Äärimmäisten ilmiöiden teoria

Syksy 2014 Harri Nyrhinen
Helsingin yliopisto
Matematiikan ja tilastotieteen laitos

1. Johdanto

Klassinen äärimmäisten arvojen teoria tutkii riippumattomien ja samoin jakautuneiden satunnaismuuttujien $X_1, \dots, X_n$ maksimin

$$M_n = \max(X_1, \dots, X_n)$$

jakaumaa. Maksimi esiintyy usein sovelluksissa, joten motivaatio tarkastelulle on selkeä (esim. X:t voivat kuvata veden pinnan korkeutta, jolloin riittävän suuri maksimi aiheuttaa ongelmia; X:t eivät tässä sovelluksessa ole riippumattomia, mutta ollaan kuitenkin kohtalaisen lähellä klassisen teorian mukaista asetelmaa).

Monissa sovelluksissa on luontevaa mallintaa tarkasteltava prosessi kumulatiivisesti. Esimerkiksi riskiteoriassa X:t kuvaavat vuotuisten tappioiden määrää ja kertyminen summa on kumuloitunut tappio

$$S_n = X_1 + \dots + X_n.$$

Mikäli

$$K_n = \max(S_1, \dots, S_n)$$

ylittää alkuvarallisuuden, tekee yhtiö vararikon. Saman tyyppisiä kysymyksiä esiintyy myös jonoteoriassa.

Muodollisesti M_n:n ja K_n:n tutkimisessa on kyse samasta asiasta. Matemaattisesti näiden kahden maksimin analysointi on kuitenkin erilaista. Mikäli toisaalta liikutaan jakaumien häntäalueilla ts. tarkastellaan todennäköisyyksiä, että maksimi ylittää korkean tason, ovat K_n ja M_n tietyillä jakaumatyypeillä ylittävien samaa suuruusluokkaa. Tyypillinen jakaumatyyppi on ns. paksuhäntäisten jakaumien luokka. Tällöin X_i:t voivat saada suuria arvoja 'kohtalaisen suurella' todennäköisyydellä. Paksuhäntäiset jakaumat ovat varsin suosittuja mallinnuksessa (esimerkiksi riskiteoriassa).

Kurssilla käydään läpi klassisen teorian perustulokset koskien M_n:ää. Toisena asiana käsitellään myös summia ja vertaillaan kevyt- ja paksuhäntäisten jakaumien käyttäytymistä ääritilanteessa.

2. Maksimin rajakäyttäytymisestä

Olkoot $X, X_1, X_2, \ldots$ riippumattomia ja samoin jakautuneita satunnaismuuttujia. Yhteinen kertymäfunktio olkoon F. Merkitään lisäksi

$$\bar{F}(x) = 1 - F(x) = \mathbb{P}(X > x), \quad \forall x \in \mathbb{R}.$$

Klassisessa äärimmäisten arvojen teoriassa (extreme value theory) kiinnostuksen kohteena on maksimin

$$M_n = \max(X_1, \ldots, X_n)$$

asymptoottinen käyttäytyminen. Selvästi

$$(2.1) \quad \mathbb{P}(M_n \le x) = F(x)^n, \quad \forall x \in \mathbb{R}.$$

Periaatteessa maksimin jakauma tunnetaan siis tarkasti, kun F on annettu. Usein on hyödyllistä hahmottaa jakaumaa karkeammalla tarkkuudella. Pyritään seuraavassa etsimään reaalilukujonot (a_n) ja (b_n) sekä ei-degeneroitunut kertymäfunktio G, joille

$$(2.2) \quad \lim_{n\to\infty} \mathbb{P}\left(\frac{M_n - b_n}{a_n} \le x\right) = G(x)$$

kaikissa G:n jatkuvuuspisteissä. Kyseessä on siis heikko suppeneminen (kertymäfunktio on ei-degeneroitunut, ellei se ole muotoa $G(x) = 1(x \ge c)$ jollain $c \in \mathbb{R}$).

Tavoitteena on siis löytää affiinit M_n:n muunnokset, joilla on rajajakauma. Suurilla n:n arvoilla voidaan M_n:n jakaumaa approksimoida keskeisen raja-arvolauseen tapaan.

Seuraavissa tarkasteluissa X, (X_n), F, $\bar{F}$ ja (M_n) ovat edellä esitetyn mukaisia. Heikosta suppenemisesta käytetään merkintää $\xrightarrow{d}$. Tätä käytetään ilman sekaannuksen vaaraa satunnaismuuttujille ja kertymäfunktioille. Esimerkiksi (2.2) kirjoitetaan muodossa

$$(2.3) \qquad \frac{M_n - b_n}{a_n} \xrightarrow{d} G.$$

Mikäli (2.3) pätee jonoille (a_n), $a_n > 0, \forall n$, ja (b_n) ja G on ei-degeneroitunut, sanotaan, että F (tai X) <u>kuuluu jakauman G vaikutuspiiriin maksimin suhteen</u> (belongs to the maximum domain of attraction of G). Merkitään myös $F \in MDA(G)$.

Jonot (a_n) ja (b_n) samoin kuin kertymäfunktio G eivät ole (2.3):sta yksikäsitteisiä. Tätä asiaa selvittää seuraava yleinen tulos, lause 2.1. Sitä ennen todistetaan tekninen lemma.

<u>Lemma 2.1.</u> Olkoon F ei-degeneroitunut kertymäfunktio ja

$$F(ax+b) = F(a'x+b'), \ \forall x \in \mathbb{R},$$

missä $a, a' > 0$ ja $b, b' \in \mathbb{R}$. Silloin $a = a'$ ja $b = b'$.

