

SVEUČILIŠTE U ZAGREBU
PRIRODOSLOVNO–MATEMATIČKI FAKULTET
MATEMATIČKI ODSJEK

Ivana Antoliš

**RAZDIOBA DUGIH CIKLUSA I VELIKIH
DJELITELJA**

Diplomski rad

Voditelj rada:
prof. dr. sc. Bojan Basrak

Zagreb, rujan, 2015.

Ovaj diplomski rad obranjen je dana _____ pred ispitnim povjerenstvom u sastavu:

1. _____, predsjednik
2. _____, član
3. _____, član

Povjerenstvo je rad ocijenilo ocjenom _____.

Potpisi članova povjerenstva:

1. _____
2. _____
3. _____

Sadržaj

Sadržaj	iv
Uvod	1
1 Pripremni rezultati	2
1.1 Funkcija Γ i njezine primjene	2
1.2 Neki pojmovi teorije vjerojatnosti	5
1.3 Neki pojmovi teorije brojeva	14
1.4 Formula uključivanja i isključivanja	16
2 Motivacijski primjeri	18
2.1 Ciklusi slučajnih permutacija	18
2.2 Djelitelji slučajnih brojeva	25
2.3 Prostor Δ	33
3 Poisson-Dirichletova razdioba	37
3.1 Povezanost Dirichletove i Poisson-Dirichletove razdiobe	37
3.2 Tehnički rezultati	43
3.3 Svojstva Poisson-Dirichletove razdiobe	55
4 Zaključak	64
Bibliografija	68

Uvod

Centralni objekt ovog rada je Poisson-Dirichletova razdioba koju je prvi uveo Kingman 1975. godine. Pokazuje se da su njezine primjene mnogobrojne u područjima kao što su populacijska genetika, kombinatorika, teorija brojeva, ekonomija, statistika i druge. U ovom radu uvest ćemo je kroz dva različita modela koja je proučavao Billingsley u svojoj knjizi *Convergence of Probability Measures* iz 1968. i kasnijem izdanju iz 1999. godine.

Prvi model bazira se na proučavanju nasumičnih permutacija fiksne duljine, koje dolaze iz uniformne distribucije te na njihovom rastavu na cikluse. Zanimat će nas asimptotska svojstva duljina ciklusa normiranih s duljinom permutacije koja će se posebno važnima pokazati u slučaju nekoliko najduljih ciklusa.

Drugi model koji ćemo proučavati dolazi iz teorije brojeva, konkretno promatrat ćemo rastav nasumično odabranog broja iz skupa $\{1, 2, \dots, n\}$ na proste faktore. Pokazat ćemo da su asimptotska svojstva skaliranih vrijednosti najvećih prostih faktora, nakon logaritmiranja, ista onima u slučaju dugih ciklusa.

U poglavlju 1 definirat ćemo pojmove i dokazati alate nužne za razumijevanje nastavka rada. Budući da se u radu obrađuju rezultati iz više matematičkih područja u ovom se poglavlju povezuju pojmovi iz teorije brojeva, analize, teorije vjerojatnosti i teorije mjere.

U poglavlju 2 detaljno obrazložimo i dokazujemo sve rezultate potrebne kako bi se objasnilo asimptotsko ponašanje slučajnih vektora iz dva spomenuta modela. Uvest ćemo prostor na kojem su promatrani objekti definirani i dokazati da oba modela konvergiraju prema istoj distribuciji.

U poglavlju 3 pomoću ranije pokazanih rezultata definirat ćemo centralni objekt ovog rada, Poisson-Dirichletovu razdiobu. Pokazat ćemo kako je povezana s Dirichletovom razdiobom te pomoću nje izvesti formulu za gustoću, marginalne gustoće i momente Poisson-Dirichletove razdiobe.

U poglavlju 4 vratit ćemo se originalnim motivacijskim primjerima te argumentirati i ilustrirati kako se dobiveni rezultati na njih odražavaju.

Svi rezultati bit će potkrijepljeni detaljnim objašnjenjima i primjerima. Za razumijevanje dokaza teorema u radu potrebno je predznanje iz područja teorije mjere i teorije vjerojatnosti.

Poglavlje 1

Pripremni rezultati

U ovom poglavlju definirat ćemo pojmove i navesti, a jednim dijelom i dokazati, njihova svojstva koja će se pokazati važnima u nastavku rada.

1.1 Funkcija Γ i njezine primjene

Definicija 1.1. Funkciju $\Gamma : \langle 0, \infty \rangle \rightarrow \mathbb{R}$ definiramo pomoću nepravog Riemannovog integrala

$$\Gamma(t) := \int_0^{\infty} x^{t-1} e^{-x} dx. \quad (1.1)$$

Pokazuje se da je funkcija Γ dobro definirana te da joj se domena može i proširiti. Za okvire ovog rada ovakva definicija će biti dostatna.

Propozicija 1.2. Funkcija Γ zadovoljava sljedeće jednakosti

(i) $\Gamma(t+1) = t\Gamma(t)$, za $t > 0$

(ii) $\Gamma(n+1) = n!$, za $n \in \mathbb{N}$

(iii) $\frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)} = \int_0^1 u^{x-1} (1-u)^{y-1} du$, za $x, y > 0$.

Dokaz. (i)

$$\Gamma(t+1) = \int_0^{\infty} x^t e^{-x} dx = \left(-e^{-x} x^t\right)\Big|_0^{\infty} + \int_0^{\infty} t x^{t-1} e^{-x} dx = t\Gamma(t)$$

(ii) Budući da je

$$\Gamma(1) = \int_0^{\infty} x^0 e^{-x} dx = (-e^{-x}) \Big|_0^{\infty} = 1 \quad (1.2)$$

koristeći (i) rekurzivno dobivamo $\Gamma(n+1) = n\Gamma(n) = \dots = n!\Gamma(1) = n!$

(iii)

$$\begin{aligned} \Gamma(x)\Gamma(y) &= \int_0^{\infty} t^{x-1} e^{-t} dt \cdot \int_0^{\infty} s^{y-1} e^{-s} ds = \\ &= \iint_{t,s>0} t^{x-1} e^{-t} s^{y-1} e^{-s} ds dt = \left[f(u, v) = (t, s) = (uv, v(1-u)), \left[\begin{array}{cc} v & u \\ -v & 1-u \end{array} \right] = v \right] \\ &= \iint_{\substack{0 < u < 1 \\ 0 < v}} u^{x-1} v^{x-1} e^{-uv} v^{y-1} (1-u)^{y-1} e^{-v(1-u)} v du dv = \\ &= \int_0^1 \int_0^{\infty} u^{x-1} (1-u)^{y-1} v^{x+y-1} e^{-v} dv du = \int_0^1 u^{x-1} (1-u)^{y-1} du \int_0^{\infty} v^{x+y-1} e^{-v} dv = \\ &= \int_0^1 u^{x-1} (1-u)^{y-1} du \cdot \Gamma(x+y). \end{aligned}$$

Dijeljenjem izraza s $\Gamma(x+y)$ slijedi tvrdnja propozicije.

□

Označimo s $B_k(x) \subset \mathbb{R}^k$ definiran na sljedeći način

$$B_k(x) := \left\{ (t_1, \dots, t_k) \in \mathbb{R}^k : t_1, \dots, t_k > 0, \sum_{i=1}^k t_i < x \right\}.$$

Uz ovu definiciju vrijedi sljedeći teorem

Teorem 1.3. *Neka je dana funkcija $f : \mathbb{R} \rightarrow \mathbb{R}$ i $\alpha_1, \dots, \alpha_k > 0$, $k \in \mathbb{N}$, tada vrijedi*

$$\int_{B_k(x)} f(t_1 + \dots + t_k) t_1^{\alpha_1-1} \dots t_k^{\alpha_k-1} dt_1 \dots dt_k = \frac{\Gamma(\alpha_1) \dots \Gamma(\alpha_k)}{\Gamma(\alpha_1 + \dots + \alpha_k)} \int_0^x f(t) t^{\sum_{i=1}^k \alpha_i - 1} dt. \quad (1.3)$$

Dokaz. Tvrdnju pokazujemo matematičkom indukcijom. Za $k = 1$ tvrdnja trivijalno slijedi. Za $k = 2$ vrijedi

$$\int_{\substack{t_1, t_2 > 0 \\ t_1 + t_2 < x}} f(t_1 + t_2) t_1^{\alpha_1 - 1} t_2^{\alpha_2 - 1} dt_1 dt_2 = \left[f(s_1, s_2) = ((1 - s_1) s_2, s_1 s_2), \left[\begin{array}{c|c} -s_2 & 1 - s_1 \\ \hline s_2 & s_1 \end{array} \right] = -s_2 \right]$$

(Uočimo da uvjet $t_1 + t_2 < x$ povlači $s_2 = (1 - s_1) s_2 + s_1 s_2 < x$.)

$$\begin{aligned} & \int_{\substack{0 < s_1 < 1 \\ 0 < s_2 < x}} f(s_2) (1 - s_1)^{\alpha_1 - 1} s_2^{\alpha_1 - 1} s_1^{\alpha_2 - 1} s_2^{\alpha_2 - 1} |-s_2| ds_1 ds_2 = \\ & \int_0^1 (1 - s_1)^{\alpha_1 - 1} s_1^{\alpha_2 - 1} ds_1 \int_0^x f(s_2) s_2^{\alpha_1 + \alpha_2 - 1} ds_2 \stackrel{1.2}{=} \frac{\Gamma(\alpha_1) \Gamma(\alpha_2)}{\Gamma(\alpha_1 + \alpha_2)} \int_0^x f(t) t^{\alpha_1 + \alpha_2 - 1} dt. \end{aligned}$$

Pretpostavimo sada da tvrdnja vrijedi za svaki $k < n$ te pokažimo da vrijedi za n . Budući da su slučajevi $n = 1, 2$ već pokazani pretpostavimo da je $n > 2$. Označimo s $a = t_3 + \dots + t_n$ tada je

$$\begin{aligned} & \int_{B_n(x)} f(t_1 + t_2 + a) t_1^{\alpha_1 - 1} t_2^{\alpha_2 - 1} \dots t_n^{\alpha_n - 1} dt_1 dt_2 \dots dt_n \stackrel{n-2 > 0}{=} \\ & \int_{B_{n-2}(x)} t_3^{\alpha_3 - 1} \dots t_n^{\alpha_n - 1} \int_{\substack{t_1, t_2 > 0 \\ t_1 + t_2 < x - a}} f(t_1 + t_2 + a) t_1^{\alpha_1 - 1} t_2^{\alpha_2 - 1} dt_1 dt_2 dt_3 \dots dt_n = \end{aligned}$$

(Primjenimo li slučaj $k = 2$ uz $\tilde{x} = x - a$ i $\tilde{f}(t) = f(t + a)$.)

$$\frac{\Gamma(\alpha_1) \Gamma(\alpha_2)}{\Gamma(\alpha_1 + \alpha_2)} \int_{B_{n-2}(x)} t_3^{\alpha_3 - 1} \dots t_n^{\alpha_n - 1} \int_0^{x-a} f(t + a) t^{\alpha_1 + \alpha_2 - 1} dt dt_3 \dots dt_n =$$

$$\frac{\Gamma(\alpha_1) \Gamma(\alpha_2)}{\Gamma(\alpha_1 + \alpha_2)} \int_{B_{n-1}(x)} f(t + t_3 + \dots + t_n) t^{(\alpha_1 + \alpha_2) - 1} t_3^{\alpha_3 - 1} \dots t_n^{\alpha_n - 1} dt dt_3 \dots dt_n \stackrel{\text{po pretpostavci indukcije}}{=}$$

$$\frac{\Gamma(\alpha_1) \Gamma(\alpha_2)}{\Gamma(\alpha_1 + \alpha_2)} \cdot \frac{\Gamma(\alpha_1 + \alpha_2) \Gamma(\alpha_3) \dots \Gamma(\alpha_n)}{\Gamma((\alpha_1 + \alpha_2) + \alpha_3 + \dots + \alpha_n)} \int_0^x f(t) t^{(\alpha_1 + \alpha_2) + \sum_{i=3}^n \alpha_i - 1} dt =$$

$$\frac{\Gamma(\alpha_1) \dots \Gamma(\alpha_n)}{\Gamma(\alpha_1 + \dots + \alpha_n)} \int_0^x f(t) t^{\sum_{i=1}^n \alpha_i - 1} dt.$$

Po principu matematičke indukcije slijedi tvrdnja teorema za svaki $k \in \mathbb{N}$. □

U radu ćemo češće koristiti sljedeći korolar.

Korolar 1.4. Neka je dana funkcija $f : \mathbb{R} \rightarrow \mathbb{R}$, $\alpha > 0$ i $k \in \mathbb{N}$, tada vrijedi

$$\int_{B_k(x)} f(t_1 + \dots + t_k) (t_1 \dots t_k)^{\alpha k - 1} dt_1 \dots dt_k = \frac{\Gamma^k(\alpha)}{\Gamma(k\alpha)} \int_0^x f(t) t^{k\alpha - 1} dt.$$

Dokaz. Tvrdnja slijedi izravno iz (1.3) ako uzmemo $\alpha_1 = \dots = \alpha_k = \alpha > 0$. □

Može se pokazati da je funkcija Γ definirana pomoću (1.1) analitička, što će se kasnije pokazati važnim. Uz nju vežemo i pojam *Euler-Mascheronove konstante* γ . Numeričkim metodama može se odrediti njezina približna vrijednost, odnosno pokazati da je $\gamma \approx 0.5772$, a za nas će biti važno njezino svojstvo koje daje vazdu između funkcije Γ i konstante γ . Vrijedi

$$\gamma = -\Gamma'(1). \tag{1.4}$$

1.2 Neki pojmovi teorije vjerojatnosti

Definicija 1.5. Neka je $B : \langle 0, \infty \rangle \times \langle 0, \infty \rangle \rightarrow \mathbb{R}$, takva da vrijedi

$$B(x, y) = \int_0^1 u^{x-1} (1-u)^{y-1} du. \tag{1.5}$$

Za B kažemo da je *beta-funkcija*.

Definicija 1.6. Neka su $p > 0, q > 0$ fiksni. Neprekidna slučajna varijabla X ima *beta-distribuciju* s parametrima p i q ako joj je gustoća f dana sa

$$f(x) = \begin{cases} \frac{x^{p-1} (1-x)^{q-1}}{B(p, q)}, & 0 < x < 1, \\ 0, & x \leq 0 \text{ ili } x \geq 1. \end{cases}$$

Iz (1.5) se lako vidi da je f zaista vjerojatnosna funkcija gustoće, a ako uvažimo i formulu iz propozicije 1.2 slijedi alternativna formula

$$f(x) = \begin{cases} \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)} x^{p-1} (1-x)^{q-1}, & 0 < x < 1, \\ 0, & x \leq 0 \text{ ili } x \geq 1. \end{cases} \tag{1.6}$$

Promotrimo kako izgleda beta-distribucija u posebnom slučaju $p = q = 1$. Budući da je po propoziciji 1.2 $\Gamma(n + 1) = n!$ iz (1.6) slijedi

$$f(x) = \begin{cases} \frac{1 \cdot 1}{(2-1)!} x^{1-1} (1-x)^{1-1}, & 0 < x < 1, \\ 0, & x \leq 0 \text{ ili } x \geq 1 \end{cases} = \begin{cases} 1, & 0 < x < 1, \\ 0, & x \leq 0 \text{ ili } x \geq 1. \end{cases}$$

Iz ovoga zaključujemo da je beta-distribucija s parametrima $p = q = 1$ zapravo uniformna distribucija na $\langle 0, 1 \rangle$. Više informacija, odnosno svojstava beta-distribucije može se naći u Sarapa [4].

Okrenimo se sada rezultatima ključnima za daljnji nastavak rada koji govore o graničnim svojstvima slučajnih varijabli.

Definicija 1.7. Reći ćemo da niz slučajnih varijabli ili vektora $\{X_n\}$ konvergira po distribuciji prema slučajnoj varijabli, odnosno vektoru X i pišemo $X_n \Rightarrow X$ ako za svaku neprekidnu, ograničenu funkciju f vrijedi

$$\mathbb{E}[f(X_n)] \rightarrow \mathbb{E}[f(X)] \text{ kada } n \rightarrow \infty.$$

Detaljna svojstva i motivacija za ovakvu definiciju mogu se naći u Billingsley [1]. Sljedeći teorem daje nam nekoliko alternativnih definicija koje će se u kasnijim poglavljima pokazati korisnima.

Teorem 1.8. Za $X_n, X : \Omega \rightarrow \mathbb{R}^r$, $r \in \mathbb{N}$ sljedećih pet uvjeta su ekvivalentni

- (i) $X_n \Rightarrow X$.
- (ii) $\mathbb{E}[f(X_n)] \rightarrow \mathbb{E}[f(X)]$ kada $n \rightarrow \infty$ za sve ograničene, neprekidne funkcije f .
- (iii) $\limsup_n \mathbb{P}[X_n \in F] \leq \mathbb{P}[X \in F]$ za svaki zatvoreni skup F .
- (iv) $\liminf_n \mathbb{P}[X_n \in U] \geq \mathbb{P}[X \in U]$ za svaki otvoreni skup U .
- (v) $\mathbb{P}[X_n \in A] \rightarrow \mathbb{P}[X \in A]$ za svaki skup A takav da je $\mathbb{P}[X \in \partial A] = 0$.

Dokaz. (i) \Rightarrow (ii) Vrijedi izravno po definiciji.

(ii) \Rightarrow (iii) Neka je $F \subset \mathbb{R}^r$ zatvoren skup i neka je $\varepsilon > 0$ proizvoljan, tada je $f : \mathbb{R}^r \rightarrow \mathbb{R}$

$$f(x) = \left(1 - \frac{d(x, F)}{\varepsilon}\right)^+,$$

gdje d označava metriku na \mathbb{R}^r , a $x^+ = x \mathbb{1}_{x>0}$.

Budući da je $d(x, F) \geq 0$ za svaki x slijedi da je $0 \leq f \leq 1$. Budući da je f kompozicija neprekidnih funkcija i sama je neprekidna. Također ukoliko s označimo $F^\varepsilon := \{x \in \mathbb{R}^r : d(x, F) \geq \varepsilon\}$ vrijedi

$$\mathbb{1}_F(x) \leq f(x) \leq \mathbb{1}_{F^\varepsilon}(x). \quad (1.7)$$

Zaista, ukoliko je $x \in F$ tada je $d(x, F) = 0$ pa je $f(x) = 1 \geq \mathbb{1}_F(x)$, za $x \notin F$ je $\mathbb{1}_F(x) = 0$, a kako je $f \geq 0$ nejednakost trivijalno slijedi. Za $x \notin F^\varepsilon$ je $\frac{d(x, F)}{\varepsilon} > \frac{\varepsilon}{\varepsilon} = 1$ pa je $f(x) = 0 \leq \mathbb{1}_F^\varepsilon(x)$, a za $x \in F^\varepsilon$ je $\mathbb{1}_{F^\varepsilon}(x) = 1$, a kako je $f \leq 1$ nejednakost trivijalno slijedi.

Iz (1.7) slijedi

$$\begin{aligned} \limsup_n \mathbb{P}[X_n \in F] &= \limsup_n \mathbb{E}[\mathbb{1}_F(X_n)] \leq \limsup_n \mathbb{E}[f(X_n)] \stackrel{(ii)}{=} \\ &= \mathbb{E}[f(X)] \leq \mathbb{E}[\mathbb{1}_{F^\varepsilon}(X)] = \mathbb{P}[X \in F^\varepsilon]. \end{aligned}$$

Budući da je F zatvoren puštanjem $\varepsilon \searrow 0$ zbog neprekidnosti vjerojatnosti s obzirom na padajući niz događaja slijedi (iii).

(iii) \Rightarrow (iv) Budući da je U otvoren slijedi da je U^c zatvoren pa imamo

$$\liminf_n \mathbb{P}[X_n \in U] = 1 - \limsup_n \mathbb{P}[X_n \in U^c] \stackrel{(iii)}{\geq} 1 - \mathbb{P}[X \in U^c] = \mathbb{P}[X \in U].$$

(iii) i (iv) \Rightarrow (v) Označimo s $\text{Int}(A)$ interior skupa A i s $\text{Cl}(A)$ zatvarač skupa A . Tada vrijedi

$$\begin{aligned} \mathbb{P}[X \in A] &\leq \mathbb{P}[X \in \text{Int}(A)] + \mathbb{P}[X \in \partial A] = \mathbb{P}[X \in \text{Int}(A)] \stackrel{(iv)}{\leq} \liminf_n \mathbb{P}[X_n \in \text{Int}(A)] \leq \\ &\limsup_n \mathbb{P}[X_n \in \text{Cl}(A)] \stackrel{(iii)}{\leq} \mathbb{P}[X \in \text{Cl}(A)] \leq \mathbb{P}[X \in A] + \mathbb{P}[X \in \partial A] = \mathbb{P}[X \in A], \end{aligned}$$

iz čega slijedi (v).

(v) \Rightarrow (i) Neka je $f : \mathbb{R}^r \rightarrow \mathbb{R}$ ograničena i neprekidna funkcija, pa postoje konstante $a < b \in \mathbb{R}$ takve da je $a \leq f(x) \leq b$ za svaki $x \in \mathbb{R}^r$. Tada za funkciju

$$g(x) := \frac{f(x) - a}{b - a}$$

vrijedi $0 \leq g(x) \leq 1$ za svaki $x \in \mathbb{R}^r$. Zato je

$$\mathbb{E}[f(X)] = (b - a)\mathbb{E}[g(X)] + a = (b - a) \int_0^1 x dF_{g(X)} + a = (b - a) \int_0^1 \int_0^x dt dF_{g(X)} + a =$$

$$(b-a) \int_0^1 \int_t^1 dF_{g(X)} dt + a = (b-a) \int_0^1 \mathbb{P}[g(X) \leq 1] - \mathbb{P}[g(X) \leq t] dt + a =$$

$$(b-a) \int_0^1 \mathbb{P}[g(X) > t] dt + a = (b-a) \int_0^1 \mathbb{P}[X \in g^{-1}(\langle t, \infty \rangle)] dt + a,$$

gdje je $F_{g(X)}$ funkcija distribucije od $g(X)$. Analogan račun vrijedi i za X_n .

