan za sve grupe talenata.

Uslovna funkcija preživljavanja prihoda, za one čiji je talenat z, u standardizovanoj formi je

$$\bar{F}\left[\frac{x - m(z)}{\lambda m(z)}\right], \tag{2.1}$$

gde $\lambda m(z)$ predstavlja uslovnu standardnu devijaciju prihoda, a bezuslovna funkcija raspodele je

$$\bar{F}(x) = \int_0^\infty h(z)\bar{F}\left[\frac{x - m(z)}{\lambda m(z)}\right]dz. \tag{2.2}$$

Kada stavimo da je smena $x - m(z) = \lambda m(z)v$ dobija se

$$m(z) = \frac{x}{1 + \lambda v}. \tag{2.3}$$

Ovim uslovom se dobija mogućnost da se z računa kao funkcija od x i od v. Ali, pored toga potrebno je odrediti i funkciju raspodele za X. Da bi se izračunala funkcija raspodele koristi se

$$\frac{\partial m(z)}{\partial v}\frac{dz}{dv} = -\frac{\lambda x}{(1 + \lambda v)^2}. \tag{2.4}$$

Uvodjenjem oznaka $M(v, x) = \partial m(z)/\partial v$ i $P(v, x) = h(z)$ dobija se

$$\int_{v_1}^{v_2} \frac{P(v, x)\bar{F}(v)\lambda x}{M(v, x)(1 + \lambda v)^2}dv. \tag{2.5}$$

Nove granice integrala se dobijaju iz jednačine (2.3). Ukoliko se za vrednost z uzme najmanja moguća vrednost odnosno 0, da talenat ne postoji, jednačina dobija oblik $m_0 = m(0) = x/(1 + \lambda v_2)$, iz ovoga sledi da je $v_2 = (x/m_0 - 1)/\lambda$. Slično, kada uzmemo najveću moguću vrednost za z, $z \to \infty$, tada i $m(z) \to \infty$, odakle je $v_1 = -1/\lambda$. Odavde (2.5) može biti napisano kao

$$\int_{-1/\lambda}^{(x/m_0 - 1)/\lambda} \frac{P(v, x)\bar{F}(v)\lambda x}{M(v, x)(1 + \lambda v)^2}dv. \tag{2.6}$$

Primetimo da je gornja granica funkcija od x. Ovaj način predstavljanja nije jednostavan bez dodatnih pretpostavki. Ukoliko $P(v,x)/M(v,x)$ možemo da predstavimo u formi $\phi(x)\psi(v)$ integral prelazi u

$$\lambda x\phi(x)\int_{-1/\lambda}^{(x/m_0-1)/\lambda}\frac{\psi(v)\bar{F}(v)}{(1+\lambda v)^2}dv=\lambda x\phi(x)\chi(x). \tag{2.7}$$

Rodes je razmatrao specijalan slučaj gde je $\chi(x)$ aproksimativno konstantno za veliko x, što implicira da je $\bar{F}(x)$ proporcionalno funkciji $cx\phi(x)$ za neko $c>0$. Ostaje da odredimo $\phi(x)$. Da bismo to uradili uvodimo smenu $w(z)=M(v,x)/P(v,x)$, odatle sledi da je $w\phi\psi=1$. Logaritmovanjem i diferenciranjem po x dobija se

$$\frac{w'(z)}{w(z)}\frac{dz}{dx}=-\frac{\phi'(x)}{\phi(x)}. \tag{2.8}$$

Sa druge strane $M(v,x)(dz/dx)(1+\lambda v)=1$, pa je

$$\frac{w'(z)}{w(z)}\frac{1}{M(v,x)(1+\lambda v)}=-\frac{\phi'(x)}{\phi(x)}, \tag{2.9}$$

odnosno, korišćenjem $M(v,x)=h(z)w(z)$ dobijamo

$$\frac{w'(z)}{w^2(z)h(z)}\frac{1}{1+\lambda v}=-\frac{\phi'(x)}{\phi(x)}. \tag{2.10}$$

Kako $-\frac{\phi'(x)}{\phi(x)}$ ne zavisi od v, daljim diferenciranjem po v se dobija

$$\left[\frac{w'(z)}{w^2(z)h(z)}\right]'\frac{dz}{dv}\frac{1}{1+\lambda v}-\frac{w'(z)}{w^2(z)h(z)}\frac{1}{(1+\lambda v)^2}=0, \tag{2.11}$$

Koristeći jednakosti (2.3) i (2.4) dobija se

$$\left[\frac{w'(z)}{w^2(z)h(z)}\right]'\Big/\left[\frac{w'(z)}{w^2(z)h(z)}\right]=-\frac{M(v,x)}{m(z)}. \tag{2.12}$$

Sve funkcije u ovoj jednačini su funkcije od z. Integracijom se dobija

$$\frac{w'(z)}{w^2(z)h(z)}=\frac{a}{m(z)}, \tag{2.13}$$

za neku konstantu a. Ukoliko opet koristimo $M(v,x)=w(z)h(z)$, dobijamo

$$\frac{w'(z)}{w(z)}=\frac{aM(v,x)}{m(z)}. \tag{2.14}$$

Odatle proizilazi da je $w(z)=cm(z)^a$ za neku konstantu c. Takodje,

$$\frac{w'(z)}{w^2(z)h(z)}\frac{1}{1+\lambda v}=-\frac{\phi'(x)}{\phi(x)}, \tag{2.15}$$

odatle sledi

$$\frac{\phi'(x)}{\phi(x)} = -\frac{a}{m(z)(1+\lambda v)} = -\frac{a}{x}. \tag{2.16}$$

Odatle je $\phi(x) = kx^{-a}$, za neku konstantu k. Iz definicije $w\phi\psi = 1$, dobijamo $\psi(v)cm(z)^a kx^{-a} = 1$, dakle

$$\psi(v) = \frac{1}{ck}\left[\frac{x}{m(z)}\right]^{\alpha} = \frac{1}{ck}(1+\lambda v)^a, \tag{2.17}$$

što je funkcija samo od v. Ovim se jednačina (2.7) može napisati u sledećem obliku

$$\chi(x) = \frac{1}{ck}\int_{-1/\lambda}^{(x/m_0-1)/\lambda} \bar{F}(v)(1+\lambda v)^{-a-2}dv, \tag{2.18}$$

i ukoliko je ova jednakost približno konstantna za veliko x, konačno dobijamo

$$\bar{F}(x) \approx x\phi(x) = kx^{-a-1}. \tag{2.19}$$

Ova raspodela prihoda je Paretova raspodela.

Ideja da se objasni veličina prihoda preko sposobnosti, koristi se i u literaturi vezanoj za raspodelu veličine firmi. Lukas (*Lukas*, 1978.) je predstavio model u kojem je tvrdio da posmatrana raspodela veličine može rešiti problem kako postaviti produktivne faktore izmedju menadžera različitih sposobnosti kako bi poboljšali proizvodnju. Ukoliko "menadžerski talenat" ima Paretovu raspodelu, onda je i raspodela veličine takodje ovog oblika.

2.2 Markovljev proces koji vodi do Paretove raspodele

Čempernov (*Champernowne*, 1953.) je pokazao da će pod odredjenim pretpostavkama, stacionarna raspodela prihoda Markovljevog procesa, koji je na odgovarajući način definisan, biti približna Paretovoj raspodeli nezavisno od početne raspodele.

Čempernov posmatra prihode kao diskretan Markovljev lanac: prihod za tekući period, odnosno stanje Markovljevog lanca, zavisi samo od prihoda za poslednji period i slučajnog uticaja. On pretpostavlja da postoji neki minimum prihoda x_0 i da intervali prihoda definišu geometrijski niz (granice klase j su veće od granice $j-1$ za odredjeni faktor c). Ovim je osoba u klasi j ukoliko je njen prihod izmedju x_0c^{j-1} i x_0c^j. Verovatnoća promene p_{ij} je definisana kao verovatnoća da u trenutku $t+1$ bude u klasi j, ukoliko je u trenutku t bila u klasi i.

Čempernov je pretpostavio da verovatnoća prelaska iz jedne klase u drugu zavisi samo od širine skoka, a ne i od pozicije sa koje počinje (oblik homogenosti). Drugim rečima verovatnoća prelaska p_{ij} je funkcija od $j - i = k$, koja je nezavisna od i. Ukoliko je $x_j(t)$ broj primalaca prihoda iz klase prihoda j za period t proces se razvija u skladu sa

$$x_j(t+1) = \sum_{k=-\infty}^{j} x_{j-k}(t)p_k. \tag{2.20}$$

Čempernov dalje pretpostavlja (njegova "osnovna pretpostavka") da su prelasci mogući jedino izmedju $-n$ i 1. Ukoliko proces traje duže vreme, raspodela prihoda dostiže ravnotežu u kojoj delovanje matrice prelaska ostavlja raspodelu nepromenjenom. Stanje ravnoteže je opisano na sledeći način

$$x_j = \sum_{k=-n}^{1} x_{j-k}p_k, \qquad j > 0. \tag{2.21}$$

Rešenje ove jednačine se dobija smenom $x_j = z^j$ i rešavanjem karakteristične jednačine

$$g(z) = \sum_{k=-n}^{1} z^{l-k}p_k - z = 0. \tag{2.22}$$

Ova jednačina ima dva pozitivna realna rešenja, jedno od njih je jednako 1. Da bi drugo rešenje bilo izmedju 0 i 1, Čempernov je uveo pretpostavku stabilnosti

$$g'(1) = -\sum_{k=-n}^{1} kp_k > 0. \tag{2.23}$$

Pošto je $g(0) = p_1 > 0$ i važi uslov stabilnosti $g'(1) > 0$, drugo rešenje mora da zadovolji uslov

$$0 < b < 1, \tag{2.24}$$

odakle je stacionarna raspodela

$$x_j = b^j. \tag{2.25}$$

Stoga je ukupan broj prihoda $1/(1-b)$ i za bilo koji drugi broj prihoda N stacionarna raspodela postaje

$$x_j = N(1-b)b^j. \tag{2.26}$$

Broj prihoda koji su veći ili jednaki $\widetilde{x}_j = x_0c^j$ je

$$N_{\widetilde{x}_j} = Nb^j \tag{2.27}$$

ili

$$\log N_{\widetilde{x}_j} = \log N + j\log b. \tag{2.28}$$

Uvodjenjem oznake $\alpha = -\log b / \log c$ i $\gamma = \log N + \alpha \log x_0$, dolazi se do

$$\log N_{\widetilde{x}_j} = \gamma - \alpha \log \widetilde{x}_j. \tag{2.29}$$

To znači da je logaritam broja prihoda koji su veći od $\widetilde{x}_j$ linearna funkcija od $\log \widetilde{x}_j$, što predstavlja Paretovu raspodelu.

Suštinska odlika ovog modela je uslov stabilnosti (2.23), koji znači da je očekivanje mogućeg prelaza uvek smanjenje prihoda, bez obzira na iznos prihoda sa kog se polazi. Stajndl (*Steindl*, 1965.) tvrdi da je ekonomsko opravdanje ove pretpostavke implicitno objasnio Čempernov u sledećem modelu. U tom modelu on tvrdi da je broj prihoda konstantan računajući i one koji neće preživeti, jer posle svakog primaoca prihoda koji "ne preživi" postoji naslednik njegovog prihoda. Ali to znači da će pri promeni starijeg mladjim, često doći do pada u prihodu, posebno ukoliko se radi o većim prihodima.

Čempernov je razmatrao više generalizacija njegovog osnovnog modela. Ukoliko su prelasci mogući u produženom intervalu, izmedju $-n$ i $m, m > 1$, jedina asimptotska raspodela može biti Paretova raspodela. Ovo je takodje moguće, ako je ljudima dozvoljeno da upadnu u grupe (prema godinama ili prema zanimanju) i ako su dozvoljena pomeranja iz jedne u drugu grupu. Medjutim Markovljev proces ne bi imao stacionarnu raspodelu ukoliko matrica verovatnoća prelaska nije konstantna. Teško je i zamisliti društvo čiji je institucionalni okvir toliko statičan. Takodje, osnovna pretpostavka je da je verovatnoća rasta ili pada nezavisna od veličine prihoda.