<u>Todistus.</u> Oletuksen nojalla

$$F(x) = F\left(a\left(\frac{x}{a} - \frac{b}{a}\right)+b\right) = F\left(a'\left(\frac{x}{a} - \frac{b}{a}\right)+b'\right)$$

$$= F(\alpha x + \beta), \ \forall x \in \mathbb{R}, \text{ missä } \alpha = \frac{a'}{a} > 0 \text{ ja } \beta = b' - \frac{a'b}{a}.$$

Voidaan määrittää sellaiset $x_1 < x_2$, että

$$F(x_1) > F(x), \ \forall x < x_1, \text{ ja } \quad F(x_2) > F(x), \ \forall x < x_2.$$

Olkoon nimittäin y sellainen, että $F(y) \in (0,1)$. Valitaan

$$x_1 = \sup\{x \mid F(x) < F(y)\} \in (-\infty, y]\}.$$

Jos $x_1 = y$, niin $F(x) < F(y)$, $\forall x < y = x_1$. Jos taas $x_1 < y$, niin $F(x_1) = F(y)$ (koska F on oikealta jatkuva). Tälläinkin $F(x) < F(x_1)$, $\forall x < x_1$. Samoin löydetään x_2: valitaan ensin y, jolle $F(y) \in (F(x_1), 1]$.
Nyt

$$F(x_1) = F(\alpha x_1 + \beta) \Rightarrow \alpha x_1 + \beta \geq x_1$$
$$F(x_2) = F(\alpha x_2 + \beta) \Rightarrow \alpha x_2 + \beta \geq x_2.$$

Jos olisi $\alpha x_1 + \beta > x_1$, olisi myös $\alpha x + \beta > x_1$ eräälle $x < x_1$ ja siis $F(x) = F(\alpha x + \beta) \geq F(x_1) > F(x)$ ↯. Siis $\alpha x_1 + \beta = x_1$. Samoin $\alpha x_2 + \beta = x_2$, joten

$$\alpha(x_2 - x_1) = x_2 - x_1 \Rightarrow \alpha = 1 \text{ ja } a = a'.$$

Nähdään myös, että $x_1 + \beta = x_1$,

joten $\beta = 0$ ja siis $b' - \frac{a'b}{a} = b' - b = 0$. $\square$

Lause 2.1. Olkoon (F_n) jono kertymäfunktioita ja G ja H kaksi ei-degeneroitunutta kertymäfunktiota. Oletetaan, että

$$(2.4) \quad F_n(a_n x + b_n) \xrightarrow{d} G(x),$$

missä $a_n \in (0, \infty)$ ja $b_n \in \mathbb{R}$, $\forall n$.

Jos

$$(2.5) \quad F_n(\alpha_n x + \beta_n) \xrightarrow{d} H(x),$$

missä $\alpha_n \in (0, \infty)$ ja $\beta_n \in \mathbb{R}$, $\forall n$, niin on olemassa $a \in (0, \infty)$ ja $b \in \mathbb{R}$ siten, että

$$(2.6) \quad \frac{\alpha_n}{a_n} \to a \quad \text{ja} \quad \frac{\beta_n - b_n}{a_n} \to b, \quad n \to \infty.$$

Lisäksi

$$(2.7) \quad H(x) = G(ax + b), \quad \forall x \in \mathbb{R}.$$

Kääntäen, jos $\alpha_n \in (0, \infty)$ ja $\beta_n \in \mathbb{R}$, $n = 1, 2, \ldots$, ovat sellaisia, että raja-arvot (2.6) ovat olemassa ja $a \in (0, \infty)$, $b \in \mathbb{R}$, niin (2.5) pätee ja H on kaavan (2.7) mukainen.

Todistus. Oletetaan, että (2.5) pätee. Olkoot x_1 ja x_2 kaksi erisuurta H:n jatkuvuuspistettä ja $H(x_i) \in (0,1)$, $i = 1,2$. Oletetaan, että olisi jonon

$$(2.8) \qquad \left(\frac{\alpha_n x_1}{a_n} + \frac{\beta_n - b_n}{a_n} \right)$$

osajono siten että

$$\lim_{k\to\infty} \left(\frac{\alpha_{n_k} x_1}{a_{n_k}} + \frac{\beta_{n_k} - b_{n_k}}{a_{n_k}} \right) = \infty .$$

Tällöin (2.4):n ja (2.5):n nojalla

$$1 = \lim_{k\to\infty} F_{n_k}\left(a_{n_k}\left(\frac{\alpha_{n_k} x_1}{a_{n_k}} + \frac{\beta_{n_k} - b_{n_k}}{a_{n_k}} \right) + b_{n_k} \right)$$

$$= \lim_{k\to\infty} F_{n_k}\left(\alpha_{n_k} x_1 + \beta_{n_k} \right) = H(x_1).$$

Saatiin ristiriita, joten jono (2.8) on ylhäältä rajoitettu. Samoin nähdään, että (2.8) on alhaalta rajoitettu. Soveltamalla samaa päättelyä x_2:een nähdään, että myös jono

$$\left(\frac{\alpha_n x_2}{a_n} + \frac{\beta_n - b_n}{a_n} \right)$$

on rajoitettu. Voidaan siis määrätä osajonot ja vakiot $c_1, c_2 \in \mathbb{R}$ siten, että

$$(2.9) \quad \lim_{k\to\infty} \left(\frac{\alpha_{n_k} x_i}{a_{n_k}} + \frac{\beta_{n_k} - b_{n_k}}{a_{n_k}} \right) = c_i, \quad i = 1,2.$$