Budući da je $X \in \partial g^{-1}(\langle t, \infty \rangle) \subset X \in g^{-1}(t)$ slijedi da je $\mathbb{P}[X \in \partial g^{-1}(\langle t, \infty \rangle)] \leq \mathbb{P}[X \in g^{-1}(t)]$ pa je $\mathbb{P}[X \in \partial g^{-1}(\langle t, \infty \rangle)] = 0$ za sve osim eventualno prebrojivo mnogo $t \in \langle 0, 1 \rangle$.

Budući da je $\mathbb{P}[X \in g^{-1}(\langle t, \infty \rangle)] \leq 1$, a $\int_0^1 1 dt = 1 < \infty$ po Lebesgueovom teoremu o dominiranoj konvergenciji i tvrdnji (v) za $A = g^{-1}(\langle t, \infty \rangle)$ slijedi

$$\mathbb{E}[f(X_n)] = (b-a) \int_0^1 \mathbb{P}[X_n \in g^{-1}(\langle t, \infty \rangle)] dt + a \xrightarrow{(v)}$$

$$(b-a) \int_0^1 \mathbb{P}[X \in g^{-1}(\langle t, \infty \rangle)] dt + a = \mathbb{E}[f(X)], \text{ kada } n \rightarrow \infty.$$

Ovime je pokazano da je svih pet uvjeta međusobno ekvivalentno. \square

Sljedeća definicija zajedno s propozicijama koje je slijede a navodene su bez dokaza, koji se mogu naći u primjerima 2.3. i 2.4. u Billingsley[1], omogućuje nam da umjesto svih skupova A u (v) ponekad promatramo manju klasu.

Definicija 1.9. Reći ćemo da je klasa \mathcal{A} određuje konvergenciju ako za svaki X i svaki niz $\{X_n\}$ konvergencija $\mathbb{P}[X_n \in A] \rightarrow \mathbb{P}[X \in A]$ za svaki skup $A \in \mathcal{A}$ takav da je $\mathbb{P}[X \in \partial A] = 0$ povlači $X_n \Rightarrow X$.

Propozicija 1.10. Klasa $\mathcal{A} := \left\{ \prod_{i=1}^r \langle -\infty, x_i \rangle : x_i, i = 1, 2, \dots, r \right\}$ u \mathbb{R}^r određuje konvergenciju.

Propozicija 1.11. Klasa $\mathbb{R}_f^\infty := \left\{ \pi_k^{-1} H : k \geq 1, H \in \mathbb{R}^k \right\}$ u \mathbb{R}^∞ određuje konvergenciju. ($\pi_k : \mathbb{R}^\infty \rightarrow \mathbb{R}^k$ je prirodna projekcija)

Sada smo u mogućnosti dati ekvivalentnu definiciju konvergencije po distribuciji.

Korolar 1.12. Neka su X_n i X slučajne varijable i F_{X_n}, F_X njihove funkcije distribucije. Tada $X_n \Rightarrow X$ ako i samo ako $F_{X_n}(x) \rightarrow F_X(x)$ za svaki $x \in D_{F_X}$, gdje je D_{F_X} skup prekida funkcije F_X . Odnosno

$$X_n \Rightarrow X \text{ ako i samo ako } X_n \xrightarrow{\mathcal{D}} X.$$

Dokaz. Neka $X_n \Rightarrow X$, tada po teoremu 1.8 $\mathbb{P}[X_n \in A] \rightarrow \mathbb{P}[X \in A]$ za svaki skup A takav da je $\mathbb{P}[X \in \partial A] = 0$. Neka je $x \in D_{F_X}$, tada je

$$\mathbb{P}[X \in \partial \langle -\infty, x \rangle] = \mathbb{P}[X = x] = F_X(x) - F_X(x-) = 0.$$

Uz $A = \langle -\infty, x \rangle$ dobivamo

$$F_{X_n}(x) = \mathbb{P}[X_n \in \langle -\infty, x \rangle] \rightarrow \mathbb{P}[X \in \langle -\infty, x \rangle] = F_X(x),$$

odnosno $X_n \xrightarrow{\mathcal{D}} X$.

Pretpostavimo sada da $X_n \xrightarrow{\mathcal{D}} X$. Tada za $x \in D_{F_X}$, odnosno, po računu od ranije, za x takav da je $\mathbb{P}[X \in \partial \langle -\infty, x \rangle] = 0$ vrijedi

$$\mathbb{P}[X_n \in \langle -\infty, x \rangle] = F_{X_n}(x) \rightarrow F_X(x) = \mathbb{P}[X \in \langle -\infty, x \rangle].$$

Po propoziciji 1.10 za $r = 1$ slijedi $X_n \Rightarrow X$. □

Sljedeći teorem daje nam neke dovoljne uvjete za konvergenciju u slučaju kada promatrana klasa \mathcal{A} ne određuje nužno konvergenciju. Dokaz se može naći u teoremu 2.5 u Billingsley [1], a prije iskaza navedimo još jednu definiciju.

Definicija 1.13. Neka je X neprazan skup. Familija $\mathcal{S} \subset \mathcal{P}(X)$ zove se poluprsten (podskupova od X) ako:

$$(S1) \quad \emptyset \in \mathcal{S}.$$

$$(S2) \quad A, B \in \mathcal{S} \Rightarrow A \cap B \in \mathcal{S}.$$

$$(S3) \quad A, B \in \mathcal{S} \Rightarrow \exists n \in \mathbb{N}, \exists C_1, C_2, \dots, C_n \in \mathcal{S}, \text{ međusobno disjunktni, takvi da je } A \setminus B = C_1 \cup C_2 \cup \dots \cup C_n.$$

Teorem 1.14. Neka je dan poluprsten \mathcal{A} takav da je svaki otvoren skup unija prebrojivo mnogo skupova iz \mathcal{A} . Ako vrijedi $\mathbb{P}[X \in A] \leq \liminf_n \mathbb{P}[X_n \in A]$ za svaki $A \in \mathcal{A}$ tada $X_n \Rightarrow X$.

Korolar 1.15. Ako vrijedi $\mathbb{P}[X \in A] \leq \liminf_n \mathbb{P}[X_n \in A]$ za svaki

$$A \in \left\{ \prod_{i=1}^r \langle a_i, b_i \rangle : a_i < b_i \in \mathbb{R} \right\}$$

tada $X_n \Rightarrow X$.

Dokaz. Tvrdnja slijedi izravno iz prethodnog teorema jer je klasa

$$\mathcal{A} = \left\{ \prod_{i=1}^r \langle a_i, b_i \rangle : a_i < b_i \in \mathbb{R} \right\}$$

poluprsten, i svaki otvoreni interval se može prikazati kao prebrojiva unija poluotvorenih intervala, a svaki otvoren skup se može prikazati kao prebrojiva unija otvorenih intervala. \square

Sljedeći teorem daje nam dovoljne uvjete za donošenje zaključaka o konvergenciji kompozicije.

Teorem 1.16. Neka je h izmjeriva funkcija i D_h skup njezinih točaka prekida. Ako je $\mathbb{P}[X \in D_h] = 0$ i $X_n \Rightarrow X$ onda $h(X_n) \Rightarrow h(X)$.

Dokaz. Neka je F zatvoren skup i $x \in \text{Cl}(h^{-1}(F))$. Tada postoji niz $(x_n)_{n \in \mathbb{N}}$ takav da je $h(x_n) \in F$. Ako još dodatno zahtijevamo da je $x \in D_h^c$ to implicira da je $h(x) \in \text{Cl}(F) = F$. Zato je $D_h^c \cap \text{Cl}(h^{-1}(F)) \subset h^{-1}(F)$. Budući da je $\mathbb{P}[X \in D_h] = 0$ vrijedi $\mathbb{P}[X \in D_h^c] = 1$, iz čega slijedi

$$\begin{aligned} \limsup_n \mathbb{P}[h(X_n) \in F] &= \limsup_n \mathbb{P}[X_n \in h^{-1}(F)] \leq \limsup_n \mathbb{P}[X_n \in \text{Cl}(h^{-1}(F))] \stackrel{\text{Teorem 1.8}}{\leq} \\ &\mathbb{P}[X \in \text{Cl}(h^{-1}(F))] = \mathbb{P}[X \in D_h^c \cap \text{Cl}(h^{-1}(F))] \leq \mathbb{P}[X \in h^{-1}(F)] = \mathbb{P}[h(X) \in F]. \end{aligned}$$

Tvrdnja teorema sada slijedi iz teorema 1.8, odnosno $h(X_n) \Rightarrow h(X)$. \square

Korolar 1.17. Ako $X_n \Rightarrow X$ onda je $\mathbb{E}|X| \leq \liminf_n \mathbb{E}|X_n|$.

Dokaz. Po teoremu 1.16 $|X_n| \Rightarrow |X|$, pa kao i ranije $\mathbb{P}[|X_n| > t] \rightarrow \mathbb{P}[|X| > t]$ za sve osim prebrojivo mnogo t , tvrdnja korolara slijedi po Fatouovoj lemi. \square

Korolar 1.18. Ako vrijedi $\sup_n \mathbb{E}[|X_n|^{1+\varepsilon}] < \infty$ i $X_n \Rightarrow X$ onda $\mathbb{E}[X_n] \rightarrow \mathbb{E}[X]$.

Dokaz. Po korolaru 1.17 X je integrabilna, a po teoremu 1.16 $X_n^+ \Rightarrow X^+$ i $X_n^- \Rightarrow X^-$. Za proizvoljni α vrijedi

$$\begin{aligned} \mathbb{E}[X_n^+] &= \int_{X_n^+ < \alpha} X_n^+ d\mathbb{P} + \int_{X_n^+ \geq \alpha} X_n^+ d\mathbb{P} = \int_0^\alpha x dF_{X_n^+} + \int_{X_n^+ \geq \alpha} X_n^+ d\mathbb{P} = \int_0^\alpha \int_0^x dt dF_{X_n^+} + \int_{X_n^+ \geq \alpha} X_n^+ d\mathbb{P} = \\ &= \int_0^\alpha \int_1^t dF_{X_n^+} dt + \int_{X_n^+ \geq \alpha} X_n^+ d\mathbb{P} = \int_0^\alpha \mathbb{P}[t < X_n^+ < \alpha] dt + \int_{X_n^+ \geq \alpha} X_n^+ d\mathbb{P} \text{ i analognim postupkom} \\ \mathbb{E}[X^+] &= \int_0^\alpha \mathbb{P}[t < X^+ < \alpha] dt + \int_{X^+ \geq \alpha} X^+ d\mathbb{P}. \end{aligned}$$

Budući da je $\sup_n \mathbb{E}[|X_n|^{1+\varepsilon}] < \infty$ za proizvoljni ε postoji dovoljno velik α da vrijedi

$$\int_{X_n^+ \geq \alpha} X_n^+ d\mathbb{P} = \int_{X_n^+ \geq \alpha} X_n^+ \left(\frac{X_n^+}{X_n^+}\right)^\varepsilon d\mathbb{P} \leq \frac{1}{\alpha^\varepsilon} \int_{X_n^+ \geq \alpha} (X_n^+)^{1+\varepsilon} \leq \frac{1}{\alpha^\varepsilon} \mathbb{E}[(X_n^+)^{1+\varepsilon}] \leq \frac{1}{\alpha^\varepsilon} \mathbb{E}[|X_n|^{1+\varepsilon}] < \varepsilon$$

i $\int_{X^+ \geq \alpha} X^+ d\mathbb{P} \leq \varepsilon$. Budući da je $\mathbb{P}[t < X_n^+ < \alpha] \leq 1$, a $\int_0^\alpha 1 dt < \infty$ po Lebesgueovom teoremu o dominiranoj konvergenciji i teoremu 1.8 slijedi

$$\int_0^\alpha \mathbb{P}[t < X_n^+ < \alpha] dt \rightarrow \int_0^\alpha \mathbb{P}[t < X^+ < \alpha] dt.$$

Zbog argumenata vezano uz odabir parametra α sada zaključujemo da $\mathbb{E}[X_n^+] \rightarrow \mathbb{E}[X^+]$, a kako bi cijeli postupak analogno prošao i za X_n^- , odnosno X^- slijedi

$$\mathbb{E}[X_n] = \mathbb{E}[X_n^+] - \mathbb{E}[X_n^-] \rightarrow \mathbb{E}[X^+] - \mathbb{E}[X^-] = \mathbb{E}[X].$$

□

Zaključno s ovim korolarom pokazali smo sve potrebne tvrdnje vezane uz konvergenciju slučajnih varijabli. Okrenimo se sada konstrukciji koja će igrati ključnu ulogu u dokazima mnogih teorema u nastavku rada.

Neka je $\mathbb{X} = (X_1, X_2, \dots)$ slučajni vektor za koji vrijedi $X_k > 0$ za svaki $k \in \mathbb{N}$ i $\sum_{k=1}^{\infty} X_k = 1$. Želimo definirati slučajni vektor $\hat{\mathbb{X}} = (\hat{X}_1, \hat{X}_2, \dots)$ tako da zadovoljava

$$\begin{aligned} \mathbb{P}[\hat{X}_1 = X_i] &= X_i \text{ za svaki } i \\ \mathbb{P}[\hat{X}_2 = X_j \mid \hat{X}_1 = X_i] &= \frac{X_j}{1 - X_i} \text{ za svaki } j \neq i \\ \mathbb{P}[\hat{X}_3 = X_k \mid \hat{X}_1 = X_i, \hat{X}_2 = X_j] &= \frac{X_k}{1 - X_i - X_j} \text{ za svaki } k \neq i, j \\ &\vdots \end{aligned}$$

Formalna konstrukcija takvog vektora dana je u sljedećoj napomeni.

Napomena 1.19. Neka je $S_0 = 0$, $S_i = \sum_{k=1}^i X_k$. Budući da je $\sum_{k=1}^{\infty} X_k = 1$ vrijedi $S_i \rightarrow 1$. Promotrimo slučajne varijable Y_1, Y_2, \dots uniformne na $\langle 0, 1 \rangle$ koje su međusobno nezavisne i nezavisne sa X_k za svaki $k \in \mathbb{N}$. Neka je $\tau_r = i$ kada je $S_{i-1} < Y_r \leq S_i$. Ukoliko je $X_i = 0$ za neko $i \in \mathbb{N}$ onda je $\tau_r \neq i$ za svaki r , a ukoliko je $X_i > 0$ onda τ_r poprima vrijednost i s vjerojatnošću X_i . Neka je $\theta_1 = \tau_1$, neka je $\theta_2 = \tau_r$ gdje je r najmanji indeks za koji je $\theta_1 \neq \tau_r$, $\theta_3 = \tau_s$ gdje je s najmanji indeks za koji je $\theta_1, \theta_2 \neq \tau_s$ i tako dalje. Tada će za vektor $\hat{\mathbb{X}} = (\hat{X}_1, \hat{X}_2, \dots)$ kojeg definiramo s $\hat{X}_u = X_{\theta_u}$ zadovoljavati

$$\begin{aligned} \mathbb{P}[\hat{X}_1 = X_i] &= \mathbb{P}[X_{\theta_1} = X_i] = \mathbb{P}[\theta_1 = i] = \mathbb{P}[\tau_1 = i] = X_i, \text{ za svaki } i, \\ \mathbb{P}[\hat{X}_2 = X_j \mid \hat{X}_1 = X_i] &= \mathbb{P}[X_{\theta_2} = X_j \mid X_{\theta_1} = X_i] = \frac{\mathbb{P}[\theta_2 = j \cap \theta_1 = i]}{\mathbb{P}[\theta_1 = i]} \stackrel{\text{definicija od } \theta_2}{=} \\ \frac{\mathbb{P}\left[\bigcup_{r=2}^{\infty} \tau_1 = \dots = \tau_{r-1} = i, \tau_r = j\right]}{X_i} &\stackrel{\text{nezavisnost}}{=} \frac{\sum_{r=2}^{\infty} \mathbb{P}[\tau_r = j] \cdot \prod_{k=1}^{r-1} \mathbb{P}[\tau_k = i]}{X_i} = \\ \frac{\sum_{r=2}^{\infty} X_j \cdot \prod_{k=1}^{r-1} X_i}{X_i} &= X_j \sum_{r=2}^{\infty} X_i^{r-2} = \frac{X_j}{1 - X_i}, \text{ za svaki } j \neq i, \\ \mathbb{P}[\hat{X}_3 = X_k \mid \hat{X}_1 = X_i, \hat{X}_2 = X_j] &= \mathbb{P}[X_{\theta_3} = X_k \mid X_{\theta_1} = X_i, X_{\theta_2} = X_j] = \\ \frac{\mathbb{P}[\theta_3 = k \cap \theta_2 = j \cap \theta_1 = i]}{\mathbb{P}[\theta_1 = i \cap \theta_2 = j]} &\stackrel{\text{definicija od } \theta_2}{=} \\ \frac{\mathbb{P}\left[\bigcup_{r=2}^{\infty} \bigcup_{s=r+1}^{\infty} \tau_1 = \dots = \tau_{r-1} = i, \tau_r = j, \tau_{r+1}, \dots, \tau_{s-1} \in \{i, j\}, \tau_s = k\right]}{\mathbb{P}\left[\bigcup_{r=2}^{\infty} \tau_1 = \dots = \tau_{r-1} = i, \tau_r = j\right]} &\stackrel{\text{nezavisnost}}{=} \end{aligned}$$

$$\begin{aligned}
 & \frac{\sum_{r=2}^{\infty} \sum_{s=r+1}^{\infty} \left(\prod_{n=1}^{r-1} \mathbb{P}[\tau_n = i] \right) \cdot \mathbb{P}[\tau_r = j] \cdot \left(\prod_{m=r+1}^{s-1} \mathbb{P}[\tau_m \in \{i, j\}] \right) \cdot \mathbb{P}[\tau_s = k]}{\frac{X_i X_j}{1-X_i}} = \\
 & \frac{\sum_{r=2}^{\infty} \sum_{s=r+1}^{\infty} \left(\prod_{n=1}^{r-1} X_i \right) \cdot X_j \cdot \left(\prod_{m=r+1}^{s-1} (X_i + X_j) \right) \cdot X_k}{\frac{X_i X_j}{1-X_i}} = \frac{X_k \cdot X_j \cdot X_i \cdot \sum_{r=2}^{\infty} X_i^{r-2} \cdot \sum_{s=r+1}^{\infty} (X_i + X_j)^{s-(r+1)}}{\frac{X_i X_j}{1-X_i}} = \\
 & \frac{X_k \cdot X_j \cdot X_i \cdot \frac{1}{1-X_i} \cdot \frac{1}{1-X_i-X_j}}{\frac{X_i X_j}{1-X_i}} = \frac{X_k}{1-X_i-X_j}, \text{ za svaki } k \neq i, j,
 \end{aligned}$$

⋮

što smo i željeli pokazati. Jedini problem je što ako niz $(\theta_u)_{u \in \mathbb{N}}$ nije dobro definiran. To bi se moglo dogoditi ako za neki u ne postoji r takav da je $\tau_r \neq \theta_1, \dots, \theta_{u-1}$. To bi značilo da je $X_{\theta_1} + \dots + X_{\theta_{u-1}} = 1$. Tada recimo da je $\theta_v = \infty$ za $v \geq u$ i $\hat{X}_v = 0$. U ovom slučaju ponovno su zadovoljene tražene tvrdnje.

Iz ove napomene dobivamo tvrdnju sljedeće propozicije

Propozicija 1.20. Za ranije definiran vektor $\hat{\mathbb{X}} = (\hat{X}_1, \hat{X}_2, \dots)$ vrijedi

(i) $\hat{X}_k > 0$, za svaki $k \in \mathbb{N}$,

(ii) $\sum_{k=1}^{\infty} \hat{X}_k = 1$

(iii) $\mathbb{P}[\theta_1 = i_1, \dots, \theta_r = i_r \mid \mathbb{X}] = X_{i_1} \frac{X_{i_2}}{1-X_{i_1}} \cdots \frac{X_{i_r}}{1-X_{i_1}-\dots-X_{i_{r-1}}}$.

Dokaz. (i) Tvrdnja vrijedi jer za svaki ω i za svaki $k \in \mathbb{N}$ postoji $i \in \mathbb{N}$ takav da je $\hat{X}_k(\omega) = X_i(\omega) > 0$.

(ii) Zbog uvjeta da za θ_k biramo τ_r s najmanjim indeksom koji je različit od svih ranijih θ_i dobivamo da je svaki ω vektor $\hat{\mathbb{X}}(\omega)$ permutacija vektora $\mathbb{X}(\omega)$. Tvrdnja vrijedi jer je $\sum_{k=1}^{\infty} X_k = 1$.

(iii)

$$\begin{aligned}
 & \mathbb{P}[\theta_1 = i_1, \dots, \theta_r = i_r \mid \mathbb{X}] = \mathbb{P}[\theta_r = i_r \mid \mathbb{X} \cap \theta_1 = i_1 \cap \dots \cap \theta_{r-1} = i_{r-1}] \cdot \\
 & \mathbb{P}[\theta_{r-1} = i_{r-1} \mid \mathbb{X} \cap \theta_1 = i_1 \cap \dots \cap \theta_{r-2} = i_{r-2}] \cdots \mathbb{P}[\theta_1 = i_1 \mid \mathbb{X}] = \\
 & \frac{X_{i_r}}{1-X_{i_1}-\dots-X_{i_{r-1}}} \cdot \frac{X_{i_{r-1}}}{1-X_{i_1}-\dots-X_{i_{r-2}}} \cdots \frac{X_{i_2}}{1-X_{i_1}} \cdot X_{i_1}.
 \end{aligned}$$

□

1.3 Neki pojmovi teorije brojeva

Definicija 1.21. Ako je $g(x) > 0$ za sve $x \geq a$, kažemo da je

$$f(x) = O(g(x))$$

ako je $f(x)(g(x))^{-1}$ ograničeno za svaki $x \geq a$, odnosno postoji konstanta $M > 0$ takva da je $|f(x)| \leq Mg(x)$ za svaki $x \geq a$. Jednadžba oblika

$$f(x) = h(x) + O(g(x))$$

znači da je $f(x) - h(x) = O(g(x))$.