2.3 Lajdalov-ov model hijerarhijskog zaradjivanja

Lajdal (*Lydall*, 1959.) je pretpostavio da su ljudi koji rade u organizacijama ili firmama hijerarhijski rasporedjeni i da su njihove plate pokazatelj organizacije firme.

Pretpostavimo da su nivoi $l_i, i = 1, 2, ..., k$, rastući brojevi i neka je x_i plata na nivou l_i i y_i je broj zaposlenih na i-tom nivou. Postoje dve pretpostavke:

$$\frac{y_i}{y_{i+1}} = n, \quad \text{gde je } n > 1 \text{ konstantno za svako } i, \tag{2.30}$$

i

$$\frac{x_i}{x_{i+1}} = p, \quad \text{gde je } p < 1 \text{ konstantno za svako } i. \tag{2.31}$$

Ove pretpostavke pokazuju činjenicu da menadžeri na svakom nivou nadgledaju konstantan broj zaposlenih na nivou ispod njih (2.30) i da je plata odredjenog menadžera konstantna proporcija sumarnih plata zaposlenih koje nadgleda (2.31) . Prirodno je pretpostaviti da je $x_{i+1}/x_i = np > 1$. Na najvišem nivou će biti jedna osoba, na sledećem nivou će biti n osoba, na sledećem n^2 i tako redom. Stoga

$$y_i = n^{k-i}. \tag{2.32}$$

Ukupan broj N_i osoba na nivou l_i ili iznad je

$$N_i = 1 + n + n^2 + ... + n^{k-i} = \frac{n^{k-i+1} - 1}{n - 1}. \tag{2.33}$$

I proporcija Q_i svih zaposlenih u firmi koji rade na nivou l_i ili iznad je

$$Q_i = \frac{N_i}{N_1} = \frac{n^{k-i+1} - 1}{n^k - 1} \approx n^{1-i}. \tag{2.34}$$

Iz (2.31) vidimo da

$$x_i = (np)^{i-1} x_1. \tag{2.35}$$

Stoga je,

$$\log Q_i = \frac{\log n}{\log np} \log x_1 - \frac{\log n}{\log np} \log x_i = \log c - \alpha \log x_i, \tag{2.36}$$

odnosno,

$$Q_i = c x_i^{-\alpha}, \tag{2.37}$$

što je Paretova raspodela.

2.4 Ostali pristupi

Mandelbrot (1964.) je izveo Paretovu raspodelu iznosa štete od požara uz pretpostavku da je verovatnoća da vatra pojača svoj intenzitet konstantna u bilo kom trenutku. Pretpostavio je da je intenzitet predstavljen kao slučajna vrednost N. Kada je vrednost $N = 0$, tada nema vatre, vatra se javlja kada vrednost N postane 1. Požar se završava ukoliko N postane jednako 0 ili kada je sve što se može zapaliti već uništeno. Dalje pretpostavlja da u bilo kom trenutku, postoji verovatnoća $p = 1/2$ da se vatra susretne sa nekim novim materijalom i da intenzitet poraste do 1, a takodje postoji verovatnoća $p = 1/2$ da se zbog nedostatka materijala koji se može zapaliti ili zbog dejstva vatrogasaca intenzitet umanji za 1. U odsustvu maksimalnog stepena oštećenja i donje granice evidentirane štete, trajanje požara će biti paran broj za koji važi

$$P(D = x) = 2 \binom{1/2}{x/2} (-1)^{x/2-1}, \tag{2.38}$$

što se ponaša kao $x^{-3/2}$ za dovoljno veliko x. Kada se pretpostavi da je šteta mala, ispod praga x_0, čak možda nije ni valjano zabeležena, obim štete je dat na sledeći način

$$P(D > x) = (x/x_0)^{-1/2}, \;\; x_0 \leq x, \tag{2.39}$$

što je ustvari Paretova raspodela sa parametrom $\alpha = 1/2$. Za vrednosti α oko 0.5 je otkriveno da mogu opisati raspodelu štete u posleratnoj Švedskoj.

Spilberg (*Shpilberg*, 1977.) je predstavio argumente koji vode do toga da Paretova raspodela predstavlja raspodelu gubitka nastalog usled požara. Pretpostavimo da je "stopa mortaliteta" $\lambda(t)$ vatre konstantna, jednaka α, pa trajanje požara T ima eksponencijalnu raspodelu. Pod pretpostavkom da je šteta X eksponencijalno povezana sa trajanjem požara,

$$X = x_0 \exp(kT), \tag{2.40}$$

za neko $x_0, \; k > 0$, funkcija raspodele od X je data na sledeći način

$$F(x) = 1 - \left(\frac{x_0}{x}\right)^{\alpha/k}, \;\; x \geq x_0, \tag{2.41}$$

i ima Paretovu raspodelu.

3 Osobine Paretove raspodele

3.1 Momenti Paretove rapodele

Paretova gustina raspodele ima polinomijalni desni rep, koji je pravilno promenljiv u beskonačnosti sa indeksom -$\alpha-1$. Zbog ovoga desni rep je teži za manje α, što implicira da postoje samo momenti nižeg reda. Momenat k-tog reda Paretove raspodele jedino postoji ako je $k < \alpha$, i u tom slučaju je jednak

$$E(X^k) = \frac{\alpha x_0^k}{\alpha - k}. \tag{3.1}$$

Specijalno, očekivanje je

$$E(X) = \frac{\alpha x_0}{\alpha - 1}, \tag{3.2}$$

disperzija je

$$D(X) = \frac{\alpha x_0^2}{\alpha(\alpha-1)^2(\alpha-2)}, \tag{3.3}$$

koeficijent varijacije je

$$CV = \frac{1}{\sqrt{\alpha(\alpha-2)}}, \tag{3.4}$$

faktori oblika su

$$\sqrt{\beta_1} = 2\frac{\alpha+1}{\alpha-3}\sqrt{1-\frac{2}{\alpha}}, \quad \alpha > 3, \tag{3.5}$$

$$\beta_2 = \frac{3(\alpha-2)(3\alpha^2+\alpha+2)}{\alpha(\alpha-3)(\alpha-4)}, \quad \alpha > 4. \tag{3.6}$$

Sledi da kada $\alpha \to \infty$ onda $\sqrt{\beta_1} \to 2$ i $\beta_2 \to 9$.

Za članove klase raspodela sa ekstremno teškim repom (sa $\alpha < 1$), moraju se koristiti i druge mere lokacije osim očekivanja. Može se koristiti geometrijsko očekivanje $x_g = \exp\{E(\log X)\}$ što je u ovom slučaju

$$x_g = x_0 \exp\left(\frac{1}{\alpha}\right) \tag{3.7}$$

i harmonijska sredina $x_h = \{E(X^{-1})\}^{-1}$, što je jednako

$$x_h = x_0(1+\alpha^{-1}). \tag{3.8}$$

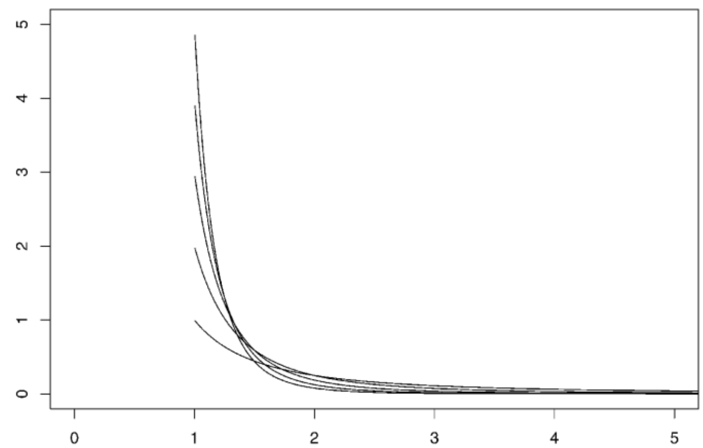
Poredjenjem sa funkcijom kvantila

$$F^{-1}(u) = x_0(1-u)^{-1/\alpha}, \qquad 0 < u < 1, \tag{3.9}$$

dolazi se do zaključka da je geometrijska sredina jednaka $(1-e^{-1})$-om kvantilu. Takodje jednakosti (3.2), (3.7), (3.8) ilustruju poznate nejednakosti $E(X) \geq x_g \geq x_h$.

Lako se može videti da se za vrednosti $\alpha_1 \leq \alpha_2$ i $x_{01} \geq x_{02}$ funkcije dvoparametarske Paretove raspodele $Par(x_{0i}, \alpha_i)$ ne seku, tako da je pod ovim uslovima

Paretova raspodela stohastički uredjena, sa $X_1 \geq_{FSD} X_2$[1]. Ovo implicira i neke druge posledice, da su i momenti takodje uredjeni, pod pretpostavkom da postoje, posebno $E(X_1) \geq E(X_2)$. Gustina raspodele opada, stoga je moda raspodele u x_0. Na slici (3.1) su prikazani neki primeri gustine Paretove raspodele. Iz (3.9), medijana je $F^{-1}(0.5) = 2^{1/\alpha}x_0$.



Slika 3.1: Gustine Paretove raspodele $x_0 = 1$ i $\alpha = 1(1)5$, odozdo na levo

Paretova raspodela je zatvorena u odnosu na minimizaciju u smislu da

$$X \sim Par(x_0, \alpha) \Rightarrow X_{1:n} \sim Par(x_0, n\alpha). \tag{3.10}$$

Medjutim, druge statistike poretka nemaju Paretovu raspodelu. Na primer, raspodela maksimuma uzorka je data na sledeći način

$$F_{n:n}(x) = \left[1 - \left(\frac{x}{x_0}\right)^{-\alpha}\right]^n, \quad 0 < x_0 \leq x. \tag{3.11}$$

Nedavno, Stopa (*Stoppa*, 1990.a, 1990.b) je predložio ovu raspodelu, za $n \in R$, kao model za raspodelu veličine ličnog prihoda.

Definicija 3.1.1. Raspodela momenta se definiše na sledeći način

$$F_k(x) = \frac{1}{E(X^k)} \int_0^x t^k f(t)dt, \; x \geq 0, \; k = 0, 1, 2, ... \tag{3.12}$$

[1]FSD - first-order stochastic dominance - Stohastička dominacija I reda. A dominira B znači da funkcija $F_A(x) \leq F_B(x)$ za svako x i da postoji neko x gde važi stroga nejednakost. Ukoliko se predje na verovatnoću onda A dominira B, ako za svako x važi $P(A \geq x) \geq P(B \geq x)$, a postoji neko x za koje važi stroga nejednakost.

Paretova raspodela je zatvorena u odnosu na formiranje raspodele momenta i u odnosu na stepenovanje, ako je

$$F \sim Par(x_0, \alpha) \Rightarrow F_k \sim Par(x_0, \alpha - k), \tag{3.13}$$

gde je $\alpha < k$ i ako je $a > 0$, onda je

$$X \sim Par(x_0, \alpha) \Rightarrow X^a \sim Par\left(x_0^a, \frac{\alpha}{a}\right). \tag{3.14}$$

Takodje, ako X ima standardnu Paretovu raspodelu i ako je $W = X^{-1}$, onda W ima gustinu raspodele

$$f(w) = \alpha x_0^{\alpha} w^{\alpha-1}, \;\; 0 < w < x_0^{-1}. \tag{3.15}$$

Opasnost od rizika ili hazard rate je data funkcijom

$$r(x) = \frac{f(x)}{1 - F(x)} = \frac{\alpha}{x}, \;\; x > x_0, \tag{3.16}$$

koja je monotono opadajuća. Hazard rate daje intenzivnost po kojoj rizik od velikih iznosa opada kada x raste. Njeno sporo opadanje je odraz teškog repa Paretove raspodele. Funkcija srednjeg prekoračenja potražnje[2], koja se definiše kao

$$e(x) = E(X - x | X > x) = \frac{\int_x^{\infty} (t - x) dF(t)}{\int_x^{\infty} dF(t)}, \; x \geq 0, \tag{3.17}$$

je

$$e(x) = \frac{x}{\alpha - 1}, \;\; x \geq x_0. \tag{3.18}$$

Iz (3.16) i (3.18) proizilazi da je proizvod opasnosti od rizika i stepena prekoračenja potražnje konstantan

$$r(x)e(x) = \frac{\alpha}{\alpha - 1}. \tag{3.19}$$

Teorija raspodele povezana sa uzorcima iz Paretove raspodele je nešto komplikovanija. Feler (*Feller*, 1971.) je našao da konvolucija[3] Paretovih raspodela takodje ima Paretov rep, medjutim, izrazi za raspodele su prilično komplikovani. Ipak, jedna značajna osobina Paretove raspodele je povezana sa centralnom graničnom teoremom: na odgovarajući način normalizovane parcijalne sume nezavisnih slučajnih veličina sa $Par(x_0, \alpha)$ raspodelom su asimptotski normalno rasporedjene, jedino ako je $\alpha \geq 2$, za $\alpha < 2$ granična raspodela je stabilna raspodela koja nije normalna Zolotarev (*Zolotarev*, 1986.).