Nähdään, että on olemassa raja-arvot

$$\lim_{k\to\infty} \frac{\alpha_{n_k}}{a_{n_k}} \doteq a$$

ja

$$\lim_{k\to\infty} \frac{\beta_{n_k} - b_{n_k}}{a_{n_k}} \doteq b.$$

Edelleen (2.4):n ja (2.5):n nojalla mielivaltaisessa H:n jatkuvuuspisteessä x pätee

$$H(x) = \lim_{k\to\infty} F_{n_k}(\alpha_{n_k} x + \beta_{n_k})$$

$$(2.10) \quad = \lim_{k\to\infty} F_{n_k}\left(a_{n_k}\left(\frac{\alpha_{n_k}}{a_{n_k}} x + \frac{\beta_{n_k} - b_{n_k}}{a_{n_k}}\right) + b_{n_k}\right)$$

$$= G(ax+b)$$

ainakin, jos $ax+b$ on G:n jatkuvuuspiste. Koska H ja G ovat heikkoja funktioita, on oltava $a > 0$ ja

$$(2.11) \quad H(x) = G(ax+b), \quad \forall x \in \mathbb{R}.$$

Jos

$$\lim_{k\to\infty} \frac{\alpha_{m_k}}{a_{m_k}} = a'$$

ja

$$\lim_{k\to\infty} \frac{\beta_{m_k} - b_{m_k}}{a_{m_k}} = b'$$

jollekin osajonolle, niin saadaan kuten edellä

$$H(x) = G(a'x + b'), \quad \forall x \in \mathbb{R}.$$

Välttämättä siis $a = a'$ ja $b = b'$. Siis (2.6) pätee ja (2.7) on jo edellä todistettu.

Oletetaan, että (2.6) pätee. Kuten (2.10):stä nähdään, että

$$F_n(\alpha_n x + \beta_n) \xrightarrow{d} G(ax + b). \quad \square$$

Seuraus 2.2. Jos F kuuluu jakauman G vaikutuspiirin maksimin suhteen, niin F kuuluu jakauman H vaikutuspiiriin, jos ja vain jos on olemassa sellaiset $a \in (0, \infty)$ ja $b \in \mathbb{R}$, että

(2.12) $$H(x) = G(ax + b), \quad \forall x \in \mathbb{R}.$$

Todistus. Valitaan lauseessa 2.1 $F_n = F^n$, jolloin

$$F(a_n x + b_n)^n = \mathbb{P}\left(\frac{M_n - b_n}{a_n} \le x\right) \xrightarrow{d} G(x).$$

Lauseen nojalla F voi kuulua vain tyyppiä (2.12) olevien jakaumien vaikutuspiiriin. Valitsemalla $\alpha_n = a a_n$ ja $\beta_n = a_n b + b_n$, $\forall n$, nähdään, että F kuuluu jakauman (2.12) vaikutuspiiriin. $\square$

Jos kertymäfunktioiden H ja G välillä on yhteys (2.12), sanotaan, että H ja G ovat samaa tyyppiä. Relaatio on ilmeisesti ekvivalenssirelaatio kaikkien kertymäfunktioiden joukossa.

Seuraava tulos esittää, että maksimin rajakäyttäytyminen määräytyy olennaisesti ottaen X:n oikeasta hännästä.

<u>Lause 2.3.</u> Olkoon $\tau \in [0, \infty]$ ja (u_n) reaalilukujono. Silloin

(2.13) $$\lim_{n \to \infty} n \bar{F}(u_n) = \tau$$

jos ja vain jos

(2.14) $$\lim_{n \to \infty} \mathbb{P}(M_n \leq u_n) = e^{-\tau}.$$

<u>Todistus.</u> Olkoon ensin $\tau < \infty$. Selvästi

$$\mathbb{P}(M_n \leq u_n) = \left(1 - \bar{F}(u_n)\right)^n$$
$$= e^{n \log(1 - \bar{F}(u_n))}.$$

Jos (2.13) pätee, niin $\bar{F}(u_n) \to 0$. Koska

$$\log(1-x) = -x + o(x), \text{ kun } x \to 0,$$

niin

(2.14.1) $$\mathbb{P}(M_n \leq u_n) = e^{-n\bar{F}(u_n)(1+o(1))}, \text{ kun } n \to \infty.$$

Tästä seuraa (2.14). Jos toisaalta (2.14) pätee, niin

$$\bar{F}(u_n) = 1 - \mathbb{P}(M_n \leq u_n)^{\frac{1}{n}}$$
$$= 1 - (e^{-\tau} + o(1))^{\frac{1}{n}}$$
$$= 1 - e^{\frac{\log(e^{-\tau} + o(1))}{n}} \longrightarrow 0, \text{ kun } n \to \infty.$$

Siis (2.14.1) pätee ja

$$-\tau = \lim_{n\to\infty} \log \mathbb{P}(M_n \le u_n) = \lim_{n\to\infty} \left(-n\bar{F}(u_n)(1+o(1))\right)$$

$$= -\lim_{n\to\infty} n\bar{F}(u_n).$$

Olkoon nyt $\tau = \infty$. Oletetaan (2.13). Jos (2.14) ei päde, voidaan määrätä $\tau' < \infty$ ja sellainen osajono, että

$$\lim_{k\to\infty} \mathbb{P}(M_{n_k} \le u_{n_k}) = e^{-\tau'}.$$

Tällöin alkuosan nojalla

$$\lim_{k\to\infty} n_k \bar{F}(u_{n_k}) = \tau',$$

mikä on ristiriita. Samoin nähdään, että (2.14):stä seuraa (2.13). $\square$

Maksimin lisäksi kiinnostavaa on tarkastella esimerkiksi 2. suurimman X:in jakaumaa. Merkitään