Uz ovu notaciju vrijedi tvrdnja sljedeće propozicije

Propozicija 1.22. Za funkciju Γ definiranu pomoću (1.1) i Euler-Mascheronovu konstantu γ vrijedi

$$\alpha\Gamma(\alpha) = \Gamma(1 + \alpha) = 1 - \gamma\alpha + O(\alpha^2), \text{ za } \alpha > 0.$$

Dokaz. Budući da je po (1.2) $\Gamma(1) = 1$, po (1.4) $\Gamma'(1) = -\gamma$, a kako je funkcija Γ analitička slijedi da je $\Gamma(1 + \alpha) = 1 - \gamma\alpha + O(\alpha^2)$. Ukoliko na dobiveno primijenimo tvrdnju propozicije 1.2 slijedi tražena tvrdnja. \square

Pokazat ćemo još dva teorema koji će se pokazati značajni u poglavlju 2.

Teorem 1.23.

$$\sum_{p \leq x} \frac{\log(p)}{p} = \log(x) + O(1), \quad x \rightarrow \infty,$$

gdje se sumacija vrši po prostim brojevima $p \leq x$.

Dokaz. Neka je

$$\Lambda(n) := \begin{cases} \log(p) & n = p^k, \\ 0, & \text{inačne.} \end{cases}$$

Po primjeru 5.4. u Dujella [2] vrijedi

$$\sum_{n \leq x} \frac{\Lambda(n)}{n} = \log(x) + O(1). \quad (1.8)$$

Također vrijedi

$$\sum_{n \leq x} \frac{\Lambda(n)}{n} - \sum_{p \leq x} \frac{\log(p)}{p} = \sum_{\substack{p^m \leq n \\ m \geq 2}} \frac{\log(p)}{p^m} \leq \sum_p \log(p) \sum_{m=2}^{\infty} \frac{1}{p^m} = \sum_p \frac{\log(p)}{p^2} \frac{p}{p-1} \leq 2 \sum_p \frac{\log(p)}{p^2} < \infty.$$

Iz dobivenog slijedi da je

$$\sum_{p \leq x} \frac{\log(p)}{p} = \sum_{n \leq x} \frac{\Lambda(n)}{n} + O(1),$$

što zajedno sa (1.8) daje tvrdnju teorema. \square

Teorem 1.24.

$$\sum_{p \leq x} \log(p) = O(x), \quad x \rightarrow \infty,$$

gdje se sumacija vrši po prostim brojevima $p \leq x$.

Dokaz. Budući da je

$$\binom{2n}{k} = \frac{(2n)!}{(n)!(n)!}$$

ako je $n < p \leq 2n$ onda prost broj p dijeli $\binom{2n}{k}$. Zato označimo li s $\theta(x) = \sum_{p \leq x} \log(p)$ vrijedi

$$\theta(2n) - \theta(n) = \sum_{n < p \leq 2n} \log(p) = \log \left(\prod_{n < p \leq 2n} p \right) \leq \log \left(\binom{2n}{k} \right) \leq \log(2^{2n}) = 2n \log(2). \quad (1.9)$$

Za $n = 2^k$ zato vrijedi

$$\theta(2^{k+1}) \leq \theta(2^k) + 2^{k+1} \log(2) \leq \dots \leq \theta(2^0) + 2^{1+1} \log(2) + \dots + 2^{k+1} \log(2) \leq 2^{k+2} \log(2).$$

Za $2^m \leq n < 2^{m+1}$ zato vrijedi

$$\theta(n) \leq \theta(2^m) + \theta(2^{m+1}) - \theta(2^m) \leq 2^{m+1} \log(2) + 2^{m+1} \log(2) = 4 \cdot 2^m \log(2) \leq 4n \log(2),$$

iz čega po definiciji 1.21 slijedi tvrdnja teorema. \square

1.4 Formula uključivanja i isključivanja

Promotrimo skup $I \subset \{1, 2, \dots, n\}$ i označimo s $|I|$ kardinalitet skupa I . Neka je $r \in \mathbb{R}$ i neka su $S_1, \dots, S_n \subset \mathbb{R}^r$. Tada vrijedi sljedeći teorem

Teorem 1.25. *Za svaku integrabilnu funkciju $f : \mathbb{R}^r \rightarrow \mathbb{R}$ vrijedi*

$$\int_{S_1 \cap \dots \cap S_n} f d\lambda = \sum_{I \subset \{1, 2, \dots, n\}} (-1)^{|I|} \int_{\bigcap_{i \in I} S_i^c} f d\lambda, \quad (1.10)$$

gdje se u slučaju $I = \emptyset$ smatra $\bigcap_{i \in I} S_i^c = \mathbb{R}^r$.

Dokaz. Za $i = 1, \dots, n$ označimo s $A_i := S_i^c$ te s $A := S_1 \cap \dots \cap S_n = A_1^c \cap \dots \cap A_n^c$. Tada je $A^c = A_1 \cup \dots \cup A_n$ te vrijedi

$$(\mathbb{1}_{A^c} - \mathbb{1}_{A_1}) \cdot (\mathbb{1}_{A^c} - \mathbb{1}_{A_2}) \cdot \dots \cdot (\mathbb{1}_{A^c} - \mathbb{1}_{A_n}) = 0,$$

jer je za $x \notin A_1 \cup \dots \cup A_n$ vrijednost svake karakteristične funkcije 0, a za $x \in A_1 \cup \dots \cup A_n$ postoji $i \in \{1, 2, \dots, n\}$ takav da je $x \in A_i$ pa je $(\mathbb{1}_{A^c} - \mathbb{1}_{A_i}) = 0$. Raspisivanjem ovog izraza i primjenom formule $\mathbb{1}_B \cdot \mathbb{1}_C = \mathbb{1}_{B \cap C}$ dobivamo

$$\begin{aligned} \sum_{I \subset \{1, 2, \dots, n\}} (-1)^{|I|} \mathbb{1}_{A^c \cap (\bigcap_{i \in I} A_i)} &= 0 \implies \mathbb{1}_{A^c} + \sum_{I \subset \{1, 2, \dots, n\}} (-1)^{|I|} \mathbb{1}_{\bigcap_{i \in I} A_i} = 0 \implies \\ \mathbb{1} - \mathbb{1}_A + \sum_{\emptyset \neq I \subset \{1, 2, \dots, n\}} (-1)^{|I|} \mathbb{1}_{\bigcap_{i \in I} A_i} &= 0 \implies \mathbb{1}_A = \sum_{I \subset \{1, 2, \dots, n\}} (-1)^{|I|} \mathbb{1}_{\bigcap_{i \in I} A_i}. \end{aligned} \quad (1.11)$$

Za integrabilnu funkciju $f : \mathbb{R}^r \rightarrow \mathbb{R}$ zato vrijedi

$$\begin{aligned} \int_{S_1 \cap \dots \cap S_n} f d\lambda &= \int_{\mathbb{R}^r} f \mathbb{1}_{S_1 \cap \dots \cap S_n} d\lambda = \int_{\mathbb{R}^r} f \mathbb{1}_A d\lambda \stackrel{(1.11)}{=} \int_{\mathbb{R}^r} f \sum_{I \subset \{1, 2, \dots, n\}} (-1)^{|I|} \mathbb{1}_{\bigcap_{i \in I} A_i} d\lambda = \\ \sum_{I \subset \{1, 2, \dots, n\}} (-1)^{|I|} \int_{\mathbb{R}^r} f \mathbb{1}_{\bigcap_{i \in I} A_i} d\lambda &= \sum_{I \subset \{1, 2, \dots, n\}} (-1)^{|I|} \int_{\mathbb{R}^r} f \mathbb{1}_{(\bigcap_{i \in I} S_i^c)} d\lambda = \sum_{I \subset \{1, 2, \dots, n\}} (-1)^{|I|} \int_{\bigcap_{i \in I} S_i^c} f d\lambda. \end{aligned}$$

Ovime je pokazana tvrdnja teorema. □

Korolar 1.26. Neka je dan skup $X \subset \mathbb{R}^r$ i integrabilna funkcija $f : \mathbb{R}^r \rightarrow \mathbb{R}$. Tada je

$$\int_{X \cap S_1 \cap \dots \cap S_n} f d\lambda = \sum_{I \subset \{1, 2, \dots, n\}} (-1)^{|I|} \int_{X \cap (\bigcap_{i \in I} S_i^c)} f d\lambda, \quad (1.12)$$

gdje se u slučaju $I = \emptyset$ smatra $\bigcap_{i \in I} S_i^c = \mathbb{R}^r$.

Dokaz. Primijenimo li (1.12) na funkciju $\tilde{f} = f\mathbb{1}_X$ dobivamo

$$\int_{X \cap S_1 \cap \dots \cap S_n} f d\lambda = \sum_{I \subset \{1, 2, \dots, n\}} (-1)^{|I|} \int_{X \cap (\bigcap_{i \in I} S_i^c)} f d\lambda$$

što je trebalo pokazati. □

Poglavlje 2

Motivacijski primjeri

U ovom poglavlju definirat ćemo dva naizgled nepovezana vjerojatnosna modela i proučiti njihova svojstva. Kasnije će nam poslužiti za definiciju Poisson-Dirichletove razdiobe koja je centralni objekt ovog rada.

2.1 Ciklusi slučajnih permutacija

Prvi prostor koji promatramo jest prostor S_n svih permutacija skupa $\{1, 2, \dots, n\}$ s jednakom vjerojatnošću svake permutacije. Neka je dana slučajna permutacija $\sigma \in S_n$ i neka se ta permutacija rastavlja na cikluse $\{C_1, C_2, \dots\}$, gdje je taj skup konačan. Dodatno pretpostavimo da C_1 sadrži 1, C_2 sadrži najmanji element skupa $\{1, 2, \dots, n\}$ koji nije sadržan u C_1 , općenito C_k sadrži najmanji element skupa $\{1, 2, \dots, n\}$ koji nije sadržan ni u jednom C_l za $l < k$. Uvedimo oznake za duljine $(|C_1|, |C_2|, \dots)$ odgovarajućih ciklusa gdje po potrebi ovaj niz možemo interpretirati kao beskonačan ukoliko uzmemo da je $|C_k| = 0$ za k veći od broja ciklusa permutacije σ . Primijetimo da trivijalno vrijede sljedeće činjenice $\mathbb{P}(\{\sigma\}) = \frac{1}{n!}$ i $\sum_{k=1}^{\infty} |C_k| = n$. Odnosno $\sum_{k=1}^{\infty} \frac{|C_k|}{n} = 1$ i $\frac{|C_k|}{n} \geq 0$ za svaki $k \in \mathbb{N}$.

Primjer 2.1. Promotrimo kako uvedene oznake izgledaju za $n = 9$ i na odabranu permutaciju $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 3 & 9 & 4 & 1 & 5 & 8 & 2 & 6 & 7 \end{pmatrix}$, odnosno zapisano pomoću ciklusa $\sigma = (1, 3, 4)(2, 9, 7)(5)(6, 8)$. Tada je uz gornje oznake

$$C_1 = (1, 3, 4), C_2 = (2, 9, 7), C_3 = (5), C_4 = (6, 8)$$

iz čega slijedi

$$|C_1| = 3, |C_2| = 3, |C_3| = 1, |C_4| = 2.$$

Odnosno dobivamo $\sum_{k=1}^{\infty} \frac{|C_k|}{9} = \frac{3+3+1+2}{9} = 1$.

U navedenom primjeru promatrali smo svojstva fiksne permutacije σ , a sljedeća lema nam govori više o ciklusima slučajno odabrane permutacije.

Lema 2.2. Neka je definiran vjerojatnosni prostor kao gore i neka se σ rastavlja na odgovarajuće cikluse. Tada vrijedi $\mathbb{P}(\binom{|C_1|}{n} = \frac{i}{n}) = \mathbb{P}(|C_1| = i) = \frac{1}{n}$ za $1 \leq i \leq n$

Dokaz.

$$\begin{aligned} \mathbb{P}(|C_1| = i) &= \frac{\text{broj permutacija s prvim ciklusom duljine } i}{n!} = \\ &= \frac{(n-1) \cdot \dots \cdot (n-i+1) \cdot (n-i) \cdot \dots \cdot 2 \cdot 1}{n!} = \frac{1}{n}. \end{aligned}$$

Gdje druga jednakost vrijedi jer znamo da je prvi element ciklusa C_1 1, a ostale odaberemo na $(n-1) \cdot \dots \cdot (n-i+1)$ načina. Tako dobivamo C_1 duljine i a zatim na $(n-i)!$ načina ispermutiramo preostale elemente. \square

Pretpostavimo nadalje da je $|C_1| = i$, odnosno da vrijedi $C_1 = (1, a_1, a_2, \dots, a_{i-1})$. Tada preostali ciklusi čine permutaciju skupa $\{1, 2, \dots, n\} \setminus \{1, a_1, a_2, \dots, a_{i-1}\}$. Pa je uvjetno na događaj da je $|C_1| = i$, vjerojatnost da je $|C_2| = j$ upravo vjerojatnost da je prvi ciklus permutacije $(n-i)$ -članog skupa duljine j . Iz ovih napomena i prethodne leme slijedi

$$\mathbb{P}(|C_1| = i, |C_2| = j) = \mathbb{P}(|C_2| = j \mid |C_1| = i) \cdot \mathbb{P}(|C_1| = i) = \frac{1}{n-i} \cdot \frac{1}{n}$$

Uočimo da ovaj postupak možemo nastaviti te induktivno za $i_1, i_2, \dots, i_k > 0$ takve da je $\sum_{j=1}^k i_j \leq n$ zaključujemo da vrijedi:

$$\begin{aligned} \mathbb{P}(|C_1| = i_1, |C_2| = i_2, \dots, |C_k| = i_k) &= \\ \mathbb{P}(|C_k| = i_k \mid |C_1| = i_1, \dots, |C_{k-1}| = i_{k-1}) \cdot \mathbb{P}(|C_1| = i_1, \dots, |C_{k-1}| = i_{k-1}) &\stackrel{\text{induktivna pretpostavka}}{=} \\ \frac{1}{n - (i_1 + \dots + i_{k-1})} \cdot \left(\frac{1}{n} \cdot \prod_{j=1}^{k-2} \frac{1}{n - (i_1 + \dots + i_j)} \right) &= \\ \frac{1}{n} \cdot \frac{1}{n - i_1} \cdot \dots \cdot \frac{1}{n - (i_1 + i_2 + \dots + i_{k-1})} & \end{aligned}$$

Ukoliko je $\sum_{j=1}^k i_j > n$ tada je tražena vjerojatnost nula, jer je nemoguće da je zbroj duljina ciklusa veći od n .

Ovim razmatranjima dokazan je sljedeći rezultat

Propozicija 2.3. (a) Slučajna varijabla $\binom{|C_1|}{n}$ uniformno je distribuirana na skupu $\left\{\frac{i}{n}, i \leq n\right\}$.

- (b) uvjetno na $\frac{|C_1|}{n} = \frac{i}{n}$ slučajna varijabla $\frac{|C_2|}{n}$ uniformno je distribuirana na skupu $\left\{\frac{j}{n}, j \leq n - i = n - |C_1|\right\}$.
- (c) uvjetno na $\frac{|C_1|}{n} = \frac{i_1}{n}, \frac{|C_2|}{n} = \frac{i_2}{n}, \dots, \frac{|C_{k-1}|}{n} = \frac{i_{k-1}}{n}$ slučajna varijabla $\frac{|C_k|}{n} = \frac{i_k}{n}$ uniformno je distribuirana na skupu $\left\{\frac{i_k}{n}, i_k \leq n - i_1 - i_2 - \dots - i_{k-1} = n - |C_1| - |C_2| \dots - |C_{k-1}|\right\}$.

Sada smo spremni razmotriti što se događa u graničnim situacijama. Prethodna propozicija daje nam za pravo pretpostaviti da će promatrane slučajne varijable u nekom smislu konvergirati ka slučajnim varijablama koje su uniformno distribuirane na $[0, 1]$. U tu svrhu definirajmo slučajne varijable $U_1, U_2, \dots, U_k, \dots$ koje su sve nezavisne i uniformne na $[0, 1]$. Pokažimo prvo sljedeću lemu.

Lema 2.4. Neka su $|C_1|$ i U_1 slučajne varijable definirane kao gore. Tada vrijedi

$$\frac{|C_1|}{n} \Rightarrow U_1.$$

Dokaz. Neka je $f : [0, 1] \rightarrow \mathbb{R}$ neprekidna funkcija, a budući da je skup $[0, 1]$ kompaktan funkcija f je i ograničena. Tada vrijedi

$$\mathbb{E} \left[f \left(\frac{|C_1|}{n} \right) \right] = \sum_{i=1}^n \frac{1}{n} \cdot f \left(\frac{i}{n} \right) = \frac{1}{n} \sum_{i=1}^n f \left(\frac{i}{n} \right) \xrightarrow{n \rightarrow \infty} \int_0^1 f(x) dx = \mathbb{E} [f(U_1)]$$

Gdje konvergencija vrijedi jer je f neprekidna pa je Riemann integrabilna iz čega slijedi da Darbuoxova suma konvergira prema Reimannovom integralu funkcije f . Po teoremu 1.8 slijedi tvrdnja leme. \square

Sličnu tvrdnju možemo dobiti i za $\frac{|C_2|}{n}$ samo što u ovom slučaju moramo danu slučajnu varijablu u nekom smislu skalirati kako bi ponovno dobili konvergenciju prema uniformnoj na $[0, 1]$.

Lema 2.5. Neka su $|C_1|, |C_2|$ i U_1, U_2 slučajne varijable definirane kao gore. Tada vrijedi

$$\left(\frac{|C_1|}{n}, \frac{|C_2|}{n - |C_1|} \right) \Rightarrow (U_1, U_2).$$

Dokaz. Uvedimo prvo oznake koje će nam pojednostavniti zapis. Neka je $L_i^n = \frac{|C_i|}{n}$. Nadalje neka je $f : [0, 1]^2 \rightarrow \mathbb{R}$ neprekidna funkcija, a budući da je skup $[0, 1]^2$ kompaktan

funkcija f je i ograničena.. Tada vrijedi

$$\begin{aligned} \mathbb{E} \left[f \left(L_1^n, \frac{L_2^n}{1 - L_1^n} \right) \right] &= \sum_{i+j=1}^n \frac{1}{n} \cdot \frac{1}{n-i} \cdot f \left(\frac{i}{n}, \frac{\frac{j}{n}}{1 - \frac{i}{n}} \right) = \\ &= \frac{1}{n^2} \sum_{i+j=1}^n \frac{f \left(\frac{i}{n}, \frac{\frac{j}{n}}{1 - \frac{i}{n}} \right)}{1 - \frac{i}{n}} \xrightarrow{n \rightarrow \infty} \iint_{0 \leq x_1 + x_2 \leq 1} \frac{f(x_1, \frac{x_2}{1-x_1})}{1 - x_1} d(x_1, x_2) \stackrel{\text{Fubini}}{=} \\ &= \int_0^1 \int_0^{1-x_1} \frac{f(x_1, \frac{x_2}{1-x_1})}{1 - x_1} dx_2 dx_1 \left[y = \frac{x_2}{1-x_1}, dy = \frac{dx_2}{1-x_1} \right] = \\ &= \int_0^1 \int_0^1 f(x_1, y) dy dx_1 = \iint_{[0,1]^2} f(x, y) d(x, y) = \mathbb{E} [f(U_1, U_2)] \end{aligned}$$

Gdje konvergencija vrijedi jer je funkcija $(x_1, x_2) \rightarrow f(x_1, \frac{x_2}{1-x_1}) / (1 - x_1)$ neprekidna pa je Riemann integrabilna iz čega slijedi da Darbuoxova suma konvergira prema Reimannovom integralu funkcije. Po teoremu 1.8 slijedi tvrdnja leme. \square

Dokazani rezultat se može induktivno poopćiti s dvije na r slučajnih varijabli. To je formalno iskazano sljedećim teoremom.

Teorem 2.6. *Neka su $|C_1|, |C_2|, \dots, |C_r|$ i U_1, U_2, \dots, U_r slučajne varijable definirane kao gore. Tada vrijedi*

$$\left(\frac{|C_1|}{n}, \frac{\frac{|C_2|}{n}}{1 - \frac{|C_1|}{n}}, \dots, \frac{\frac{|C_r|}{n}}{1 - \frac{|C_1|}{n} - \frac{|C_2|}{n} - \dots - \frac{|C_{r-1}|}{n}} \right) \Rightarrow (U_1, U_2, \dots, U_r).$$

Prije nego krenemo na dokaz teorema uočimo da bi nam ponovno, kao i u dokazu prethodne leme, uvođenje novih oznaka bitno pojednostavilo zapis u iskazu teorema. Zato prvo navedimo sljedeću definiciju.

Definicija 2.7. *Neka su $|C_1|, |C_2|, \dots, |C_r|$ slučajne varijable koje odgovaraju duljinama ciklusa permutacije n -članog skupa kao što je definirano ranije. Neka su tada $L_1^n, L_2^n, \dots, L_r^n$ slučajne varijable na istom vjerojatnosnom prostoru definirane relacijom*

$$L_i^n := \frac{|C_i|}{n}, i \in \{1, 2, \dots, r\}$$

i neka je $\mathbb{L}^n = (L_1^n, L_2^n, \dots, L_r^n, 0, \dots)$.

Uz ove oznake za dokaz teorema potrebno je pokazati da vrijedi

$$\left(L_1^n, \frac{L_2^n}{1 - L_1^n}, \dots, \frac{L_r^n}{1 - L_1^n - L_2^n - \dots - L_{r-1}^n} \right) \Rightarrow (U_1, U_2, \dots, U_r).$$

Za dokaz teorema potrebna nam je još jedna tehnička lema koja se dokazuje metodom matematičke indukcije.