[2] average excess claim
[3] raspodela zbira

Osnovna osobina standardne $Par(x_0, \alpha)$ raspodele je njena bliska veza sa eksponencijalnom raspodelom

$$X \stackrel{d}{=} x_0 \exp\left(\frac{Y}{\alpha}\right), \tag{3.20}$$

gde je Y slučajna veličina sa standardnom eksponencijalnom raspodelom, gustine $f_Y(y) = \exp(-y), \quad y > 0$. Klasična Paretova raspodela se može posmatrati kao log-eksponencijalna raspodela. Ovo, na primer, daje vezu izmedju Paretove i gama raspodele. Neka su $X_i, \quad i = 1, ..., n$, nezavisne slučajne veličine sa Paretovom raspodelom, $X_i \sim Par(x_0, \alpha)$ onda je

$$\alpha \log \prod_{i=1}^{n} \left(\frac{X_i}{x_0}\right) = \alpha \sum_{i=1}^{n} \log \left(\frac{X_i}{x_0}\right) = \Gamma(n, 1). \tag{3.21}$$

Generalno, blizak odnos izmedju Paretove i eksponencijalne raspodele podrazumeva da se mnoge osobine mogu dobiti iz osobina druge, na primer karakterizacija.

3.2 Karakterizacija

Veliki broj karakterizacija Paretove raspodele je zasnovan na ponašanju funkcija statistika poretka, što je posledica ogromne literature o karakterizaciji eksponencijalne raspodele pomoću ovih funkcija. Ovde će biti prikazani samo najvažniji primeri.

Karakterizacija I

$$X_{i:n}, \frac{X_{i+1:n}}{X_{i:n}} \text{ su nezavisne} \Rightarrow X \sim \text{Pareto}. \tag{3.22}$$

Ovo je verzija klasične karakterizacije eksponencijalne raspodele, sa uslovima rastojanja, Fiš (*Fisz*, 1958.).

Više karakterizacija Paretove raspodele je zasnovano na linearnosti funkcije očekivanja viška (očekivanje preostalog života), mada ponekad u skrivenoj formi.

Sledeću karakterizaciju je dao Batičaria (*Bhattacharya*, 1963.):

$$P(X > y | X > z) = P\left(\frac{z}{x_0} X > y\right), \quad \text{za sve } y > z \geq x_0, \tag{3.23}$$

što predstavlja karakterizaciju $Par(x_0, \alpha)$ raspodele. Ako se pretpostavi da momenti postoje, ova karakterizacija se svodi na linearnost funkcije srednjeg prekoračenja, kao funkcija od z (Kotc (*Kotz*) i Šenbag (*Shanbhag*), 1980.).

Krisnadži (*Krishnaji*, 1970.) je dokazao opštu karakterizaciju pretpostavljajući da je prijavljeni prihod Y povezan sa tačnim prihodom X preko multiplikativne greške

$$Y \stackrel{d}{=} RX, \tag{3.24}$$

gde su R, X nezavisni. Da bi ovaj model imao smisla, neophodno je da R uzme vrednosti u intervalu izmedju 0 i 1. U aktuarstvu bi analogna stvar bilo preterivanje o potraživanju osiguranja. Ukoliko postavimo da je $Y \stackrel{d}{=} X/R$, karakterizacija koja se dobije može se primeniti uz minimalne izmene. Pretpostavljajući da R ima stepenu raspodelu, sa gustinom

$$f(r) = pr^{p-1}, \ \ 0 \leq r \leq 1, \tag{3.25}$$

gde je $p > 0$, Krisnadži je izveo sledeću karakterizaciju: ukoliko je $P(X > x_0) = 1$, za neko $x_0 > 0$ i $P(RX > x_0) > 0$, gde su R i X nezavisni, onda je

$$P(RX > y|RX > x_0) = P(X > y), \ \ y > x_0, \tag{3.26}$$

ako i samo ako $X \sim Par(x_0, \alpha)$.
Rezultat će ostati isti ukoliko uslov (3.26) zamenimo uslovom

$$E(Y - x\big|Y > x) = E(X - x\big|X > x), \ \ x > x_0 > 0, \tag{3.27}$$

sa $E(X^+) < \infty$ što sledi iz rada Kotca i Šenbaga (1980.).

Karakterizacija II

Revankar (*Revankar*), Hartli (*Hartley*) i Pagan (*Pagano*) 1974. godine, su dali drugu karakterizaciju u terminima neprijavljenih prihoda. Medjutim, suprotno Krisnadžijevom pristupu korišćenja multiplikativnog prikazivanja greške, oni su pretpostavili aditivnu vezu izmedju stvarnog i prikazanog prihoda. Neka slučajne veličine X, Y i U respektivno označavaju trenutni, neprijavljeni prihod, prijavljeni prihod i grešku i neka za njih važi

$$Y = X - U, \ \ \text{gde je } 0 < U < \max\{0, X - m\}, \tag{3.28}$$

gde je m nivo oslobodjen od poreza. Pod pretpostavkom da je srednja vrednost neprijavljenog prihoda za $X = x > m$ proporcionalna $x - m$, što je

$$E(U\big|X = x) = b(x - m) = a + bx, \tag{3.29}$$

$0 < b < 1$ i $a = -bm$ i ako važi uslov $E(X) < \infty$ sledi da je

$$E(U\big|X > y) = \alpha + \beta y, \ \ \beta > b > 0, \tag{3.30}$$

ako i samo ako X ima Paretovu raspodelu sa konačnim prvim momentom. Specijalno, za $a = \alpha$ dobija se $Par(x_0, \alpha)$. Za $a < \alpha$ dobijamo Paretovu raspodelu II tipa.

3.3 Lorencova kriva i mere nejednakosti

Lorencova kriva ili kriva koncentracije, je kriva koja grafički prikazuje procenat u raspodeli dohotka. Odnosno, pokazuje koliki procenat ukupne populacije, ima koliki procenat ukupnog bogatstva ili prihoda. Lorencova kriva definiše se na sledeći način

$$L(u) = \frac{1}{E(X)} \int_0^u F^{-1}(t)dt, \quad u \in [0,1], \tag{3.31}$$

gde je

$$F^{-1}(t) = sup\{x|F(x) \leq t\}, \; t \in [0,1]. \tag{3.32}$$

Lorencova kriva Paretove raspodele postoji uvek kada je $\alpha > 1$ i jednaka je

$$L(u) = 1 - (1-u)^{1-1/\alpha}, \quad 0 < u < 1. \tag{3.33}$$

Lorencove krive Paretove raspodele se nikada ne seku i to za $X_i \sim Par(x_0, \alpha_i)$,

$$L_{X_1}(u) \leq L_{X_2}(u), \text{ za svako } u \in [0,1] \Leftrightarrow \alpha_1 \leq \alpha_2, \tag{3.34}$$

gde je $\alpha_i > 1, \; i = 1, 2$.

Tilov (*Theil*) koeficijent je statistika koja se prvenstveno koristi za merenje ekonomske nejednakosti i drugih ekonomskih pojava. Može se posmatrati kao mera viška zaposlenih, nedostatka različitosti, izolacije, nejednakosti, neslučajnosti. Predložio ga je Henri Til (*Henri Theil*). Tilov koeficijent za Paretovu raspodelu je

$$T_1 = \int_0^\infty \frac{x}{E(X)} \log\left\{\frac{x}{E(X)}\right\} dF(x) = \frac{1}{\alpha - 1} - \log\left(\frac{\alpha}{\alpha - 1}\right). \tag{3.35}$$

Svaki od ovih izraza opada sa porastom α, pokazujući da se parametar α, pre njegov inverz, može smatrati merom nejednakosti.

4 Ocenjivanje parametara Paretove raspodele

Ocenjivanjem parametara Paretove raspodele bavili su se Arnold (1983.) i Džonson (*Johnson*), Kotc i Balakrišan (*Balakrishnan*) (1994.). Ovde će biti uključeni klasičan regresioni tip ocenjivanja, metoda maksimalne verodostojnosti, i rezultati koji su vezani za UMVUE [1] ocene.

4.1 Regresiona ocena

Metod najmanjih kvadrata se zasniva na minimiziranju kvadrata odstupanja svih empirijskih tačaka od regresione linije. Ocena najmanjih kvadrata za parametar α je

$$\hat{\alpha}_{LS} = \frac{-n\sum_{i=1}^{n}\log X_i \cdot \log \bar{F}(X_i) + \sum_{i=1}^{n}\log X_i \cdot \sum_{i=1}^{n}\log \bar{F}(X_i)}{n\sum_{i=1}^{n}(\log X_i)^2 - (\sum_{i=1}^{n}\log X_i)^2}, \tag{4.1}$$

a ocena $\hat{x_0}$ je definisana na sledeći način

$$\overline{\log \bar{F}(X)} = \hat{\alpha}_{LS} \cdot \log \hat{x_0} - \hat{\alpha}_{LS} \cdot \overline{\log X} \tag{4.2}$$

(gde linija iznad označava prosek). Ova ocena je popularna u primeni, Kvand (*Quandt*, 1966b) je pokazao da je i dosledna (konzistentna).

Pareto (1896, 1897.a,b), Benini (1897.) i Džini (*Gini*,1909.a) su koristili model Košijeve (*Cauchy*) regresije, koji se danas retko koristi. Polastri (*Pollastri*, 1990.), je pokazao da je Košijev metod neznatno bolji, od regresije najmanjih kvadrata, u pogledu srednje kvadratne greške (MSE) [2]. Hosein (*Hossain*) i Zimer (*Zimmer*) (2000.) su preporučili ocenu najmanjih kvadrata, više nego metodu maksimalne verodostojnosti i njoj slične ocene za ocenu x_0 i α, za male vrednosti parametra ($\alpha < 4$).