$$M_n^{(k)} = k. \text{ suurin havainnoista } X_1, \ldots, X_n.$$

Siis

$$M_n^{(k)} > x \text{ , jos ja vain jos}$$

$$X_{j_1} > x, \ldots, X_{j_k} > x \text{ jollain}$$

$$\{j_1, \ldots, j_k\} \subseteq \{1, \ldots, n\}, \quad j_1 < \ldots < j_k .$$

<u>Lause 2.3.1.</u> Olkoon (u_n) reaalilukujono, jolle

(2.14.2) $$\lim_{n\to\infty} n \bar{F}(u_n) = \tau \in (0, \infty).$$

Silloin

$$\lim_{n\to\infty} \mathbb{P}(M_n^{(k)} \le u_n) = e^{-\tau} + e^{-\tau}\tau + \ldots + e^{-\tau}\frac{\tau^{k-1}}{(k-1)!}$$

Rajan u_n ylittävien havaintojen lukumäärä on siis asymptoottisesti Poisson-jakautunut parametrilla τ.

<u>Todistus.</u> Olkoon

$$Z_n = \sum_{j=1}^{n} 1(X_j > u_n).$$

Tällöin Z_n on $\mathrm{Bin}(n, \bar{F}(u_n))$-jakautunut. Tunnetusti oletuksesta (2.14.2) seuraa, että

$$\lim_{n\to\infty} \mathbb{P}(Z_n = k) = e^{-\tau} \frac{\tau^k}{k!}.$$

Ilmeisesti

$$\{Z_n < k\} = \{M_n^{(k)} \le u_n\},$$

josta väite seuraa. $\square$

Olkoon Y:n kertymäfunktio G. Jos G on ei-degeneroitunut ja on olemassa $a_n \in (0,\infty)$, $b_n \in \mathbb{R}$, $n=1,2,\dots$, siten että

$$G^n(a_n x + b_n) = G(x), \quad \forall x \in \mathbb{R}, \ \forall n,$$

sanotaan, että G on stabiili maksimin suhteen (engl. max-stable).

Jos siis $Y, Y_1, Y_2, \dots$ ovat riippumattomia G-jakautuneita satunnaismuuttujia,

$$M_n' = \max(Y_1, \dots, Y_n), \quad n \in \mathbb{N},$$

ja G on stabiili maksimin suhteen, niin

$$\mathbb{P}\left(\frac{M_n' - b_n}{a_n} \leq x \right) = \mathbb{P}(Y \leq x) = G(x)$$

kaikilla $x \in \mathbb{R}$, $n \in \mathbb{N}$. Tällöin triviaalisti

$$\lim_{n \to \infty} \mathbb{P}\left(\frac{M_n' - b_n}{a_n} \leq x \right) = G(x), \quad \forall x \in \mathbb{R}.$$

Siis G esiintyy rajajakaumana kaavan (2.3) mukaisessa lausumassa. Todistetaan seuraavassa tämän käänteinen tulos.

Lause 2.4. Oletetaan, että G on ei-degeneroitunut kertymäfunktio ja että on olemassa $a_n \in (0,\infty)$ ja $b_n \in \mathbb{R}$ siten, että

$$\frac{M_n - b_n}{a_n} \xrightarrow{d} G.$$

Silloin G on stabiili maksimin suhteen.

Todistus. Olkoon x G:n jatkuvuuspiste. Silloin mielivaltaiselle $k \in \mathbb{N}$,

$$\lim_{n\to\infty} F(a_n x + b_n)^{nk} = \lim_{n\to\infty} \left(F(a_n x + b_n)^n \right)^k$$

$$= \lim_{n\to\infty} \left[\mathbb{P}\left(\frac{M_n - b_n}{a_n} \le x \right) \right]^k = G(x)^k.$$

Toisaalta

$$\lim_{n\to\infty} F(a_{nk} x + b_{nk})^{nk} = G(x).$$

Lauseen 2.1 nojalla on olemassa $c_k \in (0,\infty)$ ja $d_k \in \mathbb{R}$ siten, että

$$G(c_k x + d_k)^k = G(x), \quad \forall x \in \mathbb{R}.$$

Jos siis $Y, Y_1, Y_2, \ldots$ ovat riippumattomia G-jakautuneita satunnaismuuttujia ja

$$M_n' = \max(Y_1, \ldots, Y_n),$$

niin

$$\mathbb{P}\left(\frac{M_k' - d_k}{c_k} \le x \right) = G(c_k x + d_k)^k = \mathbb{P}(Y \le x). \quad \square$$

Mahdollisten rajajakaumien selvittäminen (2.3):ssa on siis ekvivalenttia maksimin suhteen stabiilien jakaumien karakterisoinnin kanssa. Tarkastellaan seuraavassa tätä kysymystä.

Lemma 2.5. Olkoon G maksimin suhteen stabiili ei-degeneroitunut jakauma. Silloin on olemassa funktiot $a : \mathbb{R} \to (0, \infty)$ ja $b : \mathbb{R} \to \mathbb{R}$ siten, että

$$(2.15) \qquad G(x)^t = G(a(t)x + b(t)), \quad \forall x \in \mathbb{R}, \ t > 0.$$

Lisäksi a ja b ovat jatkuvia ja toteuttavat funktionaaliyhtälöt

$$(2.16) \qquad a(t)a(u) = a(tu)$$

ja

$$(2.17) \qquad a(t)b(u) + b(t) = b(tu)$$

kaikilla $t, u > 0$.