Lema 2.8. Za proizvoljan $r \in \mathbb{N}$ i za svaku neprekidnu funkciju $f : [0, 1]^r \rightarrow \mathbb{R}$ vrijedi

$$\int_{0 \leq x_1 + \dots + x_r \leq 1} \frac{f\left(x_1, \frac{x_2}{1-x_1}, \dots, \frac{x_r}{1-x_1-\dots-x_{r-1}}\right)}{(1-x_1) \dots (1-x_1-\dots-x_{r-1})} d(x_1, \dots, x_r) = \int_{[0,1]^r} f(y_1, \dots, y_r) d(y_1, \dots, y_r).$$

Dokaz. Budući da su slučajevi kada je $r = 1, 2$ već napravljeni u dokazima prethodnih lema bazu indukcije već imamo. Zato pretpostavimo da za svaki $s < r$ i za svaku neprekidnu funkciju $f : [0, 1]^s \rightarrow \mathbb{R}$ vrijedi tražena tvrdnja. Nadalje neka je $f : [0, 1]^r \rightarrow \mathbb{R}$ neprekidna funkcija. Tada vrijedi

$$\begin{aligned} & \int_{0 \leq x_1 + \dots + x_r \leq 1} \frac{f\left(x_1, \frac{x_2}{1-x_1}, \dots, \frac{x_r}{1-x_1-\dots-x_{r-1}}\right)}{(1-x_1) \dots (1-x_1-\dots-x_{r-1})} d(x_1, \dots, x_r) \stackrel{\text{Fubini}}{=} \\ & \int_{0 \leq x_1 + \dots + x_{r-1} \leq 1} \int_0^{1-x_1-\dots-x_{r-1}} \frac{f\left(x_1, \frac{x_2}{1-x_1}, \dots, \frac{x_r}{1-x_1-\dots-x_{r-1}}\right)}{(1-x_1) \dots (1-x_1-\dots-x_{r-1})} dx_r d(x_1, x_2, \dots, x_{r-1}) = \\ & \left[y = \frac{x_r}{1-x_1-\dots-x_{r-1}}, dy = \frac{dx_r}{1-x_1-\dots-x_{r-1}} \right] \\ & \int_{0 \leq x_1 + \dots + x_{r-1} \leq 1} \frac{\int_0^1 f\left(x_1, \dots, \frac{x_{r-1}}{1-x_1-\dots-x_{r-2}}, y\right) dy}{(1-x_1) \dots (1-x_1-\dots-x_{r-2})} d(x_1, x_2, \dots, x_{r-1}) = \end{aligned}$$

Primjenom induktivne pretpostavke za $s = r - 1$ i za funkciju

$$(x_1, x_2, \dots, x_{r-1}) \rightarrow \int_0^1 f\left(x_1, \dots, \frac{x_{r-1}}{1-x_1-\dots-x_{r-2}}, y\right) dy$$

$$\begin{aligned} & \int_{[0,1]^{r-1}} \int_0^1 f(y_1, \dots, y_{r-1}, y) dy d(y_1, \dots, y_{r-1}) \stackrel{\text{Fubini}}{=} \\ & \int_{[0,1]^r} f(y_1, y_2, \dots, y_r) d(y_1, y_2, \dots, y_r). \end{aligned}$$

Odnosno tražena tvrdnja vrijedi za r čime je dokazana lema. \square

Sada smo spremni za dokaz tvrdnje teorema 2.6

Dokaz. Neka je $f : [0, 1]^r \rightarrow \mathbb{R}$ neprekidna funkcija, koja je ponovno i ograničena. Tada vrijedi

$$\begin{aligned} & \mathbb{E} \left[f \left(L_1^n, \frac{L_2^n}{1 - L_1^n}, \dots, \frac{L_r^n}{1 - L_1^n - L_2^n - \dots - L_{r-1}^n} \right) \right] = \\ & \sum_{i_1 + \dots + i_r = 1}^n \frac{1}{n} \cdot \frac{1}{n - i_1} \cdot \dots \cdot \frac{1}{n - i_1 - \dots - i_{r-1}} \cdot f \left(\frac{i_1}{n}, \frac{\frac{i_2}{n}}{1 - \frac{i_1}{n}}, \dots, \frac{\frac{i_r}{n}}{1 - \frac{i_1}{n} - \dots - \frac{i_{r-1}}{n}} \right) = \\ & \frac{1}{n^r} \sum_{i_1 + \dots + i_r = 1}^n \frac{f \left(\frac{i_1}{n}, \frac{\frac{i_2}{n}}{1 - \frac{i_1}{n}}, \dots, \frac{\frac{i_r}{n}}{1 - \frac{i_1}{n} - \dots - \frac{i_{r-1}}{n}} \right)}{\left(1 - \frac{i_1}{n}\right) \cdot \left(1 - \frac{i_1}{n} - \frac{i_2}{n}\right) \cdot \dots \cdot \left(1 - \frac{i_1}{n} - \dots - \frac{i_{r-1}}{n}\right)} \xrightarrow{n \rightarrow \infty} \\ & \int_{0 \leq x_1 + \dots + x_r \leq 1} \frac{f \left(x_1, \frac{x_2}{1 - x_1}, \dots, \frac{x_r}{1 - x_1 - \dots - x_{r-1}} \right)}{(1 - x_1) \cdot (1 - x_1 - x_2) \cdot \dots \cdot (1 - x_1 - \dots - x_{r-1})} d(x_1, x_2, \dots, x_r) \stackrel{\text{lema 2.8}}{=} \\ & \int_{[0,1]^r} f(y_1, y_2, \dots, y_r) d(y_1, y_2, \dots, y_r) = \mathbb{E}[f(U_1, U_2, \dots, U_r)]. \end{aligned}$$

Gdje konvergencija vrijedi jer je funkcija definirana svojim djelovanjem na proizvoljni vektor

$$(x_1, x_2, \dots, x_r) \rightarrow \frac{f \left(x_1, \frac{x_2}{1 - x_1}, \dots, \frac{x_r}{1 - x_1 - \dots - x_{r-1}} \right)}{(1 - x_1) \cdot (1 - x_1 - x_2) \cdot \dots \cdot (1 - x_1 - \dots - x_{r-1})}$$

neprekidna pa je Riemann integrabilna iz čega slijedi da Darbuoxova suma konvergira prema Reimannovom integralu funkcije. Ponovno po teoremu 1.8 slijedi tvrdnja teorema. \square

Iz prethodno dokazanog teorema u slučaju $r = 2$ slijedi $\left(L_1^n, \frac{L_2^n}{1 - L_1^n}\right) \Rightarrow (U_1, U_2)$, a iz toga po teoremu 1.16 zaključujemo da vrijedi

$$(L_1^n, L_2^n) \Rightarrow (U_1, (1 - U_1) U_2).$$

Za $r = 3$ smo pokazali da $\left(L_1^n, \frac{L_2^n}{1 - L_1^n}, \frac{L_3^n}{1 - L_1^n - L_2^n}\right) \Rightarrow (U_1, U_2, U_3)$ odnosno ponovno po teoremu 1.16 dobivamo

$$\begin{aligned} (L_1^n, L_2^n, L_3^n) & \Rightarrow (U_1, (1 - U_1) U_2, U_3 (1 - U_1 - (1 - U_1) U_2)) \\ & = (U_1, (1 - U_1) U_2, (1 - U_1)(1 - U_2) U_3). \end{aligned}$$

Induktivnim postupkom za proizvoljan $r \in \mathbb{N}$ dobivamo

$$(L_1^n, L_2^n, \dots, L_r^n) \Rightarrow (U_1, (1 - U_1)U_2, (1 - U_1) \dots (1 - U_{r-1})U_r).$$

Ovo nas navodi na ideju da promatramo $G_1 := U_1, G_2 := (U_1, (1 - U_1)U_2), \dots, G_r := (U_1, (1 - U_1)U_2, (1 - U_1) \dots (1 - U_{r-1})U_r)$. Uz tako definirane slučajne varijable iz prethodnih napomena slijedi tvrdnja sljedećeg korolar.

Korolar 2.9.

$$(L_1^n, L_2^n, \dots, L_r^n) \Rightarrow (G_1, G_2, \dots, G_r).$$

Korolar 2.10. Neka je vektor $\mathbb{L}^n = (L_1^n, L_2^n, \dots, L_r^n, \dots)$ i $\mathbb{G} := (G_1, G_2, \dots, G_r, \dots)$. Tada vrijedi

$$\mathbb{L}^n \Rightarrow \mathbb{G}.$$

Dokaz. Neka je funkcija $\pi_r : \mathbb{R}^\infty \rightarrow \mathbb{R}^r$ projekcija na prvih r koordinata, odnosno neka na vektor $(x_1, x_2, \dots) \in \mathbb{R}^\infty$ djeluje na sljedeći način $\pi_r(x_1, x_2, \dots) = (x_1, x_2, \dots, x_r)$. U propoziciji 1.11 rečeno je da $\mathbb{L}^n \Rightarrow \mathbb{G}$ ako i samo ako $\mathbb{P}[X_n \in \pi_r^{-1}(A)] \rightarrow \mathbb{P}[X \in \pi_r^{-1}(A)]$ za svaki $r \in \mathbb{N}$ i $A \in \mathbb{R}^r$ takav da je $\mathbb{P}[X \in \partial\pi_r^{-1}(A)] = 0$. U teoremu 1.8 je pokazano da je $\mathbb{P}[\pi_r(X_n) \in A] \rightarrow \mathbb{P}[\pi_r(X) \in A]$ za svaki $A \in \mathbb{R}^r$ takav da je $\mathbb{P}[\pi_r(X) \in \partial A] = 0$ ekvivalentno tome da $\pi_r(X_n) \Rightarrow \pi_r(X)$. Budući da je $\mathbb{P}[X_n \in \pi_r^{-1}(A)] = \mathbb{P}[\pi_r(X_n) \in A]$, $\mathbb{P}[X \in \pi_r^{-1}(A)] = \mathbb{P}[\pi_r(X) \in A]$ i $\mathbb{P}[X \in \partial\pi_r^{-1}(A)] = \mathbb{P}[\pi_r(X) \in \partial A]$ te zbog korolar 2.9 slijedi $(L_1^n, L_2^n, \dots, L_r^n, \dots) \Rightarrow (G_1, G_2, \dots, G_r, \dots)$. \square

Napomena 2.11. Slučajne varijable G_1, G_2, \dots imaju sljedeću intuitivnu interpretaciju: Promatrajmo jedinični interval $I_0 = [0, 1]$ kako je po definiciji $G_1 = U_1$ znamo da je G_1 uniformna slučajna varijabla na I_0 . Slučajna varijabla G_2 uvjetno na realizaciju slučajne varijable G_1 je uniformna na intervalu $1 - G_1$. Odnosno (G_1, G_2) možemo zamišljati kao da režemo I_0 u slučajnom omjeru $1 - U_1 : U_1$ i odbacujemo drugi dio odsječenog intervala, odnosno onaj duljine $U_1 = G_1$. Zatim početni dio intervala I_0 , odnosno interval $I_1 = [0, 1 - U_1] = [0, 1 - G_1]$ ponovno režemo u slučajnom omjeru $1 - U_2 : U_2$ te odbacujemo drugi dio intervala, odnosno onaj duljine $U_2(1 - U_1) = G_2$. Ovaj postupak možemo nastaviti, tako ćemo u sljedećem koraku uzeti početni dio intervala I_1 , odnosno interval $I_2 = [0, (1 - U_1)(1 - U_2)] = [0, 1 - G_1 - G_2]$ koji ponovno režemo. U r -tom koraku, za proizvoljan $r \in \mathbb{N}$ dobivamo interval $I_r = [0, (1 - U_1) \cdot (1 - U_2) \cdot \dots \cdot (1 - U_r)] = [0, 1 - G_1 - G_2 - \dots - G_r]$.

Dosta je intuitivno za pretpostaviti da je $\lim_{r \rightarrow \infty} I_r = \{0\}$. Također iz konstrukcije intervala I_r , odnosno njegove karakterizacije preko G_i , koristeći trivijalnu činjenicu da je

početni interval I_0 bio je duljine 1 i uz oznaku $|I_r|$ za duljinu intervala I_r uočavamo da vrijedi

$$\sum_{i=1}^{\infty} G_i = 1 - \prod_{i=1}^{\infty} (1 - U_i) = 1 - \lim_{r \rightarrow \infty} |I_r| = 1$$

Zaključci ove napomene mogu se dokazati i formalno, uz činjenicu da će konvergencija koju smo u napomeni prihvatili intuitivno sada vrijediti gotovo sigurno.

Nas zapravo zanima $\lim_{r \rightarrow \infty} (1 - U_1) \cdot (1 - U_2) \cdot \dots \cdot (1 - U_r)$ g.s. Kako su slučajne varijable $U_1, U_2, \dots, U_r, \dots$ nezavisne i uniformne na intervalu $[0, 1]$ vjerojatnost da ih je samo konačno mnogo veće ili jednako $\frac{1}{2}$ je 0. Zato postoji niz $(i_k)_{k \in \mathbb{N}}$ takav da za svaki $k \in \mathbb{N}$ vrijedi $U_{i_k} \geq \frac{1}{2}$ g.s.. Neka je r proizvoljan, označimo s $n(r)$ najveći k takav da je $i_k \leq r$. Budući da je $\lim_{k \rightarrow \infty} i_k = \infty$ slijedi da je $\lim_{r \rightarrow \infty} n(r) = \infty$. Budući da za svaki od faktora u promatranom produktu vrijedi $0 \leq (1 - U_i) \leq 1$ g.s. za proizvoljan r imamo sljedeći niz nejednakosti:

$$0 \leq (1 - U_1) \cdot (1 - U_2) \cdot \dots \cdot (1 - U_r) \leq (1 - U_{i_1}) \cdot (1 - U_{i_2}) \cdot \dots \cdot (1 - U_{i_{n(r)}}) \leq \left(\frac{1}{2}\right)^{n(r)} \text{ g.s.}$$

Odnosno kad pustimo $r \rightarrow \infty$ po teoremu o sendviču slijedi tražena konvergencija.

Prethodna napomena i ovi formalni zaključci nam zapravo govore da o rastavu slučajne permutacije na cikluse možemo razmišljati kao o rezanju intervala u slučajnom omjeru.

Ono što će nas u nastavku zanimati jest vektor \mathbb{L}^n , ali s komponentama sortiranim silazno po veličini odnosno vektor

$$\mathbb{L}_{(\cdot)}^n := (L_{(1)}^n, L_{(2)}^n, \dots, L_{(r)}^n, \dots), \quad L_{(1)}^n \geq L_{(2)}^n \geq \dots \geq L_{(r)}^n \geq \dots$$

Intuitivno očekujemo da će iz $\mathbb{L}^n \implies \mathbb{G}$ slijediti $\mathbb{L}_{(\cdot)}^n \implies \mathbb{G}_{(\cdot)}$, ali da bi mogli tako nešto i formalno ustvrditi potrebna su detaljnija teorijskih razmatranja koja ćemo obraditi u poglavlju 2.3.

2.2 Djelitelji slučajnih brojeva

Sada se okrećemo ranije najavljenom drugom primjeru koji, iako naizgled nepovezan s ciklusima i permutacijama, pokazuje mnogo sličnih svojstava.

Promatrat ćemo rastav prirodnog broja na proste faktore. To nam daje primjer drugog vjerojatnosnog prostora s kojim ćemo raditi, a definiran je na skupu svih prirodnih brojeva manjih ili jednakih nekom $n \in \mathbb{N}$. Kao i u primjeru permutacija neka je jednaka vjerojatnost odabira svakog broja $N_n \leq n$. Za dani slučajni broj N_n neka su (P_1, P_2, \dots) različiti prosti faktori od N_n takvi da vrijedi $N_n = P_1^{\alpha_1} \cdot P_2^{\alpha_2} \cdot \dots$, gdje je definirani niz konačan. Označimo s $T_n := P_1 \cdot P_2 \cdot \dots$. Uočimo u ovom trenutku da smo u prethodnom primjeru imali

prirodan poredak među promatranim ciklusima. S indeksom 1 označili smo onaj koji u sebi sadrži broj 1, s 2 onaj koji sadrži najmanji prirodan broj koji se već nije pojavio u prvom ciklusu i tako dalje. Također uočimo da kod prostih djelitelja takvog prirodnog poretka nema, pa ćemo zato u ovom primjeru odmah poredati promatrane elemente od većeg prema manjem. Što smo i u prethodnom primjeru u konačnici napravili. Ponovno je moguće dani niz interpretirati i kao beskonačan ukoliko stavimo da je $P_k = 1$ za k veći od broja različitih prostih djelitelja broja N_n . Primijetimo da trivijalno vrijede sljedeće činjenice $\mathbb{P}(\{N_n = k\}) = \frac{1}{n}$ za $k = 1, 2, \dots, n$ i $\sum_{k=1}^{\infty} \log(P_k) = \log(P_1 \cdot P_2 \cdot \dots) = \log(T_n)$. Odnosno, slično kao i u prethodnom primjeru, vrijedi $\sum_{k=1}^{\infty} \frac{\log(P_k)}{\log(T_n)} = 1$ i $\frac{\log(P_k)}{\log(T_n)} \geq 0$ za $k \geq 1$.

Primjer 2.12. Promotrimo kako uvedeni pojmovi izgledaju za $N = 3150$. Tada je uz gornje oznake

$$\begin{aligned}
 p_1 &= 2, p_2 = 3, p_3 = 5, p_4 = 5, p_5 = 7 \\
 T_n &= 1050
 \end{aligned}$$

Iz čega dobivamo

$$\sum_k \frac{\log(p_k)}{\log(T_n)} = \frac{\log(2) + \log(3) + \log(3) + \log(5) + \log(5) + \log(7)}{\log(1050)} = \frac{\log(1050)}{\log(1050)} = 1.$$

Neka je vektor $\mathbf{X}^n = (X_1^n, X_2^n, \dots)$ definiran pomoću relacije $X_i = \frac{\log(P_i)}{\log(T_n)}$. Promotrimo vektor $\hat{\mathbf{X}}^n = (\hat{X}_1^n, \hat{X}_2^n, \dots)$ koji se formalno definira kao u napomeni 1.19 te zadovoljava

$$\begin{aligned}
 \mathbb{P}[\hat{X}_1^n = X_1^n] &= X_1^n \text{ za svaki } i \\
 \mathbb{P}[\hat{X}_2^n = X_2^n \mid \hat{X}_1^n = X_1^n] &= \frac{X_2^n}{1 - X_1^n} \text{ za svaki } j \neq i \\
 \mathbb{P}[\hat{X}_3^n = X_3^n \mid \hat{X}_1^n = X_1^n, \hat{X}_2^n = X_2^n] &= \frac{X_3^n}{1 - X_1^n - X_2^n} \text{ za svaki } k \neq i, j \\
 &\vdots
 \end{aligned}$$

Sada uvodimo nove slučajne varijable D_1^n, D_2^n, \dots pomoću relacije

$$D_k^n := \exp(\hat{X}_k^n \cdot \log(T_n)).$$

Budući da \hat{X}_1^n s vjerojatnošću X_i poprima vrijednost $X_i = \frac{\log(P_i)}{\log(T_n)}$ onda D_1^n s vjerojatnošću $X_i = \frac{\log(P_i)}{\log(T_n)}$ poprima vrijednost

$$\exp\left(\frac{\log(P_i)}{\log(T_n)} \cdot \log(T_n)\right) = \exp(\log(P_i)) = P_i.$$

Slične tvrdnje vrijede za ostale indekse uz uvjet da uzimamo uvjetnu vjerojatnost. Uočimo još da su slučajne varijable $\{P_i : i \in \mathbb{N}\}$ sortirane silazno dok za $\{D_i^n : i \in \mathbb{N}\}$ to ne vrijedi.

Ukoliko je $N_n = p_1 \cdot \dots \cdot p_r$ gdje su p_1, \dots, p_r međusobno različiti prosti brojevi onda je

$$\begin{aligned} \mathbb{P}[D_1^n = p_1, \dots, D_r^n = p_r \mid N_n] &= \frac{\log(p_1)}{\log(T_n)} \cdot \frac{\frac{\log(p_2)}{\log(T_n)}}{1 - \frac{\log(p_1)}{\log(T_n)}} \cdot \dots \cdot \frac{\frac{\log(p_r)}{\log(T_n)}}{1 - \frac{\log(p_1)}{\log(T_n)} - \dots - \frac{\log(p_{r-1})}{\log(T_n)}} = \\ &= \frac{\log(p_1)}{\log(T_n)} \cdot \frac{\log(p_2)}{\log(T_n) - \log(p_1)} \cdot \dots \cdot \frac{\log(p_r)}{\log(T_n) - \log(p_1) - \dots - \log(p_{r-1})} = \\ &= \frac{\log(p_1)}{\log(T_n)} \cdot \frac{\log(p_2)}{\log\left(\frac{T_n}{p_1}\right)} \cdot \dots \cdot \frac{\log(p_r)}{\log\left(\frac{T_n}{p_1 \cdot \dots \cdot p_{r-1}}\right)}. \end{aligned}$$

Definirajmo slučajne varijable $V_1^n := n$ i $V_i^n := \frac{n}{D_1^n \cdot \dots \cdot D_{i-1}^n}$ za $i > 1$ te slučajne varijable

$$U_i^n := \frac{\log(D_i^n)}{\log(V_i^n)} \text{ za } i \in \mathbb{N}.$$

Kao što notacija daje naslutiti pokazat ćemo da za slučajne varijable U_i^n vrijedi

$$(U_1^n, U_2^n, \dots) \Rightarrow (U_1, U_2, \dots),$$

gdje su U_i međusobno nezavisne slučajne varijable uniformno distribuirane na $[0, 1]$.

Raspišimo U_i^n za $i = 1, 2, 3$.

$$\begin{aligned} U_1^n &= \frac{\log(D_1^n)}{\log(V_1^n)} = \frac{\log(D_1^n)}{\log(n)}, \\ U_2^n &= \frac{\log(D_2^n)}{\log(V_2^n)} = \frac{\log(D_2^n)}{\log\left(\frac{n}{D_1^n}\right)} = \frac{\log(D_2^n)}{\log(n) - \log(D_1^n)} = \frac{\frac{\log(D_2^n)}{\log(n)}}{1 - \frac{\log(D_1^n)}{\log(n)}}, \\ U_3^n &= \frac{\log(D_3^n)}{\log(V_3^n)} = \frac{\log(D_3^n)}{\log\left(\frac{n}{D_1^n D_2^n}\right)} = \frac{\log(D_3^n)}{\log(n) - \log(D_1^n) - \log(D_2^n)} = \\ &= \frac{\frac{\log(D_3^n)}{\log(n)}}{1 - \frac{\log(D_1^n)}{\log(n)} - \frac{\log(D_2^n)}{\log(n)}}. \end{aligned}$$

Sada uočavamo još jednu sličnost s prethodnim primjerom ciklusa. Jedini problem je što su nas u prethodnom primjeru zanimala upravo slučajne varijable L_i , ali sada nas ne zanimaju slučajne varijable $\frac{\log(D_i^n)}{\log(n)}$ već slučajne varijable $\frac{\log(D_i^n)}{\log(T_n)}$. Pokažimo da je ta razlika zapravo nebitna. Odnosno formalno dokažimo sljedeću propoziciju.