4.2 Ocena metodom maksimalne verodostojnosti

Funkcija verodostojnosti za uzorak sa Paretovom raspodelom je

$$L = \prod_{j=1}^{n} \frac{\alpha x_0^{\alpha}}{x_j^{\alpha+1}}. \tag{4.3}$$

Ocena parametra α metodom maksimalne verodostojnosti je

$$\hat{\alpha} = n\left[\sum_{i=1}^{n}\log\left(\frac{X_j}{\hat{x_0}}\right)\right]^{-1}, \tag{4.4}$$

[1] eng. Uniformly minimum-variance unbiased estimator- UMVUE. UMVUE - ravnomerna nepristrasna ocena sa minimalnim odstupanjem je ocena koja je nepristrasna i ima najmanju disperziju od svih drugih ocena

[2] eng. Mean Square Error - MSE. MSE od ocene $\hat{\theta}$ se računa na sledeći način $MSE(\hat{\theta}) = D(\hat{\theta}) + bias(\hat{\theta})$ gde je $bias(\hat{\theta}) = E(\hat{\theta} - \theta)^2$

gde je ocena za granicu x_0

$$\hat{x_0} = X_{1:n}. \tag{4.5}$$

Treba primetiti pošto je $x_0 \leq x$ $\hat{x_0}$ je veće od x_0. Ukoliko se ova ocena koristi za odredjivanje $\hat{\alpha}$ u (4.4) vidi se da je i $\hat{\alpha} > \alpha$. Sledeća računanja to i pokazuju

$$E(\hat{\alpha}) = \frac{n\alpha}{n-2}, \quad n > 2, \tag{4.6}$$

$$D(\hat{\alpha}) = \frac{n^2\alpha^2}{(n-2)^2(n-3)}, \quad n > 3, \tag{4.7}$$

što daje

$$MSE(\hat{\alpha}) = \frac{\alpha^2(n^2+4n-12)}{(n-2)^2(n-3)}, \quad n > 3. \tag{4.8}$$

Odgovarajući rezultat za $\hat{x_0}$ je

$$E(\hat{x_0}) = \frac{nx_0\alpha}{n\alpha-1}, \quad n > \frac{1}{\alpha}, \tag{4.9}$$

$$D(\hat{x_0}) = \frac{nx_0\alpha^2}{(n\alpha-1)^2(n\alpha-2)}, \quad n > \frac{2}{\alpha}, \tag{4.10}$$

što dalje daje

$$MSE(\hat{x_0}) = \frac{2x_0^2}{(n\alpha-1)(n\alpha-2)}, \quad n > \frac{2}{\alpha}. \tag{4.11}$$

Obe ocene su konzistentne (Kvand, 1966.b). Ocena $1/\hat{\alpha}$ je asimptotski efikasna u eksponencijalnom slučaju, isto važi i za $\hat{\alpha}$. Saksena (*Saksena*) i Džonson (1984.), su pokazali da su ocene metodom maksimalne verodostojnosti kompletne.
Ukoliko je parametar α poznat, onda će $T = X_{1:n}$ biti kompletna, dovoljna statistika sa gustinom

$$f(t) = \frac{n\alpha}{t^{n\alpha+1}}x_0^{n\alpha}, \quad x_0 \leq t, \tag{4.12}$$

i $Par(n\alpha, x_0)$ raspodelom. Ocena metodom maksimalne verodostojnosti za parametar α, kada je x_0 poznato, je jednaka

$$\hat{\alpha} = \frac{n}{\sum_{i=1}^{n}\log(X_j/x_0)}. \tag{4.13}$$

Ova ocena je takodje kompletna i dovoljna. Gustina za $\hat{\alpha}$ je

$$f(x) = \frac{(\alpha n)^n}{\Gamma(n)}x^{-(n+1)}e^{-(\alpha n)/x}, \quad x > 0, \tag{4.14}$$

što je gustina Vinčijeve raspodele (*Vinci*, 1921.). Primetimo da slučaj kada je x_0 poznato nije neobičan u aktuarskoj matematici, na primer korisnik osiguranja bi bio koncentrisan samo na gubitke koji prelaze odredjenu unapred definisanu granicu.

Ukoliko su parametri α, x_0, nepoznati, tada je kompletna, dovoljna statistika (T, S), gde je

$$T = X_{1:n}, \quad S = \sum_{i=1}^{n} \log\left(\frac{X_i}{X_{1:n}}\right), \tag{4.15}$$

i imaju zajedničku gustinu raspodele

$$f(s,t) = \frac{n\alpha^n x_0^{n\alpha}}{\Gamma(n-1)} \cdot \frac{s^{n-2}e^{-\alpha s}}{t^{n\alpha+1}}, \quad x_0 \leq t, 0 \leq s, \tag{4.16}$$

što pokazuje da su $\hat{\alpha}$ i $\hat{x_0}$ nezavisni. Ocena $\hat{\alpha}$ ima marginalnu gustinu rapodele

$$f(x) = \frac{(\alpha n)^{n-1}}{\Gamma(n-1)} x^{-n} e^{-\alpha n/x}, \quad x > 0. \tag{4.17}$$

Neka su $x_0 = 1 < x_1 < x_2 < ... < x_{k+1} = \infty$ granice $k+1$ grupe, i predstavljaju broj zapažanja, u slučajnom uzorku obima n, od kojih n_i upada u interval $[x_i, x_{i+1})$ $(\sum_{i=0}^{k-1} n_i = n)$. Tada je ocena parametra α metodom maksimalne verodostojnosti rešenje jednačine

$$\sum_{i=0}^{k-1} n_i \left(\frac{x_{i+1}^{-\alpha} \log x_{i+1} - x_i^{-\alpha} \log x_i}{x_{i+1}^{-\alpha} - x_i^{-\alpha}}\right) + n_k \log x_k = 0. \tag{4.18}$$

Treba napomenuti da ako granica formira geometrijski niz, $x_{i+1} = cx_i, \ i = 1, ..., k$, tada se ocena metodom maksimalne verodostojnosti može izraziti u obliku

$$\hat{\alpha} = \frac{1}{\log c} \log\left(1 + \frac{n}{\sum_{i=0}^{k} i n_i}\right). \tag{4.19}$$

4.3 Nepristrasna ocena Paretovih karakteristika

Poboljšanje ocena dobijenih metodom maksimalne verodostojnosti dobija se uklanjanjem pristrasnosti, odnosno bias-a. U slučaju kada su oba parametra nepoznata, $\hat{x}_0$ se može zameniti ocenom

$$x_0^* = X_{1:n}\left\{1 - \frac{1}{(n-1)\hat{\alpha}}\right\}, \tag{4.20}$$

a $\hat{\alpha}$ sa

$$\alpha^* = \left(1 - \frac{2}{n}\hat{\alpha}\right). \tag{4.21}$$

Ovo su u stvari UMVU ocene parametara, što je diretna posledica činjenice da je $(\hat{\alpha}, \hat{x_0})$ dovoljna statistika, a (α^*, x_0^*) je funkcija od $(\hat{\alpha}, \hat{x_0})$. Do ovih ocena je došao Likeš (*Likeš*, 1969.), koje je pojednostavio Bakster (*Baxter*, 1980.). Disperzije ovih ocena su

$$D(\alpha^*) = \frac{\alpha^2}{n-3}, \quad n > 3, \tag{4.22}$$

i

$$D(x_0^*) = \frac{x_0^2}{\alpha(n-1)(\alpha n-2)}, \quad n > \frac{2}{\alpha}. \tag{4.23}$$

Jednakost (4.22) pokazuje da UMVU ocena skoro dostiže Rao Kramerovu granicu[3] za α, koja je jednaka α^2/n. Reskaliranjem ocene koja je dobijena metodom maksimalne verodostojnosti, UMVU ocena parametra α takodje ima inverznu gama raspodelu. Takodje iako je x_0 pozitivno, x_0^* je negativno ako je $\hat{\alpha} < 1/(n-1)$, što se dešava ukoliko je $2n\alpha/\hat{\alpha} > 2n(n-1)\alpha$. Pošto $2n\alpha/\hat{\alpha}$ ima inverznu gama raspodelu, to se dešava sa malom verovatnoćom.

Ukoliko je x_0 poznato, UMVU ocena od α je

$$\alpha^* = \left(1 - \frac{1}{n}\right)\hat{\alpha}, \tag{4.24}$$

sa disperzijom

$$D(\alpha^*) = \frac{\alpha^2}{n-2}, \quad n > 2. \tag{4.25}$$

Slično, ukoliko je poznato α, UMVU ocena od x_0 je

$$x_0^* = X_{1:n}\left\{1 - \frac{1}{n\alpha}\right\}, \tag{4.26}$$

sa disperzijom

$$D(x_0^*) = \frac{x_0^*}{\alpha n(\alpha n-2)}, \quad n > \frac{2}{\alpha}. \tag{4.27}$$

4.4 Robusna ocena

Problem sa ocenama dobijenim metodom maksimalne verodostojnosti (i više klasičnih ocena), je da su veoma osetljive na ekstremno pomeranje i odstupanje, kao što je greška u podacima. Viktoria (*Viktoria-Feser*, 1993, 1994.) zajedno

[3]Nejednakost Rao Kramera daje donju granicu disperzije nepristrasnih ocena jednog parametra

sa Rončetijem (*Ronchetti*) je predložila robusne[4] alternative za metodu maksimalne verodostojnosti u slučaju modela raspodela prihoda. Hampel (*Hampel*, 1986.) ocenio je robusnost statistike $T_n = T_n(x_1, ..., x_n)$ u terminima funkcije uticaja. Da bi se definisala ova funkcija, pogodno je posmatrati T_n kao funkciju od empirijske funkcije raspodele

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} \delta_{x_i}(x), \tag{4.28}$$

gde δ_x označava masu tačke u x. Ukoliko je $T(F_n) = T_n(x_1, ..., x_n)$, funkcija uticaja (IF), parametarski model F_θ, $\theta \in \Theta \subseteq R^k$ se definiše preko $T(F_n)$, odnosno $T(F_\theta)$, na sledeći način

$$IF(x; T; F_\theta) = \lim_{\epsilon \to 0} \frac{T[(1-\epsilon)F_\theta + \epsilon\delta_x] - T(F_\theta)}{\epsilon}. \tag{4.29}$$

IF opisuje efekat malog upliva na ocenu, $\epsilon\delta_x$, u tački x, standardizovanu preko mase ϵ. Stoga, linearna aproksimacija $\epsilon IF(x; T; F_\theta)$ meri asimptotski loš uticaj izazvan uplivom $\epsilon\delta_x$.

U slučaju metode maksimalne verodostojnosti, IF je proporcionalna vrednosti funkcije $s(x; \theta) = (\partial/\partial\theta) \log f(x; \theta)$. U Paretovom slučaju se dobija

$$s(x; \alpha) = \frac{1}{\alpha} - \log x + \log x_0, \tag{4.30}$$

što je neograničeno u odnosu na x. Dakle, jedna tačka može odvesti ocenu dobijenu metodom maksimalne verodostojnosti proizvoljno daleko. Ovo važi i za mnoge druge raspodele veličine. Poželjna robusna osobina za jednu ocenu je ograničena IF. Ocene, koje poseduju ovu osobinu se odnose na nepristrasnu robusnu ocenu, odnosno B-robusnu ocenu. Optimalnu B-robusnu ocenu (OBRE) je definisao Hampel (1986.), kao jednu od ocena koje pripadaju klasi M[5] ocena, i predstavlja rešenje T_n iz sledeće jednakosti

$$\sum_{i=1}^{n} \psi(x_i; T_n) = 0 \tag{4.31}$$

za neku funkciju ψ.
OBRE je optimalna u smislu da je M ocena koja minimizuje trag asimptotske matrice kovarijacije, pod uslovom da ima ograničenu funkciju uticaja. Postoji više različitih varijanti ove ocene u zavisnosti na koji način je odabrana granica za IF. Viktoria i Rončeti (1994.) su se bavili standardizovanom ocenom OBRE. Ukoliko je granica za IF jednaka c, onda je ocena definisana na sledeći način

[4] Osobina ocene pod kojom se podrazumeva da se izmenom ekstremnih vrednosti u uzorku ne menja vrednost ocene

[5] M klasa ocena je klasa ocena koje se računaju preko minimuma sume funkcije podataka

$$\sum_{i=1}^{n} \psi(x_i; T_n) = \sum_{i=1}^{n} \{s(x_i; \theta) - a(\theta)\} W_c(x_i; \theta) = 0, \tag{4.32}$$

gde je

$$W_c(x_i; \theta) = \min \left\{ 1; \frac{c}{||A(\theta)\{s(x; \theta) - a(\theta)\}||} \right\}. \tag{4.33}$$

Matrica $A(\theta)$ dimenzija $k \times k$ i vektor $a(\theta)$ dimenzije $k \times 1$ su definisani na sledeće načine

$$E\{\psi(x; \theta)\psi(x; \theta)^T\} = \{A(\theta)A(\theta)^T\}^{-1} \tag{4.34}$$

i

$$E\psi(x; \theta) = 0. \tag{4.35}$$

Ideja OBRE ocene je da se dobije ocena koja je slična, koliko je to moguće, oceni dobijenoj metodom maksimalne verodostojnosti, na podacima najvećeg obima, zbog efikasnosti.