<u>Todistus.</u> Olkoon $t>0$ kiinteä ja $x \in \mathbb{R}$. Tällöin

(2.18) $$\lim_{n\to\infty} G^{[tn]}(a_n x + b_n)$$

$$= \lim_{n\to\infty} \left[G^n(a_n x + b_n)\right]^{[tn]/n} = G(x)^t.$$

Toisaalta

(2.19) $$G^{[tn]}(a_{[tn]}x + b_{[tn]}) = G(x), \quad \forall n \geq \frac{1}{t}.$$

Lauseen 2.1 nojalla

(2.20) $$G(x)^t = G(a(t)x + b(t)), \quad \forall x \in \mathbb{R},$$

eräälle $a(t) > 0$ ja $b(t) \in \mathbb{R}$. Saadaan siis (2.15).

Olkoot nyt $t, u > 0$ kiinteitä. Tällöin

$$G(x)^{tu} = \left(G(x)^t\right)^u$$

$$= G(a(t)x + b(t))^u$$

$$= G(a(u)(a(t)x + b(t)) + b(u))$$

$$= G(a(t)a(u)x + a(u)b(t) + b(u)).$$

Toisaalta soveltamalla yhtälöä (2.15) parametrilla tu saadaan

(2.21) $$G(x)^{tu} = G(a(tu)x + b(tu)).$$

Lemman 2.1 nojalla

$$a(t)a(u) = a(tu)$$

ja

$$a(u)b(t) + b(u) = b(tu).$$

Ensimmäinen yhtälö on (2.16) ja (2.17) saadaan jälkimmäisestä vaihtamalla t:n ja u:n roolit.

On vielä todistettava a:n ja b:n jatkuvuus. Olkoon $t > 0$ kiinteä ja $t_n \to t$, kun $n \to \infty$. Tällöin

$$\lim_{n\to\infty} G(x)^{t_n} = G(x)^t, \quad \forall x \in \mathbb{R}.$$

Alkuosan nojalla

$$\lim_{n\to\infty} G(a(t_n)x + b(t_n)) = G(a(t)x + b(t)), \ \forall x \in \mathbb{R}.$$

Sama raja-arvo saadaan triviaalisti korvaamalla $a(t_n)$ ja $b(t_n)$ luvuilla $a(t)$ ja $b(t)$. Soveltamalla Lausetta 2.1 kertymäfunktioihin $F_n = G$, $\forall n$, saadaan

$$\lim_{n\to\infty} \frac{a(t_n)}{a(t)} = a', \quad \lim_{n\to\infty} \frac{b(t_n) - b(t)}{a(t)} = b', \quad a' > 0,\ b' \in \mathbb{R},$$

ja lisäksi

$$G(a(t)x + b(t)) = G(a(t)(a'x + b') + b(t))$$

$$= G(a'a(t)x + b'a(t) + b(t)).$$

Välttämättä $a' = 1$, joten a on jatkuva pisteessä t. Samoin $b(t) = b'a(t) + b(t)$, joten $b' = 0$. Siis $b(t_n) \to b(t)$, kun $n \to \infty$, joten b on jatkuva. $\square$

Maksimin suhteen stabiilit jakaumat saadaan nyt yhtälöiden (2.16) ja (2.17) ratkaisuista. Osoittautuu, että kysymykseen tulevat jakaumat ovat jotain seuraavaa tyyppiä.

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

II. Frechet: $G(x) = \begin{cases} 0, & x \leq 0 \\ e^{-x^{-\alpha}}, & x > 0 \end{cases}$

missä $\alpha > 0$ on vakio.

III. Weibull: $G(x) = \begin{cases} e^{-(-x)^{\alpha}}, & x \leq 0 \\ 1, & x > 0 \end{cases}$

missä $\alpha > 0$ on vakio.

Kaikki mainitut jakaumat ovat jatkuvia.

Huomautus. Weibull-jakaumaksi kutsutaan usein myös tyypin III analogiaa, joka keskittyy alueeseen $[0, \infty)$.

Lause 2.6. (Extremal Types Theorem). Ei-degeneroitunut kertymäfunktio on stabiili maksimin suhteen jos ja vain jos se on tyyppiä I, II tai III.

Todistus. Todistetaan, että luokan II jakaumat ovat stabiileja maksimin suhteen. Todistus luokille I ja III on analoginen.

On siis etsittävä a_n ja b_n siten, että

$$(2.21) \quad G(a_n x + b_n)^n = G(x), \quad \forall x \in \mathbb{R}.$$

Jos $a_n x + b_n > 0$, niin

$$G(a_n x + b_n)^n = \left[e^{-(a_n x + b_n)^{-\alpha}} \right]^n$$

$$= e^{-n(a_n x + b_n)^{-\alpha}} = e^{-\left(\frac{a_n x + b_n}{n^{1/\alpha}}\right)^{-\alpha}}$$

$$= G(x), \text{ jos } a_n = n^{1/\alpha} \text{ ja } b_n = 0.$$

Näillä valinnoilla (2.21) toteutuu, $\forall x \in \mathbb{R}$.