Propozicija 2.13. Za niz slučajnih varijabli $(T_n)_{n \in \mathbb{N}}$ vrijedi

$$\frac{\log(T_n)}{\log(n)} \Rightarrow 1 \text{ kada } n \rightarrow \infty.$$

Dokaz. U dokazu koristimo tvrdnje iz teorije brojeva koje su ranije pokazane u teoremu 1.23, odnosno teoremu 1.24.

$$\sum_{p \leq x} \frac{\log(p)}{p} = \log(x) + O(1), \quad x \rightarrow \infty, \quad (2.1)$$

$$\sum_{p \leq x} \log(p) = O(x), \quad x \rightarrow \infty. \quad (2.2)$$

Označimo s $t(m)$ produkt svih različitih prostih faktora prirodnog broja m . Tada je

$$t(N_n) = T_n. \quad (2.3)$$

Budući da vrijedi

$$\begin{aligned} \mathbb{E}[\log(T_n)] &= \sum_{m=1}^n \frac{\log(t(m))}{n} = \sum_{m=1}^n \frac{1}{n} \sum_{p|m} \log(p) = \\ &= \sum_{p \leq n} \frac{1}{n} \left\lfloor \frac{n}{p} \right\rfloor \log(p) \geq \sum_{p \leq n} \frac{\log(p)}{p} - \frac{1}{n} \sum_{p \leq n} \log(p) \end{aligned}$$

koristeći (2.1) i (2.2) imamo da je $\mathbb{E}[\log(n) - \log(T_n)] \leq \log(n) - \log(n) - O(1) + \frac{1}{n}O(n) = O(1)$ kada $n \rightarrow \infty$. Po definiciji 1.21 postoji konstanta M takva da za dovoljno veliki n vrijedi $\mathbb{E}[\log(n) - \log(T_n)] \leq M$. Budući da je $\log(n) > \log(T_n)$ dobivamo

$$0 \leq \mathbb{E} \left[1 - \frac{\log(T_n)}{\log(n)} \right] \leq \frac{M}{\log(n)}.$$

Po teoremu o sendviču slijedi $\mathbb{E} \left[\left| 1 - \frac{\log(T_n)}{\log(n)} \right| \right] = \mathbb{E} \left[1 - \frac{\log(T_n)}{\log(n)} \right] \rightarrow 0$ kada $n \rightarrow \infty$. Budući da konvergencija u L^1 povlači konvergenciju po distribuciji zbog korolara 1.12 dobivamo traženu tvrdnju. \square

Kao i u primjeru ciklusa promatrat ćemo slučajne varijable

$$G_1 := U_1,$$

$$G_i := (1 - U_1)(1 - U_2) \dots (1 - U_{i-1}) U_i, \text{ za } i = 2, 3, \dots$$

U skladu s prethodnom propozicijom ukoliko pokažemo da slučajne varijable U^n uistinu konvergiraju ka uniformnima istim računom kao i u slučaju ciklusa slijedit će da vrijedi

$$\left(\frac{\log(D_1^n)}{\log(T_n)}, \frac{\log(D_2^n)}{\log(T_n)}, \dots \right) \Rightarrow (G_1, G_2, \dots).$$

Pokažimo sada da je uporaba oznaka U_i^n bila opravdana, odnosno dokažimo sljedeći teorem.

Teorem 2.14. *Za slučajne varijable U_i^n vrijedi*

$$(U_1^n, U_2^n, \dots) \Rightarrow (U_1, U_2, \dots),$$

gdje su U_i međusobno nezavisne slučajne varijable uniformno distribuirane na $[0, 1]$.

Dokaz. Po korolaru 1.15 dovoljno je dokazati da za svaki $r \in \mathbb{N}$ i $0 < a_i < b_i < 1$ vrijedi

$$\liminf_n \mathbb{P}[a_i < U_i^n \leq b_i, i \leq r] \geq \prod_{i=1}^r (b_i - a_i).$$

Budući da je $U_i^n := \frac{\log(D_i^n)}{\log(V_i^n)}$, imamo

$$\begin{aligned} \mathbb{P}[a_i < U_i^n \leq b_i, i \leq r] &= \mathbb{P}\left[a_i < \frac{\log(D_i^n)}{\log(V_i^n)} \leq b_i, i \leq r\right] = \\ &= \mathbb{P}\left[\log((V_i^n)^{a_i}) < \log(D_i^n) \leq \log((V_i^n)^{b_i}), i \leq r\right] = \\ &= \mathbb{P}\left[(V_i^n)^{a_i} < D_i^n \leq (V_i^n)^{b_i}, i \leq r\right] = \sum_{m=1}^n \mathbb{P}\left[N_n = m, (V_i^n)^{a_i} < D_i^n \leq (V_i^n)^{b_i}, i \leq r\right] = \\ &= \sum_{m=1}^n \sum_{\substack{(v_i)^{a_i} < p_i \leq (v_i)^{b_i} \\ p_1, \dots, p_r | m \\ p_i \neq p_j}} \mathbb{P}[N_n = m, D_i^n = p_i, i \leq r] \end{aligned}$$

gdje je $v_i = n(p_1 \dots p_{i-1})^{-1}$. Ponovno označimo s $t(m)$ produkt svih različitih prostih faktora prirodnog broja m .

$$\begin{aligned} \mathbb{P}[N_n = m, D_i^n = p_i, i \leq r] &= \mathbb{P}[D_1^n = p_1, D_2^n = p_2, \dots, D_r^n = p_r \mid N_n = m] \cdot \mathbb{P}(N_n = m) = \\ &= \frac{\log(p_1)}{\log(t(m))} \cdot \frac{\log(p_2)}{\log\left(\frac{t(m)}{p_1}\right)} \cdot \dots \cdot \frac{\log(p_r)}{\log\left(\frac{t(m)}{p_1 \dots p_{r-1}}\right)} \cdot \frac{1}{n}. \end{aligned}$$

Budući da iz $m \leq n$ slijedi $t(m) \leq n$ dobivamo ogradu

$$\mathbb{P}[a_i < U_i^n \leq b_i, i \leq r] \geq \sum_{m=1}^n \frac{1}{n} \sum_{\substack{(v_i)^{a_i} < p_i \leq (v_i)^{b_i} \\ p_1 \dots p_r \mid m \\ p_i \neq p_j}} \prod_{i=1}^r \frac{\log(p_i)}{\log\left(\frac{n}{p_1 \dots p_{r-1}}\right)}. \quad (2.4)$$

Budući da su svi pribrojnici s desne strane od (2.4) pozitivni ukoliko smanjimo broj elemenata p_i po kojima sumiramo dobit ćemo izraz manji od desne strane u (2.4). Zato definiramo $v_i(\varepsilon) := n\varepsilon(p_1 \dots p_{i-1})^{-1} = v_i \cdot \varepsilon$ za $i \leq r$ i $0 < \varepsilon < 1$.

Za $i = r$ tada $p_r \leq (v_r(\varepsilon))^{b_r}$ i činjenica da je $b_r \in \langle 0, 1 \rangle$ implicira da je $p_r \leq p_r^{\frac{1}{b_r}} \leq n\varepsilon(p_1 \dots p_{r-1})^{-1}$ odnosno

$$n\varepsilon \geq p_1 \cdot p_2 \dots p_r. \quad (2.5)$$

Nadalje vrijedi

$$\left\lfloor \frac{n}{p_1 \dots p_r} \right\rfloor \geq \frac{n}{p_1 \dots p_r} - 1 = \frac{n - p_1 \dots p_r}{p_1 \dots p_r} \stackrel{(2.5)}{\geq} \frac{n - n\varepsilon}{p_1 \dots p_r} = \frac{n(1 - \varepsilon)}{p_1 \dots p_r}. \quad (2.6)$$

Iz (2.4) i (2.6) slijedi

$$\begin{aligned} \mathbb{P}[a_i < U_i^n \leq b_i, i \leq r] &\stackrel{(2.4)}{\geq} \sum_{m=1}^n \frac{1}{n} \sum_{\substack{(v_i)^{a_i} < p_i \leq (v_i(\varepsilon))^{b_i} \\ p_1 \dots p_r \mid m \\ p_i \neq p_j}} \prod_{i=1}^r \frac{\log(p_i)}{\log\left(\frac{n}{p_1 \dots p_{r-1}}\right)} = \\ &= \frac{1}{n} \sum_{\substack{(v_i)^{a_i} < p_i \leq (v_i(\varepsilon))^{b_i} \\ p_1 \dots p_r \leq n \\ p_i \neq p_j}} \left\lfloor \frac{n}{p_1 \dots p_r} \right\rfloor \prod_{i=1}^r \frac{\log(p_i)}{\log\left(\frac{n}{p_1 \dots p_{r-1}}\right)} \stackrel{(2.6)}{\geq} \\ &= \frac{1}{n} \sum_{\substack{(v_i)^{a_i} < p_i \leq (v_i(\varepsilon))^{b_i} \\ p_i \neq p_j}} \frac{n(1 - \varepsilon)}{p_1 \dots p_r} \prod_{i=1}^r \frac{\log(p_i)}{\log\left(\frac{n}{p_1 \dots p_{r-1}}\right)} = (1 - \varepsilon) \sum_{\substack{(v_i)^{a_i} < p_i \leq (v_i(\varepsilon))^{b_i} \\ p_i \neq p_j}} \prod_{i=1}^r \frac{\log(p_i)}{p_i \cdot \log\left(\frac{n}{p_1 \dots p_{r-1}}\right)} \end{aligned}$$

Neka je $n_i := n(p_1 \dots p_{i-1})^{-1}$ za $i = 1, 2, \dots, r$. Za n_i vrijedi $n = n_1 > n_2 > \dots > n_r$ i

$$\frac{n_i}{n_i + 1} = \frac{\frac{n}{p_1 \dots p_{i-1}}}{\frac{n}{p_1 \dots p_i}} = p_i.$$

Za $p_i \leq (v_i(\varepsilon))^{b_i} = (n\varepsilon(p_1 \dots p_{i-1})^{-1})^{b_i} \leq n_i^{b_i}$ vrijedi nejednakost $n_{i+1} \geq n_i^{1-b_i}$. Ukoliko je promijenimo redom za $i = r, r-1, \dots, 1$ dobivamo da je $n_r \geq n_1^{(1-b_{r-1})(1-b_{r-2})\dots(1-b_1)} = n^\delta$, gdje smo s δ označili $\prod_{i=1}^{r-1} (1-b_i) > 0$.

Ponovno ćemo koristiti rezultat iz teorije brojeva, odnosno teorem 1.23

$$\sum_{p \leq x} \frac{\log(p)}{p} = \log(x) + O(1), \quad x \rightarrow \infty.$$

Odnosno u našem slučaju

$$\begin{aligned} \sum_{p \leq (\varepsilon x)^{b_i}} \frac{\log(p)}{p} &= b_i \cdot \log(\varepsilon) + b_i \cdot \log(x) + O(1), \\ \sum_{p \leq (x)^{a_i}} \frac{\log(p)}{p} &= a_i \cdot \log(x) + O(1). \end{aligned}$$

Iz čega slijedi

$$\begin{aligned} \sum_{(x)^{a_i} < p \leq (\varepsilon x)^{b_i}} \frac{\log(p)}{p} &= b_i \cdot \log(\varepsilon) + b_i \cdot \log(x) - a_i \cdot \log(x) + O(1) \implies \\ \sum_{(x)^{a_i} < p \leq (\varepsilon x)^{b_i}} \frac{\log(p)}{p \log(x)} &= (b_i - a_i) + b_i \frac{\log(\varepsilon)}{\log(x)} + \frac{O(1)}{\log(x)}. \end{aligned}$$

Zbog čega postoji x_ε takav da za svaki $x > x_\varepsilon$ vrijedi

$$\sum_{(x)^{a_i} < p \leq (\varepsilon x)^{b_i}} \frac{\log(p)}{p \log(x)} \geq (b_i - a_i)(1 - \varepsilon).$$

Budući da je $n_i \geq n^\delta$ za $n \geq x_\varepsilon^{\frac{1}{\delta}}$ je $n_i \geq x_\varepsilon$, odnosno upravo dokazana nejednakost vrijedi za svaki $n_i, i = 1, 2, \dots, r$.

$$\mathbb{P}[a_i < U_i^n \leq b_i, i \leq r] \geq (1 - \varepsilon) \sum_{\substack{(v_i)^{a_i} < p_i \leq (v_i(\varepsilon))^{b_i} \\ p_i \neq p_j}} \prod_{i=1}^r \frac{\log(p_i)}{p_i \cdot \log(n_i)} =$$

$$\begin{aligned}
 & (1 - \varepsilon) \sum_{\substack{(v_i)^{a_i} < p_i \leq (v_i(\varepsilon))^{b_i} \\ p_i \neq p_j \\ i=1, \dots, r-1}} \prod_{i=1}^{r-1} \frac{\log(p_i)}{p_i \cdot \log(n_i)} \sum_{\substack{(v_i)^{a_r} < p_r \leq (v_r(\varepsilon))^{b_r} \\ p_i \neq p_r}} \frac{\log(p_r)}{p_r \cdot \log(n_r)} \geq \\
 & (1 - \varepsilon) \sum_{\substack{(v_i)^{a_i} < p_i \leq (v_i(\varepsilon))^{b_i} \\ p_i \neq p_j \\ i=1, \dots, r-1}} \prod_{i=1}^{r-1} \frac{\log(p_i)}{p_i \cdot \log(n_i)} \left(\sum_{\substack{(v_i)^{a_r} < p_r \leq (v_r(\varepsilon))^{b_r} \\ p_i \neq p_r}} \frac{\log(p_r)}{p_r \cdot \log(n_r)} - \sum_{i=1}^{r-1} \frac{\log(p_i)}{p_i \cdot \log(n_i)} \right) \geq \\
 & (1 - \varepsilon) \sum_{\substack{(v_i)^{a_i} < p_i \leq (v_i(\varepsilon))^{b_i} \\ p_i \neq p_j \\ i=1, \dots, r-1}} \prod_{i=1}^{r-1} \frac{\log(p_i)}{p_i \cdot \log(n_i)} \left((b_r - a_r)(1 - \varepsilon) - r \cdot \frac{1}{\log(n_r)} \right) \geq \\
 & (1 - \varepsilon) \left((b_r - a_r)(1 - \varepsilon) - r \cdot \frac{1}{\log(n^\delta)} \right) \sum_{\substack{(v_i)^{a_i} < p_i \leq (v_i(\varepsilon))^{b_i} \\ p_i \neq p_j \\ i=1, \dots, r-1}} \prod_{i=1}^{r-1} \frac{\log(p_i)}{p_i \cdot \log(n_i)} \geq \dots \geq \\
 & (1 - \varepsilon) \prod_{i=1}^r \left((b_i - a_i)(1 - \varepsilon) - \frac{r}{\log(n^\delta)} \right).
 \end{aligned}$$

Iz čega slijedi

$$\liminf_n \mathbb{P} [a_i < U_i^n \leq b_i, i \leq r] \geq \prod_{i=1}^r (b_i - a_i).$$

□

Iz ovog teorema i ranijih napomena izravno dobivamo tvrdnju sljedećeg korolara.

Korolar 2.15.

$$\left(\frac{\log(D_1^n)}{\log(T_n)}, \frac{\log(D_2^n)}{\log(T_n)}, \dots \right) \implies (G_1, G_2, \dots).$$

Ono što ponovno kao i u primjeru ciklusa preostaje za pokazati jest što se događa sa

$$\left(\frac{\log(P_1)}{\log(T_n)}, \frac{\log(P_2)}{\log(T_n)}, \dots \right),$$

odnosno hoće li se sortiranjem vektora

$$\left(\frac{\log(D_1^n)}{\log(T_n)}, \frac{\log(D_2^n)}{\log(T_n)}, \dots \right)$$

očuvati konvergencija. Budući da su oba primjera svedena na rješavanje istog problema sada je razumno preći na općeniti model koji će obuhvatiti relevantna svojstva oba primjera.

2.3 Prostor Δ

Iskažimo i formalno definiciju čija se važnost dala naslutiti iz dosadašnjih primjera, a pokazat će se centralnim prostorom u ovom poglavlju.

Definicija 2.16. Prostor

$$\Delta := \{\mathbf{x} \in \mathbb{R}^\infty : x_1, x_2, \dots \geq 0, x_1 + x_2 + \dots = 1\}$$

promatramo kao potprostor od \mathbb{R}^∞ od kojega nasljeđuje topologiju konvergencije po koordinatama.

Topologija konvergencije po koordinatama se tako naziva jer niz $(\mathbf{x}_n)_{n \in \mathbb{N}} \in \Delta$ konvergira k $\mathbf{x} \in \mathbb{R}^\infty$ ako i samo ako niz $(x_k^n)_{n \in \mathbb{N}}$ konvergira k x_k gdje smo s x_k^n odnosno x_k označili k -tu koordinatu vektora \mathbf{x}_n odnosno \mathbf{x} .

Na prostoru Δ definirat ćemo funkciju rangiranja koja uzima proizvoljni element iz prostora Δ i njegove elemente sortira od većeg prema manjem. S obzirom na razmatranja u prethodna dva primjera jasno je da su svojstva ove funkcije ključna u formalnom dokazivanju tvrdnji koje smo ranije naslutili.

Definicija 2.17. Funkciju rangiranja $\rho : \Delta \rightarrow \Delta$ definiramo pomoću relacije

$$\rho(x_1, x_2, \dots) = (y_1, y_2, \dots),$$

za koju postoji permutacija $\tau : \mathbb{N} \rightarrow \mathbb{N}$ takva da je $(x_1, x_2, \dots) = (y_{\tau(1)}, y_{\tau(2)}, \dots)$ i vrijedi $y_1 \geq y_2 \geq \dots \geq y_k \geq \dots$

Napomena 2.18. Funkcija ρ je dobro definirana.

Iz $x_1 + x_2 + \dots = 1$ i $x_1, x_2, \dots \geq 0$ slijedi da postoji $\max_{k \in \mathbb{N}} x_k$ i da se taj maksimum postiže za neki $n < \infty$ indeksa k . Iz toga slijedi da je dobro definirano $y_1 = \dots = y_n = \max_{k \in \mathbb{N}} x_k$. Nastavimo li ovaj postupak dobivamo egzistenciju vektora (y_1, y_2, \dots) , a jedinstvenost je očita iz definicije.

Teorem 2.19. Funkcija rangiranja je neprekidna.

Dokaz. Budući da je prostor Δ metrički kako bi dokazali da je funkcija ρ neprekidna dovoljno je pokazati da ukoliko niz $(\mathbf{x}_n)_{n \in \mathbb{N}} \in \Delta$ konvergira ka $\mathbf{x} \in \Delta$ onda niz $(\rho(\mathbf{x}_n))_{n \in \mathbb{N}} \in \Delta$ konvergira k $\rho(\mathbf{x}) \in \Delta$. Po definiciji prostora Δ ove nizovne konvergencije zapravo znače konvergenciju po svakoj od pojedinih koordinata. Iz tog razloga koordinate možemo permutirati po želji pa zato nije smanjenje općenitosti ako pretpostavimo da za vektor $\mathbf{x} = (x_1, x_2, \dots) \in \Delta$ vrijedi $x_1 \geq x_2 \geq \dots$. Odnosno da je $\rho(\mathbf{x}) = \mathbf{x}$.

Dakle pretpostavimo da za svaki $m \in \mathbb{N}$ vrijedi $x_m^n \rightarrow x_m$ kada $n \rightarrow \infty$ i dokažimo da tada za svaki $l \in \mathbb{N}$ vrijedi $\rho(\mathbf{x}_n)_l \rightarrow \rho(\mathbf{x})_l = x_l$ kada $n \rightarrow \infty$.

Dokažimo prvo tvrdnju za $l = 1$. Iz definicije funkcije ρ znamo da je tada $\rho(\mathbf{x}_n)_1 = \max_{k \in \mathbb{N}} x_k^n$. Ukoliko je $x_1 = x_2 = \dots = x_L > x_{L+1}$ pokažimo da za dovoljno velik n vrijedi $\rho(\mathbf{x}_n)_1 = \max_{k \leq L} x_k^n$. Neka je $\varepsilon > 0$, budući da je $\sum_{i=1}^{\infty} x_i = 1$, postoji $K_0 \in \mathbb{N}$ takav da je $\sum_{i=K_0}^{\infty} x_i < \varepsilon$. Neka je $N_0 \in \mathbb{N}$ takav da za svaki $n \geq N_0$ i za svaki $i < K_0$ vrijedi $|x_i^n - x_i| < \frac{\varepsilon}{K_0}$ iz čega slijedi

$$\sum_{i=K_0}^{\infty} x_i^n = 1 - \sum_{i=1}^{K_0-1} x_i^n < 1 - \sum_{i=1}^{K_0-1} x_i - \frac{\varepsilon}{K_0} = K_0 \cdot \frac{\varepsilon}{K_0} + 1 - \sum_{i=1}^{K_0-1} x_i = \varepsilon + \sum_{i=K_0}^{\infty} x_i < 2\varepsilon, \forall n \geq N_0.$$

Iz činjenice da je $x_1 > 0$ i da $x_1^n \rightarrow x_1$ kada $n \rightarrow \infty$ slijedi da postoji $N_1 \in \mathbb{N}$ takav da za svaki $n > N_1$ vrijedi $|x_1^n - x_1| < \frac{x_1}{3}$. Tada ako u gornjoj nejednakosti uzmemo $\varepsilon = \frac{x_1}{3}$ za $n > \max\{N_0, N_1\}$ vrijedi

$$x_1^n > x_1 - \frac{x_1}{3} = 2 \cdot \frac{x_1}{3} > \sum_{i=K_0}^{\infty} x_i^n \geq x_k^n, \forall k \geq K_0.$$

Odnosno za $n > \max\{N_0, N_1\}$ je $\rho(\mathbf{x}_n)_1 = \max_{k < K_0} x_k^n$. Sada ovu ocjenu možemo i poboljšati, jer budući da je $x_1 > x_{L+1} \geq x_{L+2} \geq \dots \geq x_{K_0}$ onda postoji N_2 takav da za $n > N_2$ vrijedi $x_1^n > x_{L+1}^n, x_1^n > x_{L+2}^n, \dots, x_1^n > x_{K_0}^n$.