4.5 Ostale ocene

Kang i Čo (*Cho*, 1997.) su izveli MSE-optimalnu ocenu parametra α, za nepoznato x_0, u klasi ocena oblika $c/\log(X_j/\hat{x})$, na sledeći način

$$\widetilde{\alpha} = \frac{n-3}{\sum_{i=1}^{n} \log(X_i/\hat{x_0})}, \tag{4.36}$$

sa MSE

$$MSE(\widetilde{\alpha}) = \frac{\alpha^2}{n-2}, \quad n > 2. \tag{4.37}$$

Ovu ocenu su nazivali ocenom minimalnog rizika (MRE[6]). Poredeći je sa ocenom koja je dobijena metodom maksimalne verodostojnosti, ona pored toga što ima manji MSE ima manji i bias

$$bias(\widetilde{\alpha}) = -\frac{\alpha}{n-2}, \tag{4.38}$$

što predstavlja polovinu bias-a ocene maksimalne verodostojnosti, po apsolutnoj vrednosti.

[6] minimum risk estimator

5 Empirijski rezultati

5.1 Empirijski rezultati dobijeni iz podataka o prihodima

U ispitivanju vezanom za Paretovu raspodelu II tipa, Kredi (*Creddy*, 1977.) je primetio protivrečnosti u Paretovim iskazima. On je izračunao ocenu parametra lokacije μ koja je veoma značajna i ima suprotan znak, u poredjenju sa Paretovim radom, što je kontradiktorno Paretovim napomenama da je parametar pomeranja μ pozitivan kada su u pitanju zarade na poslu.

U studiji koja je za cilj imala preispitivanje da li zarada aproksimativno ima lognormalnu raspodelu, Harison (*Harrison*, 1981.) je posmatrao bruto nedeljnu zaradu 91 968 muškaraca preko 21 godine starosti, koji su radili puno radno vreme, iz britanskog popisa zarade 1972. godine. Bili su razvrstani u 34 grupe prema prihodima. Takodje za date podatke glavni deo raspodele pokriva 85% ukupnog broja zaposlenih i "prihvatljivo je dobro opisan" lognormalnom raspodelom sa gornjim repom, ali je u poredjenju sa Paretovom raspodelom znatno lošije.

Ratc (*Ratz*) i van Šerenberg (*van Scherrenberg*) (1981.) su ispitivali raspodelu prihoda registrovanih profesionalnih inženjera, u svim strukama, u Ontariju u Kanadi, godišnje, za period od 1955-1978. godine. Modeliranjem veze izmedju parametara raspodele i godina iskustva regresionim tehnikama, pronašli su negativan odnos izmedju godina iskustva i parametra α. Ovim se dolazi do toga da je zarada inženjera sa više iskustva veća od zarade inženjera početnika, što je sasvim intuitivni zaključak.

Ransom (*Ransom*) i Kramer (*Cramer*) (1983.) su predložili da počnu da koriste model merenja greške, posmatrajući prihod y kao sumu dve nezavisne promenljive, $y = x+u$, gde je x sistematska komponenta a u je greška sa raspodelom $N(0, \sigma^2)$. Podešavajući ovakav model sa sistematskom komponentom koja ima Paretovu raspodelu, na prihodima u SAD-u za 1960. i 1969. godinu, otkrili su da greška odgovorna za trećinu totalne varijacije, dovodi do toga da osnovna raspodela postaje besmislena. Ovi rezultati su još gori i za druga tri parametarska modela, koji su definisani kao Sing-Madala (*Singh - Maddala*) raspodela.

Kavl (*Cowell*), Ferera (*Ferreira*) i Ličifild (*Litchfield*) (1998.) su izučavali raspodele prihoda u Brazilu oko 1980. godine, period medjunarodne krize. Brazil je posebno zanimljiva zemlja za naučna istraživanja raspodele prihoda zato što ima jednu od najneravnomernijih raspodela u svetu, 51% ukupnog prihoda ide bogatiima kojih ima 10%, a samo 2.1% ide siromašnima kojih ima 20% prema podacima iz 1995. godine. Raspodele prihoda većih od $1000 su modelirane Paretovom raspodelom.

Analizirajući ekstreman desni rep prihoda u Japanu i poreza na dobit za 1998. godinu, Ajoma (*Aoyoma*, 2000.) je izračunao da je α oko 2.

Iz ovih istraživanja proizilazi da je Paretova raspodela u velikom broju slučajeva nepodesna za aproksimaciju cele raspodele prihoda, ali da se raspodela može koristiti uspešno za interpolaciju grupisanih podataka o prihodima, gde je obično poželjno uvesti pretpostavku da su prihodi veći od \$1000000.

5.2 Empirijski rezultati dobijeni iz podataka o bogatstvu

Stajndl (1972.) je izračunao ocenu za α oko 1.7, na podacima o bogatstvu u Švedskoj 1955. i 1968. godine. Za podatke u Danskoj, otkrio je da je Paretov koeficijent neznatno porastao u odnosu na 1959. godinu kada je iznosio 1.45 i 1967. godinu kada je iznosio 1.52.

Čester (*Chester*, 1979.) je ispitivao Paretov model tipa I, na podacima o bogatstvu u Irskoj 1966. godine, koje je grupisano u 26 kategorija, a parametar α je ocenio sa 0.45. Ispostavilo se da je lognormalna raspodela bolja za aproksimaciju na ovim podacima.

5.3 Empirijski rezultati o veličinama firmi

U klasičnoj studiji o veličinama firmi, Stajndl (1965.) je izračunao Paretov koeficijent, koji se kreće od 1 do 1.5. Za sve firme u SAD-u 1931. i 1955. godine posmatrajući ih po aktivi, parametar α je jednak 1.1; za nemačke firme 1950. i 1959. godine posmatrajući promet, je jednak 1.1 u proizvodnji i 1.3 u maloprodaji, a za nemačke firme 1954. godine po zaposlenosti u proizvodnji je oko 1.2.

Kvand 1966. godine je istraživao raspodele veličine firmi, veličina se meri u odnosu na aktivu, u SAD-u. Koristeći listu Fortune od 500 najvećih kompanija 1955. i 1960. godine, i 30 primera reprezentativnih industrija, zaključio je da je Paretova raspodela tipa I, II, III odgovarajući model samo za pola uzoraka. Najbolje tri su modelirane tipom III, dok klasična raspodela I tipa, odgovara samo za 6 slučajeva. Paretova raspodela I i II tipa je prikladna samo za dva Fortunina primera, stoga se može reći da se lognormalna raspodela može smatrati boljom od Paretove raspodele za ovaj tip podataka.

Engval (*Engwall*, 1968. godine) je izučavao najveće firme, u pogledu prodaje, za godinu 1965. na pet područja SAD, zemlje van SAD-a, Evropa, Skandinavija, Švedska. Uspeo je da izračuna parametar α koji se kreće u intervalu od 1 do 2 u svim slučajevima.

Okajam (*Okuyama*) i Takajasi (*Takayasu*) (1999.) su izračunali Paretov koeficijent u intervalu od 0.7 do 1.4. U obzir su uzeli godišnji prihod kompanija, koristeći *Moody's* podatke o kompanijama i *Moody's* podatke o internacionalnim kompanijama, gde Japanske kompanije imaju prihod preko 40 miliona jena. Ispostavilo se da postoji razlika ne samo izmedju zemalja, već i izmedju različitih indstrijskih sektora.

Zaključak je da je Paretov koeficijent za veličine firmi, manji nego za prihod firme, sa granicom koja je veća za 2, što implicira da raspodele veličine firme pripadaju domenu privlačnosti stabilnog zakona koji nije normalan.

5.4 Empirijski rezultati dobijeni iz podataka o gubicima u osiguranju

Jedan broj istraživača je predložio Paretovu raspodelu kao prihvatljiv model za gubitak od požara.

Benkert i Sternberg (1957.) su predložili Paretovu raspodelu za raspodelu gubitka od požara u kućama u Švedskoj u periodu od 1948. do 1952. godine, vrednost ocene parametra je približno 0.5 za štetu četiri tipa kuća.

Anderson (*Andersson*, 1971.) je koristio Paretovu raspodelu da modelira gubitke usled požara u Danskoj, Finskoj, Norveškoj i Švedskoj, za period od 1951. do 1958. godine i od 1959. do 1966. godine, sa ocenom koeficijenta koja se kreće izmedju 1.25 do 1.76 i potvrdjujući medjunarodni trend porasta broja velikih šteta tokom prvog perioda u odnosu na drugi period, a opadanje vrednosti koeficijenta α.

Kao što je bio slučaj sa veličinama firmi, vredi pomenuti da iz podataka o gubicima usled požara, proizilazi α koje je manje od 2, što vodi ka stabilnom zakonu koji nije normalan. Benktander (*Benktander*, 1962.) je za podatke o osiguranjima automobila izračunao veliku vrednost za parametar $\hat{\alpha} = 2.7$.

Veliki broj studija koje koriste Paretovu raspodelu, u aktuarskoj literaturi, ne koriste klasičnu Paretovu raspodelu, već Paretovu raspodelu tipa II.

6 Raspodele Paretovog tipa

6.1 Stopina raspodela

Stopa (1990.a, 1990.b) je predložio generalizaciju klasične Paretove raspodele, uvodjenjem stepene transformacije Paretove funkcije raspodela. Tako je Stopina raspodela oblika

$$F(x) = \left[1 - \left(\frac{x}{x_0}\right)^{-\alpha}\right]^{\theta}, \quad 0 < x_0 \leq x, \tag{6.1}$$

gde su α, $\theta > 0$. Klasična Paretova raspodela je slučaj kada je $\theta = 1$. Gustina Stopine raspodele je

$$f(x) = \theta\alpha x_0^{\alpha} x^{-\alpha-1}\left[1 - \left(\frac{x}{x_0}\right)^{-\alpha}\right]^{\theta-1}, \quad 0 < x_0 \leq x, \tag{6.2}$$

a funkcija kvantila je

$$F^{-1}(u) = x_0(1 - u^{1/\theta})^{-1/\alpha}, \quad 0 < u < 1. \tag{6.3}$$

Stopina generalizacija Paretove raspodele može biti izvedena iz diferencijalne jednačine za elastičnost raspodele $\eta(x, F)$. Ukoliko se pretpostavlja da je (1) gustina prihoda ili unimodalna ili opadajuća, (2) domen raspodele je $[x_0, \infty)$, za neko $x_0 > 0$, (3) $\eta(x, F)$ je opadajuća funkcija po $F(x)$, gde je $\lim_{x \to x_0} \eta(x, F) = \infty$ i $\lim_{x \to \infty} \eta(x, F) = 0$, onda je diferencijalna jednačina

$$\eta(x, F)\left[= \frac{F^{'}(x)}{x} \cdot x\right] = \alpha\theta\frac{1 - [F(x)]^{1/\theta}}{[F(x)]^{1/\theta}}, \quad \alpha, \theta > 0, \tag{6.4}$$

što vodi do gustine raspodele (6.2).

Za celobrojne vrednsti η i θ raspodela se može posmatrati kao raspodela $X_{n:n}$ Paretove raspodele. Dakle, sama raspodela je zatvorena za maksimum, odnosno

$$X \sim Stoppa(x_0, \alpha, \theta) \Rightarrow X_{n:n} \sim Stoppa(x_0, \alpha, n\theta). \tag{6.5}$$

Suprotno od Stopine raspodele, Paretova raspodela je zatvorena za minimum.