Olkoon nyt G mielivaltainen ei-degeneroitunut maksimin suhteen stabiili jakauma. Tällöin (2.15), (2.16) ja (2.17) toteutuvat. Erityisesti $a(1) = 1$ yhtälön (2.16) nojalla. Olkoon $a(e) = \xi$. Selvästi

$$(2.22) \qquad a(t^n) = a(t)^n, \quad \forall n \in \mathbb{N}, \; t > 0.$$

Lisäksi

$$a(t)\, a\left(\tfrac{1}{t}\right) = 1, \quad \forall\, t > 0,$$

joten

$$(2.23) \qquad a(t^{-n}) = \frac{1}{a(t)^n} = a(t)^{-n}.$$

Edelleen mielivaltaiselle rationaaliluvulle $\frac{m}{n} > 0$, $m, n > 0$,

$$a\left(t^{\frac{m}{n}}\right) = a\left(\left(t^{\frac{1}{n}}\right)^m\right) \overset{(2.22)}{=} a\left(t^{\frac{1}{n}}\right)^m \overset{(2.22)}{=} \left(a(t)^{\frac{1}{n}}\right)^m = a(t)^{\frac{m}{n}}.$$

Lisäksi (2.23):n nojalla

$$a\left(t^{-\frac{m}{n}}\right) = a\left(\left(t^{\frac{1}{n}}\right)^{-m}\right) = a\left(t^{\frac{1}{n}}\right)^{-m} = a(t)^{-\frac{m}{n}}.$$

Nähdään, että

$$a(e^r) = \xi^r, \quad \forall\, r \in \mathbb{Q}.$$

Koska a on jatkuva, on

$$(2.24) \qquad a(t) = a\left(e^{\log t}\right) = \xi^{\log t} = t^{\alpha}, \quad \forall\, t > 0,$$

missä $\alpha = \log \xi$.

Sijoittamalla saadut a yhtälöön (2.17) saadaan

$$(2.25) \quad t^{\alpha} b(u) + b(t) = b(tu).$$

Jos $\alpha = 0$, saadaan

$$b(u) + b(t) = b(tu), \quad \forall t, u > 0.$$

Merkitsemällä $c(t) = e^{b(t)}$ saadaan

$$c(u)\, c(t) = c(tu).$$

Tämä on sama kuin (2.16), joten

$$(2.26) \quad c(t) = t^{\beta}, \quad \forall t > 0,$$

eräille $\beta \in \mathbb{R}$ ja siis

$$(2.22) \quad b(t) = \beta \log t, \quad \forall t > 0.$$

Jos $\alpha \neq 0$, niin (2.25):n nojalla

$$t^{\alpha} b(u) + b(t) = b(tu)$$

$$= b(ut) = u^{\alpha} b(t) + b(u).$$

Siis

$$(t^{\alpha} - 1)\, b(u) = (u^{\alpha} - 1)\, b(t),$$

joten

$$\frac{b(t)}{t^{\alpha} - 1} = \gamma = \text{vakio}, \quad \forall t \neq 1.$$

Siis

$$(2.28)\qquad b(t) = \varkappa(t^{\alpha}-1), \quad \forall t>0.$$

Tarkastellaan saatuja ratkaisuja lähemmin.

a) $a(t)\equiv 1$, $\quad b(t)=\beta\log t \quad$ ((2.24), kun $\alpha=0$ ja $b(t)=\beta\log t$).

Tällöin

$$(2.28.1)\qquad G(x)^{t} = G(x+\beta\log t), \quad \forall t>0, \; x\in\mathbb{R}.$$

Olkoon $c\in\mathbb{R}$ sellainen, että $G(c)\in(0,1)$. Valitsemalla $x=c$ saadaan (2.28.1):stä

$$\log t + \log(-\log G(c)) = \log(-\log G(c+\beta\log t)), \quad \forall t>0.$$

Jos siis $h(y) = \log(-\log G(y))$, niin

$$\log t + h(c) = h(c+\beta\log t), \quad \forall t>0.$$

Nähdään, että $\beta\neq 0$ ja

$$h(c+\beta y) = y + h(c), \quad \forall y\in\mathbb{R},$$

joten

$$h(y) = h\left(c+\beta\left(\frac{y}{\beta}-\frac{c}{\beta}\right)\right) = \frac{y-c}{\beta} + h(c), \quad \forall y\in\mathbb{R}.$$

Siispä

$$G(x) = e^{-e^{h(x)}} = e^{-e^{\frac{x}{\beta} + h(c) - \frac{c}{\beta}}}, \quad \forall x \in \mathbb{R}.$$

Jos $\beta > 0$, niin $G(x) \to 0$, kun $x \to \infty$, joten G ei ole kertymäfunktio. Näin ollen G on tyyppiä I.

b) $a(t) = t^{\alpha}$, $\alpha < 0$, $b(t) = \mu(t^{\alpha} - 1)$.

Tällöin

$$G(x)^t = G(t^{\alpha}(x + \mu) - \mu), \quad \forall x \in \mathbb{R}, \ t > 0.$$

Jos $x \leq -\mu$, on vasen puoli vähenevä ja oikea puoli kasvava t:n funktio. Molemmat ovat siis vakioita alueessa $x \leq -\mu$. Koska G on kertymäfunktio, niin $G(x) = 0$, $\forall x \leq -\mu$.

Olkoon $c > -\mu$ sellainen, että $G(c) \in (0,1)$. Tällöin

$$G\left(\frac{c+\mu}{t^{\alpha}} - \mu\right) = G(c)^{\frac{1}{t}}, \quad \forall t > 0,$$

joten

$$G((c+\mu)t - \mu) = e^{t^{\alpha} \log G(c)}$$

ja

$$G\left(\frac{(c+\mu)t}{(-\log G(c))^{1/\alpha}} - \mu\right) = e^{-t^{\alpha}}, \quad \forall t > 0.$$

Nähdään, että G on tyyppiä II.

c) $a(t) = t^\alpha$, $\alpha > 0$, $b(t) = \gamma(t^\alpha - 1)$.

Tällöin päädytään tyyppiin III kuten kohdassa b). $\square$

<u>Huomautus 2.1.</u> Lause 2.4 sanoo sinänsä vain, että mahdollinen rajajakauma on samaa <u>tyyppiä</u> kuin I, II tai III. Rajajakauma on siis muotoa

$$G(ax+b), \quad a>0, \quad b \in \mathbb{R},$$

missä G on I:n, II:n tai III:n mukainen. Valitsemalla jonot (a_n) ja (b_n) sopivasti, saadaan kuitenkin täsmälleen I:n, II:n ja III:n mukaisia rajajakaumia (seuraus 2.2).