Budući da za svaki $k \leq L$ vrijedi $x_k^n \rightarrow x_k = x_1$ kada $n \rightarrow \infty$ iz gore pokazanog slijedi i da $\rho(\mathbf{x}_n)_1 = \max_{k \leq L} x_k^n \rightarrow x_1 = \rho(\mathbf{x})_1$ kada $n \rightarrow \infty$.

Uočimo da promatranjem l -te najveće vrijednosti umjesto maksimuma isti argumenti povlače da tvrdnja vrijedi za svaki $l \leq L$. Odnosno možemo zaključiti da za dovoljno veliki n i za svaki $l \leq L$ vrijedi $\rho(\mathbf{x}_n)_l \in \{x_1^n, \dots, x_L^n\}$.

Pokažimo da tvrdnja vrijedi i za $l = L+1$, a onda će zbog analognih argumenata tvrdnja vrijediti i općenito primjenom matematičke indukcije.

Ukoliko je $x_{L+1} = 0$ onda je $x_i = 0$ za svaki $i \geq L+1$. Tada za svaki $\varepsilon > 0$ postoji $N_3 \in \mathbb{N}$ takav da za svaki $n \geq N_3$ i za svaki $i \leq L$ vrijedi $|x_i^n - x_i| < \frac{\varepsilon}{L+1}$ iz čega slijedi

$$\begin{aligned} \sum_{i=L+1}^{\infty} x_i^n &= 1 - \sum_{i=1}^L x_i^n < 1 - \sum_{i=1}^L \left(x_i - \frac{\varepsilon}{L+1}\right) = \\ &= (L+1) \cdot \frac{\varepsilon}{L+1} + 1 - \sum_{i=1}^L x_i = \varepsilon + \sum_{i=L+1}^{\infty} x_i = \varepsilon, \forall n \geq N_3. \end{aligned}$$

Budući da smo ranije već dokazali da za dovoljno veliki n vrijedi $\rho(\mathbf{x}_n)_l \in \{x_1^n, \dots, x_L^n\}$ sada možemo zaključiti da za proizvoljni $\varepsilon > 0$ postoji $N_4 \in \mathbb{N}$ takav da za svaki $n \geq N_4$ vrijedi $\rho(\mathbf{x}_n)_{L+1} \leq \sum_{i=L+1}^{\infty} x_i^n < \varepsilon$. Odnosno da $\rho(\mathbf{x}_n)_{L+1} \rightarrow 0 = \rho(\mathbf{x})_{L+1}$ kada $n \rightarrow \infty$.

Preostaje pokazati da tvrdnja teorema vrijedi i u slučaju $x_{L+1} \neq 0$.

Ponovno istim argumentima kao i ranije dobivamo da za svaki $\varepsilon > 0$ postoji $K_1, N_5 \in \mathbb{N}$ takav da je $\sum_{i=K_1}^{\infty} x_i < \varepsilon$ i za svaki $n \geq N_5$ je $\sum_{i=K_1}^{\infty} x_i^n < 2\varepsilon$. Iz činjenice da je $x_{L+1} > 0$ i da $x_{L+1}^n \rightarrow x_{L+1}$ kada $n \rightarrow \infty$ slijedi da postoji $N_6 \in \mathbb{N}$ takav da za svaki $n > N_6$ vrijedi $|x_{L+1}^n - x_{L+1}| < \frac{x_{L+1}}{3}$. Tada ako u gornjoj nejednakosti uzmemo $\varepsilon = \frac{x_{L+1}}{3}$ za $n > \max\{N_5, N_6\}$ vrijedi

$$x_{L+1}^n > x_{L+1} - \frac{x_{L+1}}{3} = 2 \cdot \frac{x_{L+1}}{3} > \sum_{i=K_1}^{\infty} x_i^n \geq x_k^n, \forall k \geq K_1.$$

Odnosno ponovno koristeći činjenicu da za dovoljno veliki n vrijedi $\rho(\mathbf{x}_n)_l \in \{x_1^n, \dots, x_L^n\}$ dobivamo da je $\rho(\mathbf{x}_n)_{L+1} = \max_{L < k < K_1} x_k^n$. Sada ovu ocjenu možemo i poboljšati, jer uz oznake $x_{L+1} = \dots = x_S > x_{S+1} \geq x_{S+2} \geq \dots \geq x_K$ postoji N_7 takav da za $n > N_7$ vrijedi $x_{L+1}^n > x_{S+1}^n, x_{L+1}^n > x_{S+2}^n, \dots, x_{L+1}^n > x_{K_1}^n$. Iz čega slijedi da ponovno za dovoljno veliki n vrijedi $\rho(\mathbf{x}_n)_{L+1} = \max_{L < k \leq S} x_k^n$.

Budući da za svaki $L < k \leq S$ vrijedi $x_k^n \rightarrow x_k = x_{L+1}$ kada $n \rightarrow \infty$ iz gore pokazanog slijedi i da $\rho(\mathbf{x}_n)_{L+1} = \max_{L < k \leq S} x_k^n \rightarrow x_{L+1} = \rho(\mathbf{x})_{L+1}$ kada $n \rightarrow \infty$. Čime je dokaz gotov. \square

Sada smo u mogućnosti vratiti se ranije pretpostavljenoj tvrdnji u primjeru ciklusa. Ustvrdili smo da intuitivno očekujemo da će iz $L^n \implies G$ slijediti da isto vrijedi i za silazno rangirane cikluse. A sljedeći korolar nam tu činjenicu i formalno dokazuje.

Korolar 2.20.

$$\mathbb{L}_{(\cdot)}^n \implies \mathbb{G}_{(\cdot)}.$$

Dokaz. Od prije znamo da je $\mathbb{L}^n \in \Delta$ i da je $\mathbb{P}[\mathbb{G} \in \Delta] = 1$, pa ukoliko \mathbb{G} izmijenimo na skupu vjerojatnosti 0 dobivamo da su $\mathbb{L}^n, \mathbb{G} \in \Delta$. Kako je $\mathbb{L}_{(\cdot)}^n = \rho(\mathbb{L}^n)$ i $\mathbb{G}_{(\cdot)} = \rho(\mathbb{G})$, a funkcija ρ po teoremu 2.19 neprekidna onda iz teorema 1.16 i korolara 2.10 slijedi tražena tvrdnja. \square

Analogan rezultat dobivamo i u primjeru djelitelja

Korolar 2.21.

$$\left(\frac{\log(P_1)}{\log(T_n)}, \frac{\log(P_2)}{\log(T_n)}, \dots \right) \implies \mathbb{G}_{(\cdot)}.$$

Dokaz. Od prije znamo da je

$$\left(\frac{\log(P_1)}{\log(T_n)}, \frac{\log(P_2)}{\log(T_n)}, \dots \right) \in \Delta$$

i da je $\mathbb{P}[G \in \Delta] = 1$. Ako \mathbb{G} izmijenimo na skupu vjerojatnosti 0 dobivamo da je i $\mathbb{G} \in \Delta$. Kako je

$$\left(\frac{\log(P_1)}{\log(T_n)}, \frac{\log(P_2)}{\log(T_n)}, \dots \right) = \rho \left(\frac{\log(D_1^n)}{\log(T_n)}, \frac{\log(D_2^n)}{\log(T_n)}, \dots \right)$$

i $\mathbb{G}_{(\cdot)} = \rho(\mathbb{G})$, a funkcija ρ po teoremu 2.19 neprekidna onda iz teorema 1.16 i korolara 2.15 slijedi tražena tvrdnja. \square

Ovime je završena rasprava iz oba primjera i u nastavku se posvećujemo proučiti svojstva slučajnog procesa $\mathbb{G}_{(\cdot)}$ na prostoru Δ . Naravno da je to ključno u proučavanju graničnih svojstava za primjere ciklusa i djelitelja.

Poglavlje 3

Poisson-Dirichletova razdioba

U ovom poglavlju cilj nam je definirati, a potom opisati te dokazati glavna svojstva Poisson-Dirichletove razdiobe čija je važnost uočena u prethodnom poglavlju.

3.1 Povezanost Dirichletove i Poisson-Dirichletove razdiobe

Promatrat ćemo nešto općenitiji model nego u prethodnom poglavlju. Do sada smo za definiciju slučajne varijable G_k koristili uniformnu distribuciju, a sada uzimamo beta distribuciju definiranu u 1.6.

Definicija 3.1. Neka je dan vjerojatnosni prostor $(\Omega, \mathcal{F}, \mathbb{P})$ i na njemu nezavisne slučajne varijable B_1, B_2, \dots s beta- $(1, \theta)$, $\theta > 0$ distribucijom. Definiramo slučajne varijable G_1, G_2, \dots tako da vrijedi

$$\begin{aligned} G_1 &:= B_1, \\ G_k &:= (1 - B_1) \cdot \dots \cdot (1 - B_{k-1}) B_k, \quad k \geq 2. \end{aligned}$$

Neka je također $\mathbb{G} := (G_1, G_2, \dots)$.

Ova definicija se poklapa s ranijom u slučaju $\theta = 1$ budući da je uniformna distribucija beta- $(1, 1)$ distribucija, a formule kojom su definirane G_k su jednake. Za ovako definirane, odnosno poopćene slučajne varijable G_1, G_2, \dots također vrijedi $\sum_{k=1}^{\infty} G_k = 1 - \prod_{k=1}^{\infty} (1 - B_k)$, što se pokazuje isto kao i u posebnom slučaju iz prethodnog poglavlja. Također, budući da je $\sum_{k=1}^{\infty} G_k = 1$ g.s., ponovno možemo izmijeniti vrijednost od \mathbb{G} na skupu vjerojatnosti 0 tako da vrijedi $\mathbb{G} \in \Delta$. Zato je dobro definirano $\mathbb{G}_{(\cdot)} := \rho(\mathbb{G})$.

Definicija 3.2. Razdiobu slučajnog vektora $\mathbb{G}_{(\cdot)}$ na prostoru Δ nazivamo *Poisson-Dirichletova razdioba* s parametrom θ .

Napomena 3.3. U prethodnom poglavlju pokazali smo da slučajni vektori

$$\mathbb{L}_{(\cdot)}^n \text{ i } \left(\frac{\log(P_1)}{\log(T_n)}, \frac{\log(P_2)}{\log(T_n)}, \dots \right)$$

konvergiraju prema Poisson-Dirichletovoj razdiobi s parametrom 1.

Cilj ovog poglavlja je odrediti funkciju gustoće slučajnog vektora $\mathbb{G}_{(\cdot)}$ i njegovih komponenti. Pokazuje se da se to najjednostavnije postiže neizravnim putem. U tu svrhu promotrimo niz slučajnih vektora $\mathbb{B}^n = (B_1^n, B_2^n, \dots)$ na prostoru $[0, 1]^\infty$. Neka su komponente B_1^n, B_2^n, \dots međusobno nezavisne slučajne varijable, takve da B_i^n ima beta- $(\alpha + 1, (n - i)\alpha)$ distribuciju, gdje je koeficijent $\alpha > 0$. Tada je, po (1.6), funkcija gustoće od B_i^n dana formulom

$$h_i^n(x) = \frac{\Gamma((n - i + 1)\alpha + 1)}{\Gamma(\alpha + 1) \cdot \Gamma((n - i)\alpha)} x^\alpha (1 - x)^{(n-i)\alpha-1} \mathbb{1}_{[0,1]}(x). \quad (3.1)$$

Uočimo da za fiksni r kada pustimo $n \rightarrow \infty$ ukoliko dodatno zahtijevamo da vrijedi i $n\alpha \rightarrow \theta$ za svaki $1 \leq i \leq r$ imamo $B_i^n \Rightarrow B_i$, a onda i $(B_1^n, \dots, B_r^n) \Rightarrow (B_1, \dots, B_r)$.

Napomena 3.4. U nastavku kada koristimo parametar α podrazumijevat ćemo da ako $n \rightarrow \infty$ onda $n\alpha \rightarrow \theta$, pa i $\alpha \rightarrow 0$, drugim riječima α se mijenja u ovisnosti o n čak i kada to eksplicitno nije naglašeno. Formalno možemo definirati funkciju $\alpha : \mathbb{N} \rightarrow \mathbb{R}$ takvu da je $\lim_{n \rightarrow \infty} n\alpha(n) = \theta$ i slučajne varijable B_i^n koje imaju beta- $(\alpha(n) + 1, (n - i)\alpha(n))$ distribuciju. U praksi ćemo nastaviti pisati α , ali podrazumijevamo značenje tog parametra u skladu s ovom napomenom.

Neka je nadalje slučajni vektor $\mathbb{G}^n = (G_1^n, G_2^n, \dots)$ definiran s

$$\begin{aligned} G_1^n &:= B_1^n, \\ G_k^n &:= (1 - B_1^n) \cdot \dots \cdot (1 - B_{k-1}^n) B_k^n, \quad k \geq 2. \end{aligned}$$

Po teoremu 1.16 i iz prethodnih razmatranja slijedi $(G_1^n, \dots, G_r^n) \Rightarrow (G_1, \dots, G_r)$.

Propozicija 3.5. Funkcija gustoće slučajnog vektora $(G_1^n, G_2^n, \dots, G_r^n)$ je $f_r^n : \mathbb{R}^r \rightarrow \mathbb{R}$,

$$f_r^n(z_1, z_2, \dots, z_r) = h_1^n(z_1) h_2^n\left(\frac{z_2}{1 - z_1}\right) \dots h_r^n\left(\frac{z_r}{1 - z_1 - \dots - z_{r-1}}\right) \prod_{i=1}^{r-1} (1 - z_1 - \dots - z_i)^{-1}$$

za (z_1, z_2, \dots, z_r) takve da je $z_1, z_2, \dots, z_r > 0$ i $z_1 + z_2 + \dots + z_r < 1$, a $f_r^n \equiv 0$ inače.

Dokaz. Budući da su $B_1^n, B_2^n, \dots, B_r^n$ međusobno nezavisne slučajne varijable s funkcijama gustoće $h_1^n, h_2^n, \dots, h_r^n$ redom, slijedi da je funkcija gustoće slučajnog vektora $(B_1^n, B_2^n, \dots, B_r^n)$ funkcija $\psi : \mathbb{R}^r \rightarrow \mathbb{R}$,

$$\psi(z_1, z_2, \dots, z_r) := h_1^n(z_1) h_2^n(z_2) \dots h_r^n(z_r).$$

Budući da je $(G_1^n, G_2^n, \dots, G_r^n) = (B_1^n, (1 - B_1^n) B_2^n, \dots, (1 - B_1^n) \dots (1 - B_{r-1}^n) B_r^n)$ definiramo funkciju $g : \mathbb{R}^r \rightarrow \mathbb{R}^r$,

$$g(z_1, z_2, \dots, z_r) := (z_1, (1 - z_1) z_2, \dots, (1 - z_1) \dots (1 - z_{r-1}) z_r).$$

i prostor

$$T := \left\{ (s_1, \dots, s_r) \in \mathbb{R}^r : s_1, \dots, s_r > 0, \sum_{i=1}^r s_i < 1 \right\}$$

Vrijedi $(G_1^n, G_2^n, \dots, G_r^n) = g(B_1^n, B_2^n, \dots, B_r^n)$ i funkcija $g : \langle 0, 1 \rangle^r \rightarrow T$ je Borelova i bijekcija. Nadalje g^{-1} djeluje na vektor $(z_1, z_2, \dots, z_r) \in T$ na sljedeći način

$$g^{-1}(z_1, z_2, \dots, z_r) := \left(z_1, \frac{z_2}{1 - z_1}, \dots, \frac{z_r}{1 - z_1 - \dots - z_{r-1}} \right),$$

iz čega slijedi da je $g^{-1} \in C^1(T)$. Odredimo Dg^{-1} .

$$Dg^{-1} = \begin{pmatrix} 1 & \frac{z_2}{(1 - z_1)^2} & \frac{z_3}{(1 - z_1 - z_2)^2} & \cdots & \frac{z_r}{(1 - z_1 - \dots - z_{r-1})^2} \\ 0 & \frac{1}{(1 - z_1)} & \frac{z_3}{(1 - z_1 - z_2)^2} & \cdots & \frac{z_r}{(1 - z_1 - \dots - z_{r-1})^2} \\ 0 & 0 & \frac{1}{(1 - z_1 - z_2)} & \cdots & \frac{z_r}{(1 - z_1 - \dots - z_{r-1})^2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \frac{1}{(1 - z_1 - \dots - z_{r-1})} \end{pmatrix} \quad \begin{array}{l} \text{gornje trokutasta matrica} \\ = \end{array}$$

$$\prod_{i=1}^{r-1} (1 - z_1 - \dots - z_i)^{-1} \neq 0$$

za $(z_1, z_2, \dots, z_r) \in T$. Budući da su zadovoljeni uvjeti teorema 11.8. u Sarapa [4] slijedi da je funkcija gustoće slučajnog vektora $(G_1^n, G_2^n, \dots, G_r^n)$ dana s

$$\begin{aligned} f_r^n(z_1, z_2, \dots, z_r) &= \psi g^{-1}(z_1, z_2, \dots, z_r) |Dg^{-1}| \mathbb{1}_T(z_1, z_2, \dots, z_r) = \\ &= h_1^n(z_1) h_2^n\left(\frac{z_2}{1 - z_1}\right) \dots h_r^n\left(\frac{z_r}{1 - z_1 - \dots - z_{r-1}}\right) \prod_{i=1}^{r-1} (1 - z_1 - \dots - z_i)^{-1} \mathbb{1}_T(z_1, z_2, \dots, z_r). \end{aligned}$$

□

Iz prethodne propozicije, kada uvrstimo čemu je po (3.1) jednako h_1^n slijedi da je

$$\begin{aligned} \mathbb{P}[(G_1^n, \dots, G_r^n) \in H] &= \int_H h_1^n(z_1) h_2^n\left(\frac{z_2}{1-z_1}\right) \dots h_r^n\left(\frac{z_r}{1-z_1-\dots-z_{r-1}}\right) \cdot \\ &\prod_{i=1}^{r-1} (1-z_1-\dots-z_i)^{-1} \mathbb{1}_T(z_1, z_2, \dots, z_r) dz_1 \dots dz_r = \\ &\frac{1}{\Gamma(\alpha+1)^r} \prod_{i=1}^r \frac{\Gamma((n-i+1)\alpha+1)}{\Gamma((n-i)\alpha)} \int_H z_1^\alpha \dots z_r^\alpha \cdot (1-z_1-\dots-z_r)^{(n-r)\alpha-1} \cdot \\ &\prod_{i=1}^{r-1} (1-z_1-\dots-z_i)^{-1} \mathbb{1}_T(z_1, z_2, \dots, z_r) dz_1 \dots dz_r. \end{aligned}$$

Koristeći svojstvo $\Gamma(x+1) = x \cdot \Gamma(x)$ odredimo posebno vrijednost izraza

$$\begin{aligned} \frac{1}{\Gamma(\alpha+1)^r} \prod_{i=1}^r \frac{\Gamma((n-i+1)\alpha+1)}{\Gamma((n-i)\alpha)} &= \frac{1}{\alpha^r \Gamma(\alpha)^r} \prod_{i=1}^r \frac{((n-i+1)\alpha) \Gamma((n-i+1)\alpha)}{\Gamma((n-i)\alpha)} = \\ \frac{1}{\Gamma(\alpha)^r} \cdot \frac{\Gamma((n-1+1)\alpha)}{\Gamma((n-1)\alpha)} \cdot \frac{\Gamma((n-2+1)\alpha)}{\Gamma((n-2)\alpha)} \cdot \dots \cdot \frac{\Gamma((n-r+1)\alpha)}{\Gamma((n-r)\alpha)} &\prod_{i=1}^r (n-i+1) = \\ \frac{1}{\Gamma(\alpha)^r} \cdot \frac{\Gamma(n\alpha)}{\Gamma((n-r)\alpha)} \cdot (n)_r. \end{aligned}$$

Iz pokazanog slijedi

$$\begin{aligned} \mathbb{P}[(G_1^n, \dots, G_r^n) \in H] &= \frac{\Gamma(n\alpha)}{\Gamma(\alpha)^r \Gamma((n-r)\alpha)} \int_H (n)_r \cdot z_1^{\alpha-1} \dots z_r^{\alpha-1} \cdot (1-z_1-\dots-z_r)^{(n-r)\alpha-1} \cdot \\ z_1 \frac{z_2}{1-z_1} \dots \frac{z_r}{(1-z_1-\dots-z_{r-1})} &\mathbb{1}_T(z_1, z_2, \dots, z_r) dz_1 \dots dz_r. \end{aligned} \quad (3.2)$$

Za daljnje proučavanje svojstava Poisson-Dirichletove razdiobe okrenut ćemo se modelu koji proizlazi iz populacijske genetike. Pretpostavimo da se promatrana populacija sastoji od jedinki koje pripadaju jednom od n različitih tipova. Označimo s Z_i , $i = 1, 2, \dots, n$ relativne frekvencije i -tog tipa. Budući da je $Z_1 + \dots + Z_n = 1$ dovoljno je promatrati vektor (Z_1, \dots, Z_{n-1}) . Razdioba vektora (Z_1, \dots, Z_{n-1}) naziva se *Dirichletova razdioba*, a odgovarajuća funkcija gustoće dana je formulom

$$f_n(z_1, \dots, z_{n-1}) = \frac{\Gamma(n\alpha)}{\Gamma^n(\alpha)} z_1^{\alpha-1} \dots z_{n-1}^{\alpha-1} (1-z_1-\dots-z_{n-1})^{\alpha-1}, \quad (3.3)$$

gdje vektor (z_1, \dots, z_{n-1}) zadovoljava $z_1 + \dots + z_{n-1} < 1$ i $z_1, \dots, z_{n-1} > 0$.

Parametar α u zapisu predstavlja stopu mutacije po pojedinom tipu. Nas će zanimati slučaj u kojem se broj tipova beskonačno povećava, jer se svakom mutacijom u populaciji stvaraju novi tipovi, ali ukupna stopa mutacije za cijelu populaciju ostaje fiksna. Formalno promatrat ćemo što se događa kada $n \rightarrow \infty$, ali parametar $n\alpha$ ne teži u 0 već u pozitivnu konstantu. Uočimo da se ova situacija poklapa s ranije spomenutim slučajem od interesa kada je $n \rightarrow \infty$ i $n\alpha \rightarrow \theta$. Kao i u prethodnom poglavlju zanimat će nas sortirane, odnosno konkretno najveće relativne frekvencije $Z_{(1)}, Z_{(2)}, \dots, Z_{(r)}$.