Poredeći sa klasičnom Paretovom raspodelom, Stopina raspodela je više fleksibilna, jer ima dodatni parametar θ koji omogućava i unimodalnu (kada je $\theta > 1$) i nula modalnu (kada je $\theta \leq 1$) gustinu raspodele. Moda je jednaka

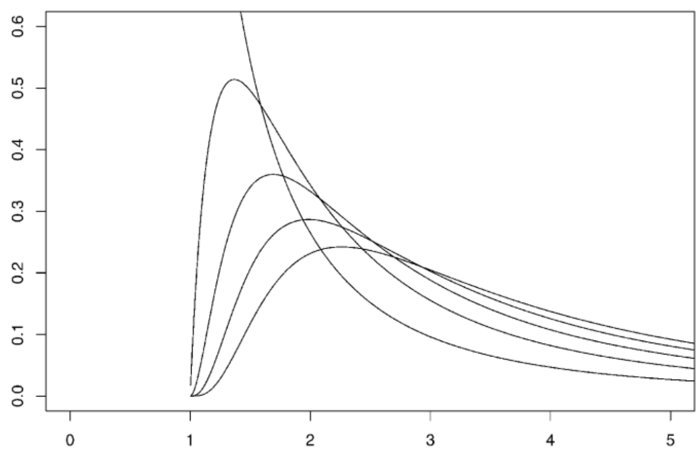
$$x_{mode} = x_0\left(\frac{1 + \theta\alpha}{1 + \alpha}\right)^{1/\alpha}, \quad \theta > 1, \tag{6.6}$$

u tački x_0. Kako θ raste, moda se pomera u desno. Slika (6.1) pokazuje efekat novog parametra θ.

Za $k < \alpha$, k-ti momenat postoji i jednak je

$$E(X^k) = \theta x_0^k B\left(1 - \frac{k}{\alpha}, \theta\right). \qquad (6.7)$$

Na slici (6.1) su prikazani grafici gustine Stopine raspodele za različite parametre.



Slika 6.1: gustina Stopa raspodele sa parametrima $x_0 = 1$, $\alpha = 1.5$ i $\theta = 1(1)5$ s leva na desno

Parametri iz jednakosti (6.1) mogu biti ocenjeni na više načina. Stopa (1990.b, 1990.c) je razmatrao nelinearnu metodu najmanjih kvadrata, kao i metodu maksimalne verodostojnosti. Odgovarajuće ocene nije moguće dobiti u eksplicitnom obliku, već moraju da se dobiju numerički. U slučaju ocene parametra x_0 metodom maksimalne verodostojnosti, definišući krajnju tačku domena funkcije F dolazi do uobičajenog problema koji je nastao zbog granice parametara, funkcija verodostojnosti je neograničena u slučaju x_0. Stopa je predložio korišćenje modifikovanog modela maksimalne verodostojnost. Početne vrednosti za ocenu metodom maksimalne verodostojnosti mogu biti izračunate iz, na primer, regresione ocene koju je predložio Stopa 1995. godine. Došao je do zaključka da funkcija raspodele može biti napisana na drugi način

$$\log\{[1 - F^{1/\theta}(x)]\} = \alpha \log x_0 - \alpha \log x, \qquad (6.8)$$

a elastičnost

$$\log \eta(F, x) = \log(\theta \alpha x_0^{\alpha}) - \alpha \log x - \frac{1}{\theta} \log F(x). \qquad (6.9)$$

Prema tome, $\log(\theta\alpha x_0^{\alpha})$, α i $1/\theta$ mogu biti ocenjeni metodom najmanjih kvadrata, a ocene za originalne parametre mogu biti izvedene rešavanjem diferencijalnih jednačina novih parametara.

Stopa (1990.b) je predložio i drugu generalizaciju Paretove raspodele, koju zovemo Stopina raspodela tipa II, njena funkcija raspodele je

$$F(x) = \left\{1 - \left[\frac{x-c}{x_0}\right]^{-\alpha}\right\}^{\theta}, \quad x > c, \tag{6.10}$$

koja je dvoparametarska Stopina raspodela tipa I sa parametrima skaliranja i pomeranja, c, x_0. Može se izvesti na sličan način kao i raspodela tipa I.

6.2 Konusna raspodela

Uprkos tome što to nije bilo objavljeno, Hautaker (*Houthakker*, 1992.) je privukao pažnju malog kruga istraživača i korisnika statističkih raspodela prihoda.

Hautaker uvodi fleksibilnu familiju generalizovane Paretove raspodele definisane u terminima konusnog preseka. Konusni presek Paretovog dijagrama je predstavljen na sledeći način

$$c_0U^2 + 2c_1UV + c_2V^2 + 2c_3U + 2c_4V + c_5 = 0, \tag{6.11}$$

gde je $U = \log\{\bar{F}(x)\}$, $V = \log x$, a $c_0, c_1, ..., c_5$ su parametri. Da bi se dobila Paretova raspodela, potreban je konusni presek sa linearnom asimptotom. Zbog toga ne treba uzeti u obzir krugove i elipse i jedini prihvatljivi konusni presek je hiperbola. Pošto $\exp U$ mora da definiše funkciju preživljavanja, moraju se definisati dalja ograničenja. Navešćemo samo osnovna svojstva konusne raspodele koristeći reparametrizaciju koju je predložio Klajber (*Kleiber*, 1994.).

Konusna raspodela je oblika

$$F(x) = 1 - \exp\left\{-\eta\sqrt{1+(\log x - \lambda)^2} + \xi(\log x - \lambda) + \mu\right\}, \quad x \geq x_0 > 0, \tag{6.12}$$

gde je $\eta > 0$, $-\infty < \xi \leq \eta$, $\mu = \eta\sqrt{1+(\log x_0 - \lambda)^2} - \xi(\log x_0 - \lambda)$. Takodje je potrebno da važi

$$\frac{\eta(\log x_0) - \lambda}{\sqrt{1+(\log x_0 - \lambda)^2}} - \xi \geq 0. \tag{6.13}$$

Odgovarajuća gustina raspodele je

$$f(x) = x^{-\xi-1}\exp\left\{\eta\sqrt{1+(\log x-\lambda)^2}-\xi\lambda+\mu\right\}\left[\frac{\eta(\log x-\lambda)}{\sqrt{1+(\log x-\lambda)^2}}-\xi\right]. \tag{6.14}$$

Hautaker je razmotrio dve potklase konusnih raspodela. To su konusno-linearna raspodela definisana uslovom $\eta - \xi = 0$ i konusno kvadratna definisana sa $f(x_0) = 0$. Stoga za nekvadratnu konusnu raspodelu uvek imamo $f(x_0) > 0$.

Nije teško primetiti da se funkcija raspodele asimptotski ponaša kao $x^{-\eta-\xi}$, medjutim k-ti momenti ove raspodele postoje za $k < \eta - \xi$ i $\eta - \xi$ ima ulogu Paretovog parametra α.

6.3 Log-prilagodjena Paretova raspodela

Zibak (*Zeibach*, 2000.) je predložio generalizaciju Paretove raspodele uvodjenjem logaritamskog prilagodjavanja što bi omogućilo fleksibilniji oblik. On je pošao od Paretovog dijagrama definišući funkciju raspodele u sledećem obliku

$$\log\{1-F(x)\} = -\alpha\log x - \beta\log(\log x) + \gamma. \tag{6.15}$$

Novi oblik ne utiče na asimptotsku linearnost u duploj logaritamskoj reprezentaciji, tako da rezultujuća raspodela ima slab Paretov zakon. Uslov $F(x_0) = 0$, za neko $x_0 > 0$, donosi $e^\gamma = x_0^\alpha(\log x_0)^\beta$, što vodi do funkcije raspodele

$$F(x) = 1-\left(\frac{x_0}{x}\right)^\alpha\left(\frac{\log x_0}{\log x}\right)^\beta,\ \ 1 < x_0 \le x, \tag{6.16}$$

gde je ili $\alpha > 0$ i $\beta \ge -\alpha\log x_0$ ili $\alpha = 0$ i $\beta > 0$. Gustina raspodele je

$$f(x) = \frac{x_0^\alpha(\log x_0)^\beta(\alpha\log x+\beta)}{x^{\alpha+1}(\log x)^{\beta+1}},\ \ 1 < x_0 \le x, \tag{6.17}$$

koja je opadajuća za $\beta \ge 0$ i unimodalna za

$$\beta \in \left[-\alpha\log x_0, -\alpha\log x_0+\frac{1}{2}\left\{\sqrt{(\log x_0+1)^2+4\alpha\log x_0}-(\log x_0+1)\right\}\right]. \tag{6.18}$$

Jedino ako je $\beta = -\alpha\log x_0$ onda je $f(x_0) = 0$, u suprotnom $f(x_0) > 0$.

Momenti ove raspodele su

$$E(X^k) = \frac{\alpha x_0^k}{\alpha-k} - k\beta(\alpha-k)^{\beta-1}x_0^\alpha(\log x_0)^\beta\Gamma[-\beta;(\alpha-k)\log x_0], \tag{6.19}$$

kada je $k < \alpha$ ili kada je $k = \alpha$ i $\beta > 1$,

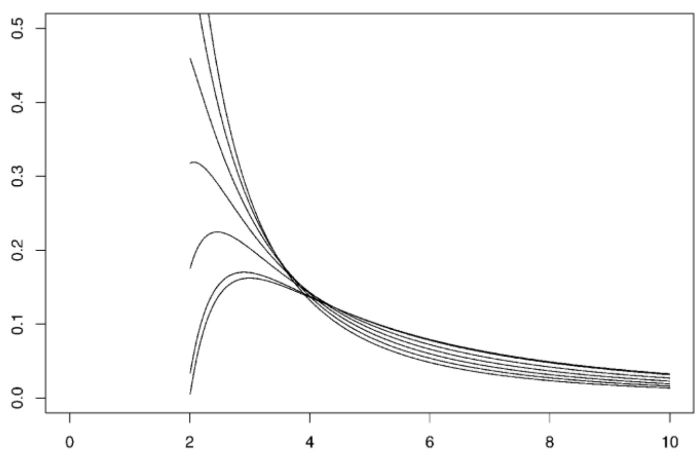
$$E(X^k) = x_0^k + kx_0^k \frac{\log x_0}{\beta - 1}, \tag{6.20}$$

u suprotnom ne postoje. Jednakost (6.19) pokazuje efekat novog parametra β. Prvi deo je k-ti momenat klasične Paretove raspodele, drugi deo opisuje prilagodjavanje zbog β.

Log-Paretova raspodela je vrlo slična log-gamma raspodeli

$$f(x) = \frac{\alpha^\beta x_0^\alpha}{\Gamma(\beta)} x^{-\alpha-1} \left\{ \log\left(\frac{x}{x_0}\right) \right\}^{\beta-1}, \quad 0 < x_0 \leq x, \tag{6.21}$$

gde je $\beta \geq 1, \alpha > 1$. Na slici (6.2) je prikazan grafik gustine log-prilagodjene raspodele.



Slika 6.2: Log-prilagodjena Paretova raspodela $x_0 = 2$, $\alpha = 1.5$ i $\beta = -1.03, -1(.2)0$, posmatarjući odozdo na levo

6.4 Stabilne raspodele

Oko 1960. godine, Mandelbrot je pretpostavio da prihodi imaju Pareto-Levi raspodelu, koja predstavlja raspodelu koja je maksimalno asimetrična stabilna raspodela sa eksponentom α koji se kreće u intervalu izmedju 1 i 2. Stabilnost znači, da ako su X_1 i X_2 nezavisne kopije od X i b_i, $i = 1, 2$, su pozitivne konstante, onda važi

$$b_1 X_1 + b_2 X_2 \stackrel{d}{=} bX + a, \tag{6.22}$$

za neko pozitivno b i realno a.

Danas postoji veoma obimna literatura o stabilnim raspodelama, delimično zbog uspešne primene, naročito u finansijama, a delimično zbog napretka u statističkom računanju. Njihova glavna prednost je da one predstavljaju granične raspodele normalizovanih suma nezavisnih i jednako rasporedjenih slučajnih promenljivih, pa stoga generalizuju normalnu raspodelu. Zaista, stabilni zakoni koji nisu normalni, nastaju kod centralno graničnog problema za nezavisne i jednako rasporedjene slučajne promenljive, ukoliko ove promenljive nemaju konačnu disperziju. U tom slučaju, granična raspodela nema konačnu disperziju, desni rep je pravilno promenljiv i Paretovog tipa, ali ne mora da ima Paretovu raspodelu, što je razlog za uključivanje stabilnih raspodela.