Luonnollinen kysymys nyt on, mitkä jakaumat kuuluvat tyyppien I, II ja III vaikutuspiireihin maksimin suhteen. Vastaus kysymykseen tunnetaan täydellisesti. Tulos esitetään seuraavassa lauseessa. Toinen luonnollinen kysymys koskee jonojen (a_n) ja (b_n) valintaa. Tätä koskeva tulos esitetään myös. Todistusten osalta rajoitutaan tyyppiin II.

Annetulle kertymäfunktiolle F merkitään

$$x_F = \sup\{x \mid F(x) < 1\} \in (-\infty, \infty].$$

Lause 2.7. Välttämättömät ja riittävät ehdot sille, että F kuuluu tyyppien I, II ja III vaikutuspiiriin maksimin suhteen ovat seuraavat.

Tyyppi I. On olemassa sellainen aidosti positiivinen funktio $g: \mathbb{R} \to \mathbb{R}$, että

$$\lim_{t \to x_F -} \frac{\bar{F}(t + x g(t))}{\bar{F}(t)} = e^{-x}, \quad \forall x \in \mathbb{R}.$$

Tyyppi II. $x_F = \infty$ ja

$$\lim_{t \to \infty} \frac{\bar{F}(tx)}{\bar{F}(t)} = x^{-\alpha}, \quad \forall x > 0, \text{ missä } \alpha > 0 \text{ on vakio.}$$

Tyyppi III. $x_F < \infty$ ja

$$\lim_{t \to \infty} \frac{\bar{F}(x_F - \frac{1}{tx})}{\bar{F}(x_F - \frac{1}{t})} = x^{-\alpha}, \quad \forall x > 0, \text{ missä } \alpha > 0 \text{ on vakio.}$$

Tyyppien II ja III kriteerit voidaan nähdä vaatimuksina säännöllisesti vaihtelevista hännistä. Määritelmän mukaan funktio $f : (0,\infty) \to (0,\infty)$ on <u>säännöllisesti vaihteleva indeksillä</u> ρ, jos

$$\lim_{t\to\infty} \frac{f(tx)}{f(t)} = x^{\rho}, \quad \forall x > 0.$$

Indeksiä ρ vastaavien säännöllisesti vaihtelevien funktioiden joukkoa merkitään symbolilla R_ρ. Jos $\rho = 0$, kutsutaan f:ää <u>hitaasti vaihtelevaksi</u>.

Tyyppiä II koskeva kriteeri on yhtäpitävä vaatimuksen

$$\bar{F} \in R_{-\alpha} \quad \text{jollain} \quad \alpha > 0$$

kanssa ja tyyppiä III vaatimuksen

$$f \in R_{-\alpha} \quad \text{jollain} \quad \alpha > 0$$

kanssa, missä

$$f(x) = \bar{F}(x_F - \tfrac{1}{x}), \quad \forall x > 0.$$

Merkitään

$$\gamma_n = \gamma_n(F) = \inf\{x \mid \bar{F}(x) \le \tfrac{1}{n}\}.$$

Lause 2.8. Lauseessa 2.7 jonot (a_n) ja (b_n) voidaan valita seuraavasti.

Tyyppi I. $a_n = g(\gamma_n)$, $b_n = \gamma_n$.

Tyyppi II. $a_n = \gamma_n$, $b_n = 0$.

Tyyppi III. $a_n = x_F - \gamma_n$, $b_n = x_F$.

Lauseiden 2.7 ja 2.8 todistus (koskien tyyppiä II).
Oletetaan, että $\bar{F} \in R_{-\alpha}$ jollain $\alpha > 0$ ja olkoon $a_n = \gamma_n$ ja $b_n = 0$, $\forall n \in \mathbb{N}$. Sovelletaan Lausetta 2.3.
Selvästi

$$\bar{F}(a_n) \leq \frac{1}{n}.$$

Toisaalta $a_n \to \infty$ kun $n \to \infty$ ja annetulle $\varepsilon > 0$

$$\bar{F}((1-\varepsilon)a_n) > \frac{1}{n}.$$

Koska $\bar{F} \in R_{-\alpha}$, niin

$$\lim_{n\to\infty} \frac{\bar{F}((1-\varepsilon)a_n)}{\bar{F}(a_n)} = (1-\varepsilon)^{-\alpha}.$$

Siis pä

$$\bar{F}(a_n) \geq \frac{(1-\varepsilon)^{\alpha}\ \bar{F}((1-\varepsilon)a_n)}{1+\varepsilon} > \frac{(1-\varepsilon)^{\alpha}}{(1+\varepsilon)n},$$

kun n on riittävän suuri. Nähdään, että

$$\lim_{n\to\infty} n\bar{F}(a_n) = 1.$$

Mielivaltaiselle $x > 0$ pätee

$$\lim_{n\to\infty} n\bar{F}(a_n x) = \lim_{n\to\infty} n\bar{F}(a_n) \cdot \frac{\bar{F}(a_n x)}{\bar{F}(a_n)} = x^{-\alpha}.$$

Lauseen 2.3 nojalla

$$(2.29) \quad \lim_{n\to\infty} \mathbb{P}\left(\frac{M_n}{a_n} \leq x\right) = e^{-x^{-\alpha}}.$$

Tämä on tyypin II mukainen raja-arvo.