Napomena 3.6. Uočimo da smo odlukom da promatramo samo prvih r slučajnih varijabli izgubili jedno od ranije bitnih svojstava da je suma elemenata u promatranom slučajnom vektoru 1.

Uvedimo zato formalno vektor $Z^n = (Z_1^n, Z_2^n, \dots)$ i neka je on slučajni element prostora Δ za koji je prvih $n-1$ komponenta definirano pomoću funkcije gustoće dane u (3.3), n -ta komponenta dana formulom $1 - Z_1^n - \dots - Z_{n-1}^n$, a za sve veće komponente vrijedi $Z_i^n = 0$. Označimo s $Z_{(\cdot)}^n := \rho Z^n$.

Za ovako definiran Z^n i fiksna r promatrat ćemo vektor

$$(Z_{(1)}^n, Z_{(2)}^n, \dots, Z_{(r)}^n),$$

kada $n \rightarrow \infty$ i $n\alpha \rightarrow \theta > 0$.

Propozicija 3.7. Ako je slučajni vektor $(Z_1^n, \dots, Z_{n-1}^n)$ dan funkcijom gustoće (3.3) onda je za $1 \leq r < n$ funkcija gustoće vektora (Z_1^n, \dots, Z_r^n) u točki (z_1, \dots, z_r) , gdje je $z_1, \dots, z_r > 0$ i $z_1 + \dots + z_r < 1$ dana formulom

$$h_r(z_1, \dots, z_r) = \frac{\Gamma(n\alpha)}{\Gamma^r(\alpha) \Gamma((n-r)\alpha)} z_1^{\alpha-1} \dots z_r^{\alpha-1} (1 - z_1 - \dots - z_r)^{(n-r)\alpha-1}.$$

Dokaz. Za proizvoljan $1 \leq r < n$ označimo s h_r funkcija gustoće vektora (Z_1^n, \dots, Z_r^n) . Označimo s $m := n - r - 1$ i $\sigma := \sum_{i=1}^r z_i$. Zbog (3.3) slijedi da za vektor (z_1, \dots, z_r) , takav da je $z_1, \dots, z_r > 0$ i $z_1 + \dots + z_r < 1$ vrijedi

$$\begin{aligned} h_r(z_1, \dots, z_r) &= \int_{\substack{s_1, \dots, s_m > 0 \\ s_1 + \dots + s_m < 1 - \sigma}} f_n(z_1, \dots, z_r, s_1, \dots, s_m) = \\ &= \int_{\substack{s_1, \dots, s_m > 0 \\ s_1 + \dots + s_m < 1 - \sigma}} \frac{\Gamma(n\alpha)}{\Gamma^n(\alpha)} (z_1 \dots z_r)^{\alpha-1} (s_1 \dots s_m)^{\alpha-1} (1 - \sigma - s_1 - \dots - s_m)^{\alpha-1} ds_1 \dots ds_m = \\ &= \frac{\Gamma(n\alpha)}{\Gamma^n(\alpha)} (z_1 \dots z_r)^{\alpha-1} \int_{\substack{s_1, \dots, s_m > 0 \\ s_1 + \dots + s_m < 1 - \sigma}} (s_1 \dots s_m)^{\alpha-1} (1 - \sigma - s_1 - \dots - s_m)^{\alpha-1} ds_1 \dots ds_m = \end{aligned}$$

(ukoliko na uglatu zgradu primjenimo tvrdnju korolara 1.4 uz $f(t) = (1 - \sigma - t)^{\alpha-1}$)

$$\begin{aligned} & \frac{\Gamma(n\alpha)}{\Gamma^n(\alpha)} (z_1 \dots z_r)^{\alpha-1} \frac{\Gamma^m(\alpha)}{\Gamma(m\alpha)} \int_0^{1-\sigma} (1 - \sigma - t)^{\alpha-1} t^{m\alpha-1} dt = \left[s = \frac{t}{1-\sigma}, \quad ds = \frac{dt}{1-\sigma} \right] = \\ & \frac{\Gamma(n\alpha)}{\Gamma(m\alpha)\Gamma^{r+1}(\alpha)} (z_1 \dots z_r)^{\alpha-1} \int_0^1 (1-\sigma)^{\alpha-1} (1-s)^{\alpha-1} (1-\sigma)^{m\alpha-1} s^{m\alpha-1} (1-\sigma) ds = \\ & \frac{\Gamma(n\alpha)}{\Gamma(m\alpha)\Gamma^{r+1}(\alpha)} (z_1 \dots z_r)^{\alpha-1} (1-\sigma)^{(n-r)\alpha-1} \int_0^1 (1-s)^{\alpha-1} s^{m\alpha-1} ds \stackrel{\text{propozicija 1.2}}{=} \\ & \frac{\Gamma(n\alpha)}{\Gamma(m\alpha)\Gamma^{r+1}(\alpha)} (z_1 \dots z_r)^{\alpha-1} (1-\sigma)^{(n-r)\alpha-1} \frac{\Gamma(\alpha)\Gamma(m\alpha)}{\Gamma((n-r)\alpha)} = \\ & \frac{\Gamma(n\alpha)}{\Gamma^r(\alpha)\Gamma((n-r)\alpha)} z_1^{\alpha-1} \dots z_r^{\alpha-1} (1 - z_1 - \dots - z_r)^{(n-r)\alpha-1}. \end{aligned}$$

□

Ako umjesto promatranja konkretnog podskupa $\{1, 2, \dots, r\}$ uzmemo proizvoljni podskup skupa $\{1, 2, \dots, n\}$ duljine r i označimo ga s $\{i_1, i_2, \dots, i_r\}$ zbog simetrije i prethodne propozicije vrijedi tvrdnja sljedećeg korolara.

Korolar 3.8. Funkcija gustoće vektora $(Z_{i_1}^n, \dots, Z_{i_r}^n)$ u točki (z_1, \dots, z_r) , gdje je $z_1, \dots, z_r > 0$ i $z_1 + \dots + z_r < 1$ dana je formulom

$$h_r(z_1, \dots, z_r) = \frac{\Gamma(n\alpha)}{\Gamma^r(\alpha)\Gamma((n-r)\alpha)} z_1^{\alpha-1} \dots z_r^{\alpha-1} (1 - z_1 - \dots - z_r)^{(n-r)\alpha-1}. \quad (3.4)$$

Označimo s $(Z_{\tau_1}^n, \dots, Z_{\tau_r}^n)$ vektor dobiven kao u napomeni 1.19 iz \mathbb{Z}^n , odnosno vektor $(\hat{Z}_1^n, \dots, \hat{Z}_r^n)$. Budući da je po propoziciji 1.20,

$$\mathbb{P}[\tau_1 = i_1, \dots, \tau_r = i_r \mid Z^n] = Z_{i_1}^n \frac{Z_{i_2}^n}{1 - Z_{i_1}^n} \dots \frac{Z_{i_r}^n}{1 - Z_{i_1}^n - \dots - Z_{i_{r-1}}^n} =: p_r(Z_{i_1}^n, \dots, Z_{i_r}^n),$$

vrijedi

$$\mathbb{P}[(\hat{Z}_1^n, \dots, \hat{Z}_r^n) \in H \mid Z^n] = \sum p_r(Z_{i_1}^n, \dots, Z_{i_r}^n) \mathbb{1}_H(Z_{i_1}^n, \dots, Z_{i_r}^n),$$

gdje se sumira po uređenim r -torkama međusobno različitih elemenata skupa $\{1, 2, \dots, n\}$. Uočimo da tih pribrojnika ima $(n)_r = n(n-1)\dots(n-r+1)$

Budući da za događaj A i slučajnu varijablu X vrijedi

$$\mathbb{P}[A] = \mathbb{E}[\mathbf{1}_A] = \mathbb{E}[\mathbb{E}[\mathbf{1}_A | X]] = \mathbb{E}[\mathbb{P}[A | X]]$$

imamo sljedeću relaciju

$$\begin{aligned} \mathbb{P}[(\hat{Z}_1^n, \dots, \hat{Z}_r^n) \in H] &= \int_H \sum p_r(z_1, \dots, z_r) h_r(z_1, \dots, z_r) dz_1 \dots dz_r \stackrel{\text{zbog simetrije}}{=} \\ &= \int_H (n)_r p_r(z_1, \dots, z_r) h_r(z_1, \dots, z_r) dz_1 \dots dz_r = \\ &= \frac{\Gamma(n\alpha)}{\Gamma(\alpha)^r \Gamma((n-r)\alpha)} \int_H (n)_r \cdot z_1^{\alpha-1} \dots z_r^{\alpha-1} \cdot (1 - z_1 - \dots - z_r)^{(n-r)\alpha-1} \cdot \\ &= z_1 \frac{z_2}{1 - z_1} \dots \frac{z_r}{(1 - z_1 - \dots - z_{r-1})} \mathbb{1}_T(z_1, z_2, \dots, z_r) dz_1 \dots dz_r \stackrel{(3.2)}{=} \mathbb{P}[(G_1^n, \dots, G_r^n) \in H]. \end{aligned}$$

Iz čega slijedi da su vektori $(\hat{Z}_1^n, \dots, \hat{Z}_r^n)$ i (G_1^n, \dots, G_r^n) jednako distribuirani.

Korolar 3.9.

$$\mathbb{Z}_{(\cdot)}^n \Longrightarrow \mathbb{G}_{(\cdot)} \quad n \rightarrow \infty.$$

Dokaz. Ranije smo pokazali da $(G_1^n, \dots, G_r^n) \Longrightarrow (G_1, \dots, G_r)$ iz čega po teoremu 2.19 slijedi da $\rho(\mathbb{G}^n) \Longrightarrow \rho(\mathbb{G})$. Ako uz tu konvergenciju uvažimo i napomene prije iskaza korolara slijedi $\rho(\hat{\mathbb{Z}}^n) \Longrightarrow \rho(\mathbb{G})$. Budući da je $\hat{\mathbb{Z}}^n$ permutacija slučajnog vektora \mathbb{Z}^n vrijedi $\rho(\hat{\mathbb{Z}}^n) = \rho(\mathbb{Z}^n)$. Dakle imamo

$$\mathbb{Z}_{(\cdot)}^n = \rho(\mathbb{Z}^n) = \rho(\hat{\mathbb{Z}}^n) \Longrightarrow \rho(\mathbb{G}) = \mathbb{G}_{(\cdot)} \quad n \rightarrow \infty,$$

čime je korolar dokazan. □

U nastavku ćemo proučavati granična svojstva vektora $\mathbb{Z}_{(\cdot)}^n$, ali da bi mogli doći do potrebnih zaključaka prvo je potrebno dokazati određen dio tehničkih tvrdnji.

3.2 Tehnički rezultati

Za fiksni $r \in \mathbb{N}$ neka je

$$M_r := \left\{ (z_1, \dots, z_r) \in \mathbb{R}^r : 1 > z_1 > \dots > z_r > 0, \sum_{i=1}^r z_i < 1 \right\}. \quad (3.5)$$

Neka je za fiksni α funkcija $g_m(\cdot; \alpha)$ konvolucija funkcije $\alpha x^{\alpha-1}$ sa samom sobom na intervalu $\langle 0, 1 \rangle$ m puta.

Raspišimo malo detaljnije funkciju $g_m(\cdot; \alpha)$ za $m = 1, 2, 3$:

$$g_1(x; \alpha) = \alpha x^{\alpha-1}$$

$$g_2(x; \alpha) = \int_0^1 \alpha y^{\alpha-1} \cdot \alpha (x-y)^{\alpha-1} dy = \alpha^2 \int_0^1 (x-y)^{\alpha-1} y^{\alpha-1} dy$$

$$g_3(x; \alpha) = \int_0^1 \alpha z^{\alpha-1} \cdot \alpha^2 \int_0^1 y^{\alpha-1} (x-z-y)^{\alpha-1} dy dz = \alpha^3 \iint_{0 < y, z < 1} (x-y-z)^{\alpha-1} y^{\alpha-1} z^{\alpha-1} dy dz$$

Matematičkom indukcijom za općeniti m dobivamo formulu

$$g_m(x; \alpha) = \alpha^m \int_{0 < s_1, \dots, s_{m-1} < 1} (x - s_1 - \dots - s_{m-1})^{\alpha-1} \prod_{i=1}^{m-1} s_i^{\alpha-1} ds_1 \dots ds_{m-1}.$$

Sada smo u mogućnosti pokazati sljedeću lemu koja će nam biti od pomoći pri određivanju traženih svojstava Poisson-Dirichletove razdiobe.

Lema 3.10. Gustoća slučajnog vektora $(Z_{(1)}^n, Z_{(2)}^n, \dots, Z_{(r)}^n)$ u točki $(z_1, z_2, \dots, z_r) \in M_r$ dana je s

$$n(n-1) \dots (n-r+1) \frac{\Gamma(n\alpha)}{\Gamma^n(\alpha)} \alpha^{-(n-r)} z_1^{\alpha-1} \dots z_{r-1}^{\alpha-1} z_r^{(n-r+1)\alpha-2} g_{n-r} \left(\frac{1-z_1-\dots-z_r}{z_r}; \alpha \right).$$

Dokaz. Označimo s $m = n - 1 - r$, $s = \sum_{i=1}^r z_i$, $c = \frac{1-s}{z_r}$. Gustoća slučajnog vektora $(Z_1^n, Z_2^n, \dots, Z_{n-1}^n)$ dana je s

$$f_n(z_1, \dots, z_{n-1}) = \frac{\Gamma(n\alpha)}{\Gamma^n(\alpha)} z_1^{\alpha-1} \dots z_{n-1}^{\alpha-1} (1 - z_1 - \dots - z_{n-1})^{\alpha-1} \cdot \mathbb{1}_T(z_1, \dots, z_{n-1}),$$

gdje je $T = \{(x_1, \dots, x_{n-1}) \in \mathbb{R}^{n-1} : x_1, \dots, x_{n-1} > 0, x_1 + \dots + x_{n-1} < 1\}$. Uz ranije definirane oznake parcijalnom integracijom želimo dobiti gustoću slučajnog vektora $(Z_{(1)}^n, \dots, Z_{(r)}^n)$. Budući da su koordinate slučajnog vektora $(Z_{(1)}^n, \dots, Z_{(r)}^n)$ sortirane zanimat će nas samo one permutacije vektora (z_1, \dots, z_{n-1}) za koje je $z_1 > z_2 > \dots > z_r > z_{r+1}, \dots, z_{n-1}, z_n$, gdje je $z_n = 1 - \sum_{i=1}^{n-1} z_i$. Kako bi dobili taj uređaj, zbog simetričnosti izraza, dovoljno je formulu dobivenu parcijalnom integracijom pomnožiti s faktorom $(n)_r = n(n-1) \dots (n-r+1)$. Odnosno dobivamo formulu za gustoću

$$(n)_r \frac{\Gamma(n\alpha)}{\Gamma^n(\alpha)} z_1^{\alpha-1} \dots z_r^{\alpha-1} \int_{\substack{0 < y_1, \dots, y_m < z_r \\ 0 < 1-s-y_1-\dots-y_m < z_r}} y_1^{\alpha-1} \dots y_m^{\alpha-1} \left(1 - s - \sum_{i=1}^m y_i \right)^{\alpha-1} dy_1 \dots dy_m,$$

za $(z_1, \dots, z_r) \in M_r$. Dalje računamo vrijednost integrala

$$\begin{aligned}
 & \int_{\substack{0 < y_1, \dots, y_m < z_r \\ 0 < 1 - s - y_1 - \dots - y_m < z_r}} y_1^{\alpha-1} \dots y_m^{\alpha-1} \left(1 - s - \sum_{i=1}^m y_i \right)^{\alpha-1} dy_1 \dots dy_m = \\
 & [y_i = z_r s_i, dy_i = z_r ds_i, i = 1, \dots, m] \\
 & \int_{\substack{0 < s_1, \dots, s_m < 1 \\ 0 < c - s_1 - \dots - s_m < 1}} (z_r s_1)^{\alpha-1} \dots (z_r s_m)^{\alpha-1} z_r^{\alpha-1} \left(c - \sum_{i=1}^m s_i \right)^{\alpha-1} z_r^m ds_1 \dots ds_m = \\
 & z_r^{(\alpha-1)(m+1)+m} \int_{\substack{0 < s_1, \dots, s_m < 1 \\ c-1 < s_1 + \dots + s_m < c}} s_1^{\alpha-1} \dots s_m^{\alpha-1} \left(c - \sum_{i=1}^m s_i \right)^{\alpha-1} ds_1 \dots ds_m = \\
 & z_r^{\alpha(m+1)-1} \alpha^{-m} \int_{\substack{0 < s_1, \dots, s_m < 1 \\ c-1 < s_1 + \dots + s_m < c}} \left(\prod_{i=1}^m \alpha s_i^{\alpha-1} \right) \left(c - \sum_{i=1}^m s_i \right)^{\alpha-1} ds_1 \dots ds_m.
 \end{aligned}$$

Ukoliko u dobivenom integralu napravimo zamjenu varijabli $x = s_1 + \dots + s_m$, $s_m = x - s_1 - \dots - s_{m-1}$, $ds_m = dx$, slijedi

$$\begin{aligned}
 & \int_{\substack{0 < s_1, \dots, s_m < 1 \\ c-1 < s_1 + \dots + s_m < c}} \left(\prod_{i=1}^m \alpha s_i^{\alpha-1} \right) \left(c - \sum_{i=1}^m s_i \right)^{\alpha-1} ds_1 \dots ds_m = \\
 & \int_{c-1 < x < c} (c-x)^{\alpha-1} \int_{0 < s_1, \dots, s_{m-1} < 1} \alpha^m (x - s_1 - \dots - s_{m-1})^{\alpha-1} \prod_{i=1}^{m-1} s_i^{\alpha-1} ds_1 \dots ds_{m-1} dx = \\
 & \alpha^{-1} \int_{c-1 < x < c} \alpha (c-x)^{\alpha-1} g_m(x; \alpha) dx = \\
 & [t = c - x, dt = -dx] \\
 & \alpha^{-1} \int_0^1 \alpha t^{\alpha-1} g_m(c-t; \alpha) dt = \alpha^{-1} g_{m+1}(c; \alpha)
 \end{aligned}$$

Odnosno dobivamo formulu za gustoću slučajnog vektora $(Z_{(1)}^n, Z_{(2)}^n, \dots, Z_{(r)}^n)$

$$(n)_r \frac{\Gamma(n\alpha)}{\Gamma^n(\alpha)} z_1^{\alpha-1} \dots z_r^{\alpha-1} z_r^{\alpha(m+1)-1} \alpha^{-(m+1)} g_{m+1}(c; \alpha).$$

Iz čega, nakon što uvrstimo čemu su jednaki c , m i s , slijedi tvrdnja leme. \square

Uvedimo još jednu novu oznaku. Za $k \geq 1$ neka je

$$C_k(x) := \left\{ (s_1, \dots, s_k) \in \mathbb{R}^k : s_1, \dots, s_k > 1, \sum_{i=1}^k s_i < x \right\}.$$

Uočimo da za $x \leq k$ uvjeti iz definicije $C_k(x)$ nisu nikada oba zadovoljeni pa je $C_k(x) = \emptyset$. U slučaju kada je $k = 0$ recimo da je $C_0(x) := \emptyset$.

U nastavku će nas zanimati granična svojstva funkcije $g_m(x; \alpha)$ čija je važnost postala očita u prethodnoj lemi. Tvrdnja sljedećeg teorema pokazat će se ključna u dobivanju traženih rezultata.

Teorem 3.11. Za $0 < x < m$, vrijedi

$$g_m(x; \alpha) = \sum_{0 \leq k < x} (-1)^k \binom{m}{k} \frac{\Gamma^{m-k}(\alpha) \alpha^m}{\Gamma((m-k)\alpha)} \cdot \int_{C_k(x)} s_1^{\alpha-1} \dots s_k^{\alpha-1} \left(x - \sum_{i=1}^k s_i \right)^{(m-k)\alpha-1} ds_1 \dots ds_k,$$

gdje se u slučaju $k = 0$ integral zamjeni s $x^{m\alpha-1}$.