Mandelbrot smatra da je ukupan prihod bilo kog primaoca, dobijen kao suma prihoda iz različitih izvora (na primer prihod porodice, predstavlja sumu prihoda svih članova porodice). Ako sve vrste prihoda imaju isti tip raspodele i ako se može napraviti pretpostavka o nezavisnosti, može se očekivati da totalni prihod bude dobro aproksimiran stabilnom raspodelom.

Ovo zvuči jako primamljivo, mana je da su stabilni zakoni generalno definisani preko svojih karakterističnih funkcija, osim u tri slučaja a to su normalna, Košijeva raspodela i specijalan slučaj inverzne gama raspodele, kada se znaju funkcije raspodele i njihove gustine. Karakteristična funkcija stabilnog zakona je

$$\log \phi(t) = \log E(e^{itx}) = i\mu t - \lambda |t|^{\alpha} \exp\big\{\frac{\pi}{2} i\gamma(t)\mathrm{sign}(t)\big\}. \tag{6.23}$$

Ovde je $0 < \alpha < 1$ ili $1 < \alpha \leq 2$. Za $\alpha = 1$, što uključuje familiju Košijeve raspodele, mora se koristiti malo drugačija forma. Parametar α se obično odnosi na karakteristični eksponent ili indeks stabilnosti, on je u odredjenom smislu najvažniji od četiri parametra jer upravlja ponašanjem repa raspodele, stoga i postojanjem momenata. Apsolutni k-ti momenat $E|X|^k$ postoji ako je $k < \alpha$. Parametar $\mu \in R$ je parametar lokacije, $\alpha > 0$ je parametar skaliranja, a γ je parametar zakrivljenja, sa $|\gamma| \leq 1 - |1 - \alpha|$. Kada je $\alpha = 2$ raspodela je normalna. Postoji mnoštvo parametrizacija stabilnih zakona. Oblik koji je ovde naveden je naveo Zolotarev (1986.).

Veza sa raspodelama prihoda je sledeća: ukoliko je $1 < \alpha \leq 2$, što je povezano sa originalnim Paretovim koeficijentom $\alpha = 1.5$, tada je $|\gamma| \leq 2 - \alpha$. Raspodela je maksimalno zakrivljena u desno ako je $\gamma = 2 - \alpha$. Mandelbrot je predložio da je ovo odgovarajući slučaj za primenu na podacima o prihodu, jer u suprotnom verovatnoća negativnih prihoda postaje prevelika.

Van Dejk (*Van Dijk*) i Kluk (*Kloek*) (1978. i 1980.) su obratili pažnju na problem ocene grupisanih podataka. Oni su iskoristili maksimum verodostojnosti i minimum χ^2 ocene koje su zasnovane na numeričkoj inverziji karakteristične funkcije, koji su usledili iz numeričke integracije rezultujućih gustina. Van Dejk i Kluk su ocenili α za raspoloživi prihod u Australiji u periodu od 1966. do 1968. godine i za ukupan prihod u Danskoj 1976. godine, dobijene vrednosti za parametar su se kretale u intervalu od 1.17 do 1.72.

Takodje su razmatrali log-stabilne raspodele, pretpostavljajući da ne prihod već njegov logaritam ima stabilnu raspodelu. Ovo je potpuno prirodno jer je normalna raspodela posebna stabilna raspodela, kao i raspodele koje su od nje nastale, lognormalna je klasična raspodela veličine.

U aktuarstvu, Holkomb (*Holcomb*, 1973.), je izložio teoriju da je zahtev za neživotno osiguranje mešavina nezavisnih slučajnih promenljivih, koje nisu iz iste raspodele i sve pripadaju domenu atrakcije stabilnog zakona sa uslovom da je $\alpha < 1$. Ovo implicira da raspodela zakona za odštetu konvergira ka stabilnom zakonu.

Holkomb je ispitivao tri skupa podataka. Prvi sadrži podatke za period od 1965. do 1970. godine, 2326 nacionalnih gubitaka usled provala u SAD-u u lancima maloprodajnih objekata. Drugi sadrži 2483 evidencija vandalizma za isti period za isti lanac, a treći sadrži 1142 evidencija kombinovanih istorija gubitaka na svu aktivu i opasnosti na pasivu različitih lanaca. Ocene za parametar α su imale vrednosti 0.82, 0.88, 0.67.

Problem sa ovim pristupom je da raspodela ima beskonačno očekivanje što je teško pomiriti sa opštim aktuarskim principima za konačan prvi momenat. Ovo može objasniti do neke mere zašto se maksimalno zakrivljena stabilna raspodela manje koristi u aktuarskoj matematici.

6.5 Ostale raspodele Paretovog tipa

Krisnan (*Krishnan*), Eng (*Ng*) i Šihade (*Shihadeh*) (1990.) su predložili generalizaciju Paretove raspodele na sledeći način

$$\frac{\partial \log\{1 - F(x)\}}{\partial \log x} = -\alpha - \sum_{i=1}^{k} \beta_i x^i, \tag{6.24}$$

što dovodi do funkcije raspodele

$$1 - F(x) \approx x^{-\alpha} \exp\Big\{ - \sum_{i=1}^{k} \beta_i x^i \Big\}, \tag{6.25}$$

koja se zove polinomijalna Paretova kriva koja obuhvata i Pareto raspodelu tipa I u specijalnom slučaju, kada je $\beta_1 = ... = \beta_k = 0$.

Raspodela veličine nagrade u popularnim lutrijama i nagradnim takmičenjima je takodje blisko povezana sa Paretovom raspodelom. Konstatujući da se nagrade u lutrijama u Nemačkoj obično karakterišu preko činjenice da je broj nagrada n_i koje pripadaju klasi i obrnuto proporcionalan njihovoj vrednosti x_i, Bomsdorfa (*Bomsdorf*, 1977.) je razmatrao diskretnu raspodelu verovatnoća

$$f(x) = \frac{p}{d_i}, \ \ x_i = d_i \cdot a, \ \ i = 1, ..., k, \tag{6.26}$$

gde je $a > 0$, $d_1 = 1$, $d_i < d_{i+1}$, $p = 1/\sum_{i=1}^{k} d_i^{-1}$, koju je nazvao raspodela nagrade na takmičenju. Analogno se dobija gustina za neprekidan slučaj

$$f(x) = \frac{c}{x}, \ \ a_1 \leq x \leq a_2. \tag{6.27}$$

Ako i interval $[a^n, a^{n+1}]$, koji je nosač, napišemo u drugačijem obliku, gde je $n = \log_a a_1 \in R$ i $a = a_2/a_1$, za konstantu normalizacije c se dobija da je jednaka $c = \log_a e$. Tako da za slučaj kada je $a = e$, stvari se pojednostavljuju do hiperboličke funkcije

$$f(x) = \frac{1}{x}, \ \ e^n \leq x \leq e^{n+1}, \tag{6.28}$$

sa funkcijom raspodele

$$F(x) = \log x - n, \ \ e^n \leq x \leq e^{n+1}, \tag{6.29}$$

što je dvostruko sasečena Paretova raspodelu sa parametrom $\alpha = 0$.

Funkcija generisanja momenata raspodele nagrade na takmičenju je data sledećom formulom

$$m(t) = 1 + \sum_{i=1}^{\infty} t^i \frac{e^{in+i}}{i \cdot i!} - \sum_{i=1}^{\infty} t^i \frac{e^{in}}{i \cdot i!}, \tag{6.30}$$

odakle se dobijaju prvi momenat i disperzija

$$E(X) = e^{n+1} - e^n \tag{6.31}$$

i

$$D(X) = 2 \cdot e^{2n+1} - \frac{e^{2n+2}}{2} - \frac{3e^{2n}}{2}. \tag{6.32}$$

Generalizovana Paretova raspodela ima oblik

$$F_c(x) = \begin{cases} 1 - (1 - \frac{cx}{b})^{1/c}, & \text{ako } c \neq 0, \\ 1 - e^{-x} & \text{ako } c = 0, \end{cases} \tag{6.33}$$

gde je

$$x \geq 0 \text{ ako je } c \geq 0, \tag{6.34}$$

$$0 \leq x \leq -\frac{1}{c} \text{ ako je } c < 0, \tag{6.35}$$

i od velike je važnosti u istraživanju ekstremnih dogadjaja. U kontekstu raspodela prihoda Dargahi-Numbri (*Dargahi - Noubary*) 1994. godine je ovaj model primenio na 1969 ličnih prihoda u 157 okruga u Teksasu, i životne zarade 50 profesionalnih igrača golfa preko 1980. godine.

Kolombi (*Colombi*) 1990. godine je predložio Paretovu lognormalnu raspodelu, koja je definisana kao raspodela proizvoda $X = Y \cdot Z$ dve nezavisne promenljive, od kojih jedna ima $\text{Par}(1, \alpha)$ raspodelu, a druga ima dvoparametarsku lognormalnu raspodelu. Gustina ove raspodele je

$$f(x) = \frac{\theta}{x^{\theta+1}} \exp\left\{\theta\left(\mu + \frac{\theta\sigma^2}{2}\right)\right\}\Lambda(x; \mu + \theta\sigma^2, \sigma), \quad x > 0, \tag{6.36}$$

gde je $\theta > 0$, $\mu \in R$, $\sigma > 0$ i

$$\Lambda(x; \mu + \theta\sigma^2, \sigma) = \Phi\left(\frac{\log x - \mu - \theta\sigma^2}{\sigma}\right), \quad x > 0, \tag{6.37}$$

označava dvoparametarsku lognormalnu raspodelu.

Za $k < 0$, k-ti momenat Paretove lognormalne raspodele postoji za $k < \theta$ i jednak je

$$E(X^k) = \frac{\theta}{\theta - k} e^{k\mu + (k\mu)^2/2}. \tag{6.38}$$

7 Zaključak

U ovom radu smo izučavali Paretovu raspodelu, koja ima više tipova, od kojih najveću primenu ima Paretova raspodela tipa I, a najmanju Paretova raspodela tipa III.

Paretova raspodela je kroz istoriju bila čest predmet rasprava i sukoba medju naučnicima, tako da je istorija Paretove raspodele i dalje živa i predmet istraživanja , a raspodela predmet daljeg izučavanja.

Nastanak raspodele se vezuje za Pareta, koji je primetio opadajuću linearnu vezu izmedju logaritma prihoda i broja primalaca prihoda. Pored izvodjenja raspodele iz prihoda, raspodela može biti izvedena iz raspodele talenta, odakle se dolazi do zaključka da veličina firme zavisi od veličine talenta. Zatim iz Markovljevog procesa odakle se vidi da je broj prihoda konstantan, jer posle svakog vlasnika postoji njegov naslednik. I iz pretpostavke o hijerarhijskom zaradjivanju u firmama, iz čega proizilazi da je zarada menadžera konstantna proporcija sumarnih plata zaposlenih koje nadgleda.

Paretova raspodela ima polinomijalni desni rep, koji je pravilno promenljiv u beskonačnosti sa indeksom $-\alpha - 1$. Momenti raspodele k-tog reda postoje ako je $k < \alpha$. Postoji veoma blizak odnos izmedju Paretove i eksponencijalne raspodele, tako da se mnoge osobine Paretove raspodele mogu dobiti iz osobina eksponencijalne, kao što je to slučaj sa karakterizacijom.

Lorencove krive ili krive koncentracije Paretove raspodele se nikada ne seku. Izvodjenjem mera nejednakosti, koje su nastale od Lorencove krive, dolazi se do zaključka da ako svaka od mera nejednakosti opada sa porastom α, onda se sam parametar α i njegov inverz mogu smatrati merom nejednakosti.

Parametri Paretove raspodele mogu biti ocenjeni različitim metodama. Ocenjivanjem metodom maksimalne verodostojnosti, se dolazi do problema da je ocena veoma osetljiva na ekstremno pomeranje i odstupanje u podacima, zbog čega su uvedene i robusne ocene.