Koska $e^{-x^{-\alpha}} \to 0$, kun $x \to 0+$, on välttämättä

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{M_n}{a_n} \le x\right) = 0, \quad \forall x \le 0.$$

Siis F on tyyppiä II.

Olkoon nyt F tyyppiä II. On osoitettava, että $\overline{F} \in R_{-\alpha}$. Koska

$$\lim_{n\to\infty} F^n(a_n x + b_n) = e^{-x^{-\alpha}}, \quad \forall x \ge 0,$$

on

$$\lim_{n\to\infty} F^n(a_{[tn]}x + b_{[tn]}) = \lim_{n\to\infty} \left[F^{[tn]}(a_{[tn]}x + b_{[tn]})\right]^{\frac{n}{[nt]}}$$

$$= \left(e^{-x^{-\alpha}}\right)^{t^{-1}} = e^{-\left(t^{\frac{1}{\alpha}}x\right)^{-\alpha}}, \quad \forall x \ge 0,\ t > 0.$$

Lauseen 2.1 nojalla

(2.30) $$\lim_{n\to\infty} \frac{a_{[tn]}}{a_n} = t^{\frac{1}{\alpha}}, \quad \forall t > 0,$$

ja

(2.31) $$\lim_{n\to\infty} \frac{b_{[tn]} - b_n}{a_n} = 0, \quad \forall t > 0.$$

Osoitetaan seuraavaksi, että

(2.32) $$\lim_{n\to\infty} \frac{b_n}{a_n} = 0.$$

Olkoon $\varepsilon > 0$ annettu. Kiinnitetään suuri N ja $C \in (0,1)$ siten, että

$$(2.33) \qquad \frac{|b_n - b_{[\frac{1}{2}n]}|}{a_{[\frac{1}{2}n]}} \le \varepsilon, \qquad \forall n \ge N$$

ja

$$(2.34) \qquad \frac{a_{[\frac{1}{2}n]}}{a_n} \le C, \qquad \forall n \ge N.$$

Merkitään

$$(2.35) \qquad A = \sup\{|b_n| \mid N \le n \le 2N+2\}.$$

Olkoon $n \ge N$. Määritellään $n_0 = n$ ja

$$n_k = [\tfrac{1}{2} n_{k-1}], \quad k = 1, 2, \dots$$

Olkoon edelleen $m \in \mathbb{N} \cup \{0\}$ sellainen, että

$$n_m \ge N \quad \text{ja} \quad n_{m+1} < N.$$

Tällöin

$$\frac{|b_n|}{a_n} \le \frac{1}{a_n}\left[\sum_{k=0}^{m-1} |b_{n_k} - b_{n_{k+1}}| + |b_{n_m}|\right]$$

$$= \sum_{k=0}^{m-1} \frac{|b_{n_k} - b_{n_{k+1}}|}{a_{n_{k+1}}} \cdot \frac{a_{n_{k+1}}}{a_{n_k}} \cdots \frac{a_{n_1}}{a_{n_0}} + \frac{|b_{n_m}|}{a_n}.$$

Koska $n_k \ge N$, $k = 0, 1, \dots, m$ ja toisaalta

$$N > n_{m+1} = [\tfrac{1}{2} n_m] \ge \tfrac{1}{2} n_m - 1$$

eli $n_m \le 2N + 2$, voidaan soveltaa epäyhtälöitä (2.33) ja (2.34) ja $|b_{n_m}| \le A$. Saadaan

$$\frac{|b_n|}{a_n} \le \sum_{k=0}^{m-1} \varepsilon \cdot C^{k+1} + \frac{A}{a_n}$$

$$\le \varepsilon C \cdot \frac{1}{1-C} + \frac{A}{a_n}, \quad \forall n \ge N.$$

Tuloksen (2.30) nojalla $a_n \to \infty$. Nimittäin, $\forall x > 0$,

$$\lim_{t\to\infty} \frac{a_{[tx]}}{a_{[t]}} = x^{\frac{1}{2}}.$$

Siis $x \mapsto a_{[x]}$ on säännöllisesti vaihteleva indeksillä $\frac{1}{2} > 0$. Tällöin $a_n \to \infty$ (kts. myöhemmin esitettävä lemma 3.1). Siis $\frac{|b_n|}{a_n} \to 0$.

Lauseen 2.1 nojalla

$$\lim_{n\to\infty} F^n(a_n x) = e^{-x^{-\alpha}}, \quad \forall x \geq 0.$$

Täysin samoin

$$\lim_{n\to\infty} \mathbb{P}(M_n \leq a_n x) = e^{-x^{-\alpha}}.$$

Lauseen 2.3 nojalla

$$(2.36) \quad \lim_{n\to\infty} n\bar{F}(a_n x) = x^{-\alpha}, \quad \forall x \geq 0.$$

Olkoon $t > 0$ ja

$$n_t = \inf\{n \mid a_{n+1} > t\}.$$

Koska $a_n \to \infty$, on n_t hyvin määritelty. Selvästi $n_t \to \infty$, kun $t \to \infty$. Koska $\bar{F}$ on vähenevä, on

$$\frac{\bar{F}(a_{n_t} x)}{\bar{F}(a_{n_t+1})} \geq \frac{\bar{F}(tx)}{\bar{F}(t)} \geq \frac{\bar{F}(a_{n_t+1} x)}{\bar{F}(a_{n_t})}, \quad \forall x \geq 0.$$

Tuloksen (2.36) nojalla äärimmäiset termit suppenevat kohti $x^{-\alpha}$:aa, kun $t \to \infty$, joten

$$\lim_{t\to\infty} \frac{\bar{F}(tx)}{\bar{F}(t)} = x^{-\alpha}.$$

Siis $\bar{F} \in R_{-\alpha}$. $\square$