Primenjujući ovu raspodelu na podacima o prihodu se dolazi do zaključka da je ona često nepodesna za aproksimaciju cele raspodele prihoda, i da je mnogo bolje koristiti je za interpolaciju grupisanih podataka. Što se tiče raspodele veličine firmi često se lognormalna raspodela smatra boljom nego Paretova raspodela. U većini istraživanja vezanih za aktuarstvo uglavnom se koristi Paretova raspodela tipa II.

Paretova raspodela je osnova za nastanak nekih drugih raspodela. Različitim transformacijama Paretove raspodele naučnici su pokušavali da naprave raspodelu koja je fleksibilnija od Paretove. Stepenom transformacijom je nastala Stopina raspodela. Zatim uvodjenjem logaritamskog prilagodjavanja je nastala log-prilagodjena raspodela.

Literatura

[1] Christian Kleiber, Samuel Kotz (2003). *Statistical Size Distributions in Economics and Actuarial Sciences*, New Jersey

[2] Vesna Jevremović (2013). *Rečnik statističkih termina*, Beograd

[3] Ammon, O. (1895). *Die Gesellschaftsordnung und ihre natürlichen Grundlagen*. Jena: Gustav Fischer.

[4] Andersson, H. (1971). An analysis of the development of the fire losses in the Northern countries after the second World War. *ASTIN Bulletin*, 6, 25−30.

[5] Aoyama, H., Souma, W., Nagahara, Y., Okazaki, M. P., Takayasu, H., and Takayasu, M. (2000). Pareto's law for income of individuals and debt of bankrupt companies. *Fractals*, 8, 293−300.

[6] Arnold, B. C. (1983). *Pareto Distributions*. Fairland, MD: International Co-operative Publishing House.

[7] Baxter, M. A. (1980). Minimum variance unbiased estimation of the parameters of the Pareto distribution. *Metrika*, 27, 133−138.

[8] Benckert, L.-G., and Sternberg, I. (1957). An attempt to find an expression for the distribution of fire damage amount. *Transactions of the 15th International Congress of Actuaries*, 2, 288−294.

[9] Benini, R. (1897). Di alcune curve descritte da fenomeni economici aventi relazione colla curva del reddito o con quella del patrimonio. *Giornale degli Economisti*, 14, 177−214. English translation in *Rivista di Politica Economica*, 87 (1997), 701−744.

[10] Benktander, G. (1962). Notes sur la distribution conditionnee du montant d'un sinistre par rapport a l'hypothese qu'il y a eu un sinistre dans l'assurance automobile. *ASTIN Bulletin*, 2, 24−29.

[11] Bhattacharya, N. (1963). A property of the Pareto distribution. *Sankhyā*, B25, 195−196.

[12] Bomsdorf, E. (1977). The prize-competition distribution: A particular L-distribution as a supplement to the Pareto distribution. *Statistische Hefte*, 18, 254−264.

[13] Bowley, A. L. (1926). *Elements of Statistics*, 5th ed. Westminster: P. S. King and Sons.

[14] Champernowne, D. G. (1953). A model of income distribution. *Economic Journal*, 53, 318−351.

[15] Chesher, A. (1979). An analysis of the distribution of wealth in Ireland. *Economic and Social Review*, 11, 1−17.

[16] Colombi, R. (1990). A new model of income distribution: The Pareto−lognormal distribution. In: C. Dagum and M. Zenga (eds.): pp. 18−32.

[17] Cowell, F. A., Ferreira, F. H. G., and Litchfield, J. A. (1998). Income distribution in Brazil 1981−1990. Parametric and non-parametric approaches. *Journal of Income Distribution*, 8, 63−76.

[18] Creedy, J. (1977). Pareto and the distribution of income. *Review of Income and Wealth*, 23, 405−411.

[19] Dargahi-Noubary, G. R. (1994). On distributions of large incomes. *Journal of Applied Statistical Science*, 1, 211−222.

[20] Davis, H. T. (1941a). *The Analysis of Economic Time Series.* Bloomington, IN: Principia Press.

[21] Davis, H. T. (1941b). *The Theory of Econometrics.* Bloomington, IN: Principia Press

[22] Engwall, L. (1968). Size distributions of firms−A stochastic model. *Swedish Journal of Economics*, 70, 138−159.

[23] Feller, W. (1971). *An Introduction to Probability Theory and Its Applications*, Vol. II, 2nd ed. New York: John Wiley.

[24] Fisz, M. (1958). Characterizations of some probability distributions. *Scandinavian Actuarial Journal*, 65−70.

[25] Gini, C. (1909a). Il diverso accrescimento delle classi sociali e la concentrazione della ricchezza. *Giornale degli Economisti*, 38, 27−83.

[26] Hampel, F. R., Ronchetti, E. M., Rousseeuw, P., and Stahel, W. A. (1986). *Robust Statistics: The Approach Based on Influence Functions.* New York: John Wiley.

[27] Harrison, A. J. (1981). Earnings by size: A tale of two distributions. *Review of Economic Studies*, 48, 621−631.

[28] Holcomb, E. W. (1973). Estimating parameters of stable distributions with application to nonlife insurance. Ph.D. thesis, University of Tennessee, Knoxville, TN.

[29] Hossain, A. M., and Zimmer, W. J. (2000). Comparison of methods of estimation for a Pareto distribution of the first kind. *Communications in Statistics− Theory and Methods*, 29, 859−878.

[30] Houthakker, H. S. (1992). Conic distributions of earnings and income. Discussion Paper No. 1598, Harvard Institute of Economic Research, Harvard University.

[31] Johnson, N. L., Kotz, S., and Balakrishnan, N. (1994). *Continuous Univariate Distributions*, Vol. 1, 2nd ed. New York: John Wiley.

[32] Kang, S.-B., and Cho, Y.-S. (1997). Estimation of the parameters in a Pareto distribution by jackknife and bootstrap methods. *Journal of Information & Optimization Sciences*, 18, 289−300.

[33] Kleiber, C. (1994). Hyperbolische Pareto-Diagramme und Einkommensverteilungen. Diploma thesis, Dept. of Statistics, Universität Dortmund, Germany.

[34] Kotz, S., and Shanbhag, D. N. (1980). Some new approaches to probability distributions. *Advances in Applied Probability*, 12, 903−921.

[35] Krishnaji, N. (1970). Characterization of the Pareto distribution through a model of underreported incomes. *Econometrica*, 38, 251−255.

[36] Krishnan, P., Ng, E., and Shihadeh, E. (1990). Some generalized forms of the Pareto curve to approximate income distributions. *ASA Proceedings of the Business and Economic Statistics Section*, 502−504.

[37] Likeš, J. (1969). Minimum variance unbiased estimation of the parameters of power-function and Pareto's distribution. *Statistische Hefte*, 10, 104−110.

[38] Lucas, R. E. (1978). On the size distribution of business firms. *Bell Journal of Economics*, 9, 508−523.

[39] Lydall, H. F. (1959). The distribution of employment incomes. *Econometrica*, 27, 110−115.

[40] Mandelbrot, B. (1960). The ParetoLévy law and the distribution of income. *International Economic Review*, 1, 79−106.

[41] Mandelbrot, B. (1964). Random walks, fire damage amount and other Paretian risk phenomena. *Operations Research*, 12, 582−585.

[42] Meidell, B. (1912). Zur Theorie des Maximums. *Transactions of the 7th International Congress of Actuaries*, Vol. 1, 85−99.

[43] Okuyama, K., Takayasu, M., and Takayasu, H. (1999). Zipf's law in income distribution of companies. *Physica*, A269, 125−131.

[44] Pareto, V. (1895). La legge della domanda. *Giornale degli Economisti*, 10, 59−68. English translation in *Rivista di Politica Economica*, 87 (1997), 691−700.

[45] Pareto, V. (1896). La courbe de la répartition de la richesse. Reprinted 1965 in G. Busoni (ed.): *Œeuvres complées de Vilfredo Pareto, Tome 3: Écrits sur la courbe de la répartition de la richesse*, Geneva: Librairie Droz. English translation in *Rivista di Politica Economica*, 87 (1997), 647−700.

[46] Pareto, V. (1897a). *Cours d'économie politique*. Lausanne: Ed. Rouge.

[47] Pareto, V. (1897b). Aggiunta allo studio della curva delle entrate. *Giornale degli Economisti*, Series 2, 14, 1526. English translation in *Rivista di Politica Economica*, 87 (1997), 645−700.

[48] Pollastri, A. (1990). A comparison of the traditional estimators of the parameter a of the Pareto distribution. In C. Dagum and M. Zenga (eds.): pp. 183−193.

[49] Quandt, R. E. (1966b). Old and new methods of estimation and the Pareto distribution. *Metrika*, 10, 55−82.

[50] Ransom, M. R., and Cramer, J. S. (1983). Income distributions with disturbances.*European Economic Review*, 22, 363−372.

[51] Ratz, H. C., and van Scherrenberg, M. (1981). The distribution of incomes for professional engineers. *IEEE Transactions on Systems, Man, and Cybernetics*, 11, 322−325.

[52] Revankar, N. S., Hartley, M. J., and Pagano, M. (1974). A characterization of the Pareto distribution. *Annals of Statistics*, 2, 599−601.

[53] Rhodes, E. C. (1944). The Pareto distribution of incomes. *Economica* (New Series), 11, 1−11.

[54] Saksena, S. K., and Johnson, A. M. (1984). Best unbiased estimators for the parameters of a two-parameter Pareto distributions. *Metrika*, 31, 77−83.

[55] Shirras, G. F. (1935). The Pareto law and the distribution of income. *Economic Journal*, 45, 663−681.

[56] Shpilberg, D. C. (1977). The probability distribution of fire loss amount. *Journal of Risk and Insurance*, 44, 103−115.

[57] Stamp, J. C. (1914). A new illustration of Paretos law. *Journal of the Royal Statistical Society*, 77, 200−204.

[58] Steindl, J. (1965). *Random Processes and the Growth of Firms.* London: Griffin.

[59] Steindl, J. (1972). The distribution of wealth after a model of Wold and Whittle. *Review of Economic Studies*, 39, 263−280.

[60] Stoppa, G. (1990a). A new generating system of income distribution models. *Quaderni di Statistica e Matematica Applicata alle Scienze Economico-Sociali*, 12, 47−55.

[61] Stoppa, G. (1990b). A new model for income size distributions. In: C. Dagum and M. Zenga (eds.): *Income and Wealth Distribution, Inequality and Poverty: Proceedings of the 2nd International Conference on Income Distribution by Size: Generation, Distribution, Measurement and Applications*. New York, Berlin, London, and Tokyo: Springer, pp. 33−41.

[62] Stoppa, G. (1990c). Proprietà campionarie di un nuovo modello Pareto generalizzato. *Atti XXXV Riunione Scientifica della Società Italiana di Statistica*, Padova: Cedam, 137−144.

[63] van Dijk, H. K., and Kloek, T. (1978). Empirical evidence on Pareto−Lévy and log stable income distributions. Report 7812/E, Econometric Institute, Erasmus University, Rotterdam.

[64] van Dijk, H. K., and Kloek, T. (1980). Inferential procedures in stable distributions for class frequency data on incomes.*Econometrica*, 48, 1139−1148.

[65] Victoria-Feser, M.-P. (1993). Robust methods for personal income distribution models. Ph.D. thesis, University of Geneva, Switzerland.

[66] Victoria-Feser, M.-P., and Ronchetti, E. (1994). Robust methods for personal income distribution models. *Canadian Journal of Statistics*, 22, 247−258.

[67] Vinci, F. (1921). Nuovi contributi allo studio della distribuzione dei redditi. *Giornale degli Economisti e Rivista di Statistica*, 61, 365−369.

[68] Ziebach, T. (2000). *Die Modellierung der personellen Einkommensverteilung mit verallgemeinerten Pareto-Kurven*. Lohmar: Josef Eul.

[69] Zolotarev, V. M. (1986). *One-dimensional Stable Distributions*. Providence, RI: American Mathematical Society.