Dokaz. Neka je

$$A_m(x) := \left\{ (s_1, \dots, s_m) \in \mathbb{R}^m : 0 < s_1, \dots, s_m < 1, \sum_{i=1}^m s_i < x \right\} \text{ i}$$

$$B_k(x) := \left\{ (s_1, \dots, s_k) \in \mathbb{R}^k : s_1, \dots, s_k > 0, \sum_{i=1}^k s_i < x \right\}.$$

Za $0 < x < m$ imamo

$$\int_0^x g_m(u; \alpha) du = \int_0^x \alpha^m \int_{0 < s_1, \dots, s_{m-1} < 1} (u - s_1 - \dots - s_{m-1})^{\alpha-1} \prod_{i=1}^{m-1} s_i^{\alpha-1} ds_1 \dots ds_{m-1} du =$$

$$\left[u - \sum_{i=1}^{m-1} s_i = s_m \quad du = ds_m \right] = \int_{0 < s_1, \dots, s_{m-1} < 1} \int_{0 < s_1 + \dots + s_m < x} \alpha^m s_1^{\alpha-1} \dots s_m^{\alpha-1} ds_m ds_1 \dots ds_{m-1}.$$

Pokažimo da se integracija u zadnjem integralu zapravo vrši po skupu $A_m(x)$. Uvjeti $0 < s_1, \dots, s_{m-1} < 1$ i $\sum_{i=1}^m s_i < x$ su trivijalno zadovoljeni. Pretpostavimo da uvjet $s_m > 0$ nije zadovoljen, odnosno budući da nas kod integracije rubovi ne zanimaju pretpostavimo da je $s_m = -\varepsilon < 0$. Budući da se integrira po svim $0 < s_1, \dots, s_{m-1}$ dozvoljen je slučaj $s_1 = \dots = s_{m-1} = \varepsilon(2m-2)^{-1}$, ali tada je $\sum_{i=1}^m s_i = \frac{\varepsilon}{2} - \varepsilon < 0$ što je kontradikcija. Dakle

$s_m > 0$. Pretpostavimo nadalje da je $s_m > 1$. Budući da je $x < m$ postoji $\varepsilon' > 0$ takav da je $x = m - \varepsilon'$. Tada za $s_1 = \dots = s_{m-1} = 1 - \varepsilon' (2m - 2)^{-1}$ imamo $\sum_{i=1}^m s_i = m - 1 - \frac{\varepsilon'}{2} + s_m > m - \frac{\varepsilon'}{2} > x$. Ovime smo pokazali da je zadovoljen i uvjet $s_m < 1$. Odnosno pokazali smo da je

$$\int_0^x g_m(u; \alpha) du = \int_{A_m(x)} \alpha^m s_1^{\alpha-1} \dots s_m^{\alpha-1} ds_1 \dots ds_m.$$

Uvedimo sljedeće pomoćne oznake

$$P_k^n := \{(s_1, \dots, s_n) \in \mathbb{R}^n : s_k < 1\}, \quad n = 1, 2, \dots, m, \quad k = 1, 2, \dots, n.$$

Uz ove oznake dobivamo da vrijede sljedeće skupovne relacije

$$\begin{aligned} A_m(x) &= B_m(x) \cap P_1^m \cap \dots \cap P_m^m, \\ C_k(x) &= B_k(x) \cap (P_1^k)^c \cap \dots \cap (P_k^k)^c. \end{aligned}$$

Iz dobivenog korištenjem korolar 1.26 za $n = m$, $X = B_m(x)$, $S_i = P_i^m$ te $f(x_1, \dots, x_m) = \alpha^m (x_1, \dots, x_m)^{\alpha-1}$ i činjenice da je funkcija f simetrična s obzirom na varijable x_1, \dots, x_m slijedi

$$\begin{aligned} \int_{A_m(x)} \alpha^m s_1^{\alpha-1} \dots s_m^{\alpha-1} ds_1 \dots ds_m &= \int_{B_m(x) \cap P_1^m \cap \dots \cap P_m^m} \alpha^m s_1^{\alpha-1} \dots s_m^{\alpha-1} ds_1 \dots ds_m = \\ &= \sum_{0 \leq k \leq m} (-1)^k \binom{m}{k} \alpha^m \int_{B_m(x) \cap (P_1^m)^c \cap \dots \cap (P_k^m)^c} s_1^{\alpha-1} \dots s_k^{\alpha-1} s_{k+1}^{\alpha-1} \dots s_m^{\alpha-1} ds_1 \dots ds_m = \end{aligned}$$

(budući da se integrira po području na kojem su $s_1, \dots, s_k > 1$ i vrijedi $\sum_{i=1}^k s_i < x$ ako je $k \geq x$ područje intergacije je prazan skup, odnosno odgovarajući pribrojnik je 0)

$$\begin{aligned} \sum_{0 \leq k < x} (-1)^k \binom{m}{k} \alpha^m \int_{B_k(x) \cap (P_1^k)^c \cap \dots \cap (P_k^k)^c} \int_{B_{m-k}(x - \sum_{i=1}^k s_i)} s_1^{\alpha-1} \dots s_k^{\alpha-1} s_{k+1}^{\alpha-1} \dots s_m^{\alpha-1} ds_{k+1} \dots ds_1 \dots ds_k = \\ \sum_{0 \leq k < x} (-1)^k \binom{m}{k} \alpha^m \int_{C_k(x)} s_1^{\alpha-1} \dots s_k^{\alpha-1} \left[\int_{B_{m-k}(x - \sum_{i=1}^k s_i)} s_{k+1}^{\alpha-1} \dots s_m^{\alpha-1} ds_{k+1} \dots ds_m \right] ds_1 \dots ds_k = \end{aligned}$$

(ukoliko na uglatu zagradu primjenimo korolar 1.4 uz funkciju $f \equiv 1$)

$$\sum_{0 \leq k < x} (-1)^k \binom{m}{k} \alpha^m \int_{C_k(x)} s_1^{\alpha-1} \dots s_k^{\alpha-1} \frac{\Gamma^{m-k}(\alpha)}{\Gamma((m-k)\alpha)} \int_0^{x - \sum_{i=1}^k s_i} t^{(m-k)\alpha-1} dt ds_1 \dots ds_k =$$

$$\sum_{0 \leq k < x} (-1)^k \binom{m}{k} \frac{\alpha^m \Gamma^{m-k}(\alpha)}{(m-k)\alpha \Gamma((m-k)\alpha)} \int_{C_k(x)} s_1^{\alpha-1} \dots s_k^{\alpha-1} \left(x - \sum_{i=1}^k s_i \right)^{(m-k)\alpha} ds_1 \dots ds_k.$$

U slučaju $k = 0$ u izrazu u drugom redu zapravo imamo samo integral po $B_m(x)$, koji zatim u zadnjem redu prelazi u $x^{m\alpha}$.

Budući da smo promatrali $\int_0^x g_m(u; \alpha) du$ zanimat će nas derivacija dobivenog izraza.

Želimo po x derivirati izraz oblika $\sum_{0 \leq k < x} f_k(x)$. Uočimo da izraz ovog oblika ne mora uvijek biti derivabilan, čak ni u slučaju kada su f_k derivabilne, no pokazuje se da u našem slučaju jest. Zaista, za $x \notin \mathbb{N}$, a to se svodi na derivaciju od f_k za fiksni k . Promotrimo prvo slučaj $x \in \mathbb{N}$. U tom slučaju tražena derivacija je $\sum_{0 \leq k < x} f'_k(x) + \lim_{h \searrow 0} f_x(x+h) \cdot h^{-1}$. Za određivanje limesa koristit ćemo argument sličan onome pomoću kojega smo zaključili da je dovoljno sumirati do x . Ako integriramo po području na kojem su $s_1, \dots, s_k > 1$ i vrijedi $\sum_{i=1}^k s_i < x+h$, gdje $h \searrow 0$ za $k = x$ imamo da je $\sum_{i=1}^x s_i = x + \delta > x+h$ za dovoljno mali h . Odnosno traženi limes je za dovoljno mali h integral po praznom skupu, odnosno on je 0. Ovime se slučaj $x \in \mathbb{N}$ svodi na $x \notin \mathbb{N}$.

S obzirom na ranije komentare za $k = 0$ tražena derivacija se svodi na derivaciju od $x^{m\alpha}$, odnosno ako uzmemo u obzir konstante dobivamo

$$(-1)^0 \binom{m}{0} \frac{\alpha^m \Gamma^m(\alpha)}{m\alpha \Gamma(m\alpha)} m\alpha x^{m\alpha-1}.$$

Uočimo da dobiveni izraz odgovara izrazu u slučaju $k = 0$ iz iskaza teorema. Preostaje promotriti što se događa u slučaju $k > 0$.

Neka je $1 \leq k < x$ fiksni. Budući da je $x < m$ vrijedi $\beta := (m-k)\alpha > 0$. Označimo sa $\sigma := \sum_{i=1}^k s_i$. Zanima nas derivacija integrala

$$\int_{C_k(x)} s_1^{\alpha-1} \dots s_k^{\alpha-1} \left(x - \sum_{i=1}^k s_i \right)^{(m-k)\alpha} ds_1 \dots ds_k$$

po varijabli x . Odnosno računamo sljedeći limes

$$\lim_{h \searrow 0} \frac{1}{h} \left[\int_{C_k(x+h)} (s_1 \dots s_k)^{\alpha-1} (x+h-\sigma)^\beta ds_1 \dots ds_k - \int_{C_k(x)} (s_1 \dots s_k)^{\alpha-1} (x-\sigma)^\beta ds_1 \dots ds_k \right] =$$

$$\lim_{h \searrow 0} \frac{1}{h} \int_{C_k(x+h) \setminus C_k(x)} (s_1 \dots s_k)^{\alpha-1} (x+h-\sigma)^\beta ds_1 \dots ds_k +$$

$$\int_{C_k(x)} (s_1 \dots s_k)^{\alpha-1} \frac{(x+h-\sigma)^\beta - (x-\sigma)^\beta}{h} ds_1 \dots ds_k.$$

Za $(s_1, \dots, s_k) \in C_k(x+h) \setminus C_k(x)$ je $\sigma = \sum_{i=1}^k s_i \geq x$ odnosno $x+h-\sigma \leq h$. Zato je

$$\begin{aligned} & \lim_{h \searrow 0} \frac{1}{h} \int_{C_k(x+h) \setminus C_k(x)} (s_1 \dots s_k)^{\alpha-1} (x+h-\sigma)^\beta ds_1 \dots ds_k \leq \\ & \lim_{h \searrow 0} h^{\beta-1} \int_{C_k(x+h) \setminus C_k(x)} (s_1 \dots s_k)^{\alpha-1} ds_1 \dots ds_k = \\ & \lim_{h \searrow 0} h^{\beta-1} \left[\int_{C_k(x+h)} (s_1 \dots s_k)^{\alpha-1} ds_1 \dots ds_k - \int_{C_k(x)} (s_1 \dots s_k)^{\alpha-1} ds_1 \dots ds_k \right] \stackrel{\text{korolar 1.4}}{=} \\ & \lim_{h \searrow 0} h^{\beta-1} \left[\frac{\Gamma(\alpha)^k}{\Gamma(k\alpha)} \int_0^{x+h} t^{k\alpha-1} dt - \frac{\Gamma(\alpha)^k}{\Gamma(k\alpha)} \int_0^x t^{k\alpha-1} dt \right] = \lim_{h \searrow 0} h^{\beta-1} \frac{\Gamma(\alpha)^k}{\Gamma(k\alpha)} \int_x^{x+h} t^{k\alpha-1} dt = \\ & \lim_{h \searrow 0} h^{\beta-1} \frac{\Gamma(\alpha)^k}{\Gamma(k\alpha) k\alpha} \left((x+h)^{k\alpha} - x^{k\alpha} \right) = \lim_{h \searrow 0} \frac{\Gamma(\alpha)^k h^\beta}{\Gamma(k\alpha) k\alpha} \frac{(x+h)^{k\alpha} - x^{k\alpha}}{h} = \\ & \lim_{h \searrow 0} \frac{\Gamma(\alpha)^k h^\beta}{\Gamma(k\alpha)} x^{k\alpha-1} = 0, \end{aligned}$$

gdje zadnja jednakost vrijedi jer je $\beta > 0$. Preostaje odrediti

$$\lim_{h \searrow 0} \int_{C_k(x)} (s_1 \dots s_k)^{\alpha-1} \frac{(x+h-\sigma)^\beta - (x-\sigma)^\beta}{h} ds_1 \dots ds_k.$$

Budući da je $\lim_{h \searrow 0} \left((x+h-\sigma)^\beta - (x-\sigma)^\beta \right) \cdot h^{-1} = \beta (x-\sigma)^{\beta-1}$ pod uvjetom da možemo zamijeniti poredak limesa i integrala odnosno pokazati da vrijedi

$$\begin{aligned} & \lim_{h \searrow 0} \int_{C_k(x)} (s_1 \dots s_k)^{\alpha-1} \frac{(x+h-\sigma)^\beta - (x-\sigma)^\beta}{h} ds_1 \dots ds_k = \\ & \int_{C_k(x)} \lim_{h \searrow 0} (s_1 \dots s_k)^{\alpha-1} \frac{(x+h-\sigma)^\beta - (x-\sigma)^\beta}{h} ds_1 \dots ds_k \end{aligned}$$

dobivamo da je derivacija k -tog člana sume dana formulom

$$(-1)^k \binom{m}{k} \frac{\alpha^m \Gamma^{m-k}(\alpha)}{(m-k)\alpha \Gamma((m-k)\alpha)} \int_{C_k(x)} (s_1 \dots s_k)^{\alpha-1} \beta (x-\sigma)^{\beta-1} ds_1 \dots ds_k.$$

Nakon uvrštavanja vrijednosti β i σ i sumacije po k dobivamo

$$g_m(x; \alpha) = \sum_{0 \leq k < x} (-1)^k \binom{m}{k} \frac{\Gamma^{m-k}(\alpha) \alpha^m}{\Gamma((m-k)\alpha)} \cdot \int_{C_k(x)} s_1^{\alpha-1} \dots s_k^{\alpha-1} \left(x - \sum_{i=1}^k s_i \right)^{(m-k)\alpha-1} ds_1 \dots ds_k,$$

gdje se u slučaju $k = 0$ integral zamjeni s $x^{m\alpha-1}$.

Ostali smo dužni opravdati zamjenu limesa i integrala za što ćemo iskoristiti Lebesgueov teorem o dominiranoj konvergenciji. Po teoremu srednje vrijednosti za funkciju $f(y) = y^\beta$ na intervalu $[x - \sigma, x - \sigma + h]$ postoji $c \in \langle x - \sigma, x - \sigma + h \rangle$ takav da je

$$\beta c^{\beta-1} = \frac{(x - \sigma + h)^\beta - (x - \sigma)^\beta}{x - \sigma + h - (x - \sigma)} = \frac{(x - \sigma + h)^\beta - (x - \sigma)^\beta}{h}.$$

Ako je $\beta \geq 1$ onda za $h < x$ imamo $\beta c^{\beta-1} \leq \beta (2x)^{\beta-1}$, a ako je $0 < \beta < 1$ onda je $\beta c^{\beta-1} \leq \beta (x - \sigma)^{\beta-1}$. U oba slučaja postoji funkcija g takva da je

$$(s_1 \dots s_k)^{\alpha-1} \frac{(x + h - \sigma)^\beta - (x - \sigma)^\beta}{h} \leq g(s_1, \dots, s_k).$$

U slučaju $\beta \geq 1$ vrijedi

$$\begin{aligned} \int_{C_k(x)} g(s_1, \dots, s_k) ds_1 \dots ds_k &\leq \int_{B_k(x)} (s_1 \dots s_k)^{\alpha-1} \beta (2x)^{\beta-1} ds_1 \dots ds_k \stackrel{\text{korolar 1.4}}{=} \\ \beta (2x)^{\beta-1} \frac{\Gamma(\alpha)^k}{\Gamma(k\alpha)} \int_0^x t^{\alpha-1} dt &= \frac{\beta}{\alpha-1} 2^{\beta-1} x^{\alpha+\beta-1} \frac{\Gamma(\alpha)^k}{\Gamma(k\alpha)} < \infty, \end{aligned}$$

a u slučaju $0 < \beta < 1$

$$\begin{aligned} \int_{C_k(x)} g(s_1, \dots, s_k) ds_1 \dots ds_k &\leq \int_{B_k(x)} (s_1 \dots s_k)^{\alpha-1} \beta (x - \sigma)^{\beta-1} ds_1 \dots ds_k \stackrel{\text{korolar 1.4}}{=} \\ \beta \frac{\Gamma(\alpha)^k}{\Gamma(k\alpha)} \int_0^x (x-t)^{\beta-1} t^{\alpha-1} dt &= [t = xs, dt = xds] = \\ \beta x^{\alpha+\beta-1} \frac{\Gamma(\alpha)^k}{\Gamma(k\alpha)} \int_0^1 (1-s)^{\beta-1} s^{\alpha-1} ds &\stackrel{1.2}{=} \beta x^{\alpha+\beta-1} \frac{\Gamma(\alpha)^{k+1} \Gamma(\beta)}{\Gamma(k\alpha) \Gamma(\alpha + \beta)} < \infty \end{aligned}$$

pa možemo primijeniti Lebesgueov teorem o dominiranoj konvergenciji. Čime je tvrdnja teorema dokazana. \square

Neka je za $k \geq 1$ i $\theta > 0$ definirana funkcija

$$J_k(x; \theta) := \int_{C_k(x)} \frac{\left(x - \sum_{i=1}^k s_i\right)^{\theta-1}}{s_1 \cdot \dots \cdot s_k} ds_1 \dots ds_k. \quad (3.6)$$

Uočimo da je za $x \leq k$ $C_k(x) = 0$, a onda i $J_k(x; \theta) = 0$. U slučaju kada je $k = 0$ recimo da je $J_0(x; \theta) := x^{\theta-1}$.

Sada smo spremni dokazati rezultate vezane uz granična svojstva funkcije $g_m(x; \alpha)$, koja su iskazana sljedećim korolarom.

Korolar 3.12. Ako $m \rightarrow \infty$ i $m\alpha \rightarrow \theta > 0$, onda $g_m(x; \alpha)$ konvergira prema

$$g_\theta(x) = \sum_{0 \leq k < x} \frac{(-1)^k \theta^k}{\Gamma(\theta) e^{\gamma\theta} k!} J_k(x; \theta) \quad (3.7)$$

za svaki pozitivni x . Definiran gornjom formulom g_θ je vjerojatnosna funkcija gustoće na $\langle 0, \infty \rangle$. Korišteni γ je Eulerova konstanta definirana pomoću relacije 1.4.

Dokaz. Promotrimo što se događa kada $m \rightarrow \infty$ i $m\alpha \rightarrow \theta > 0$. Iz (1.4) znamo da je $\Gamma'(1) = -\gamma$, a iz propozicije 1.22 da je $\alpha\Gamma(\alpha) = \Gamma(1 + \alpha) = 1 - \gamma\alpha + O(\alpha^2)$. Zato vrijedi

$$\begin{aligned} \lim_{m \rightarrow \infty} (\alpha\Gamma(\alpha))^m &= \lim_{m \rightarrow \infty} \left(1 - \gamma\alpha + O(\alpha^2)\right)^m = \lim_{m \rightarrow \infty} \left(1 + \frac{-\gamma\theta}{m} \left(\frac{m\alpha}{\theta} - O(\alpha^2) \frac{m}{\gamma\theta}\right)\right)^m = \\ \lim_{m \rightarrow \infty} \left(1 + \frac{-\gamma\theta}{m} \left(1 - O(\alpha^2) \frac{1}{\gamma\alpha}\right)\right)^m &= \lim_{m \rightarrow \infty} \left(1 + \frac{-\gamma\theta}{m} (1 - O(\alpha))\right)^m = \lim_{m \rightarrow \infty} \left(1 + \frac{-\gamma\theta}{m}\right)^m = e^{-\gamma\theta}. \end{aligned}$$

Zbog neprekidnosti funkcija Γ za proizvoljni x i svaki $k < x$ je $\lim_{m \rightarrow \infty} (\Gamma((m-k)\alpha))^{-1} = (\Gamma(\theta))^{-1}$. Budući da je $(m-k)^k \leq \frac{m!}{(m-k)!} \leq m^k$ i jer vrijedi

$$\lim_{m \rightarrow \infty} \frac{\Gamma(\alpha)^k}{(m-k)^k} = \lim_{m \rightarrow \infty} \frac{(\Gamma(\alpha+1))^k}{(\alpha m - \alpha k)^k} = \frac{1}{\theta^k} = \lim_{m \rightarrow \infty} \frac{(\Gamma(\alpha+1))^k}{(\alpha m)^k} = \lim_{m \rightarrow \infty} \frac{\Gamma(\alpha)^k}{m^k}$$

po teoremu o sendviču slijedi $\lim_{m \rightarrow \infty} \frac{m!}{(m-k)!} \Gamma(\alpha)^{-k} = \theta^k$.

Dosadašnjim razmatranjima pokazali smo da je

$$\begin{aligned} \lim_{m \rightarrow \infty} (-1)^k \binom{m}{k} \frac{\Gamma^{m-k}(\alpha) \alpha^m}{\Gamma((m-k)\alpha)} &= \lim_{m \rightarrow \infty} (-1)^k \frac{1}{\Gamma((m-k)\alpha)} (\alpha\Gamma(\alpha))^m \frac{m!}{(m-k)!} \Gamma(\alpha)^{-k} \frac{1}{k!} = \\ (-1)^k \frac{1}{\Gamma(\theta)} e^{-\gamma\theta} \theta^k \frac{1}{k!} &= \frac{(-1)^k \theta^k}{\Gamma(\theta) e^{\gamma\theta} k!}. \end{aligned}$$

Preostaje pokazati da je

$$\lim_{m \rightarrow \infty} \int_{C_k(x)} s_1^{\alpha-1} \dots s_k^{\alpha-1} \left(x - \sum_{i=1}^k s_i \right)^{(m-k)\alpha-1} ds_1 \dots ds_k = J_k(x; \theta).$$

Za $k = 0$, kao što je napomenuto u iskazu teorema 3.11, umjesto integrala uzimamo $x^{m\alpha-1}$, a kako je $J_0(x; \theta) = x^{\theta-1}$ u tom slučaju konvergencija trivijalno vrijedi. Neka je zato $1 \leq k < x$. Uvedimo novu oznaku

$$D_k(x) := \left\{ (s_1, \dots, s_k) \in \mathbb{R}^k : s_1, \dots, s_k > \frac{1}{x}, \sum_{i=1}^k s_i < 1 \right\}$$

te kao i u prethodnom dokazu neka je

$$B_k(x) = \left\{ (s_1, \dots, s_k) \in \mathbb{R}^k : s_1, \dots, s_k > 0, \sum_{i=1}^k s_i < x \right\}.$$

te pomoću nje napravimo zamjenu varijabli u promatranim integralima. Neka je $s_i = xu_i$, $ds_i = xdu_i$.

$$\begin{aligned} & \int_{C_k(x)} s_1^{\alpha-1} \dots s_k^{\alpha-1} \left(x - \sum_{i=1}^k s_i \right)^{(m-k)\alpha-1} ds_1 \dots ds_k = \\ & \int_{D_k(x)} x^{(\alpha-1)k} u_1^{\alpha-1} \dots u_k^{\alpha-1} x^{(m-k)\alpha-1} \left(1 - \sum_{i=1}^k u_i \right)^{(m-k)\alpha-1} x^k du_1 \dots du_k = \\ & x^{m\alpha-1} \int_{D_k(x)} u_1^{\alpha-1} \dots u_k^{\alpha-1} \left(1 - \sum_{i=1}^k u_i \right)^{(m-k)\alpha-1} du_1 \dots du_k = \\ & x^{m\alpha-1} \int_{D_k(x)} \frac{\left(1 - \sum_{i=1}^k u_i \right)^{(m-k)\alpha-1}}{u_1 \dots u_k} (u_1 \dots u_k)^\alpha du_1 \dots du_k, \\ & \int_{C_k(x)} \frac{\left(x - \sum_{i=1}^k s_i \right)^{\theta-1}}{s_1 \dots s_k} ds_1 \dots ds_k = \int_{D_k(x)} x^{\theta-1} \frac{\left(1 - \sum_{i=1}^k u_i \right)^{\theta-1}}{x^k u_1 \dots u_k} x^k du_1 \dots du_k = \\ & x^{\theta-1} \int_{D_k(x)} \frac{\left(1 - \sum_{i=1}^k u_i \right)^{\theta-1}}{u_1 \dots u_k} du_1 \dots du_k. \end{aligned}$$

Budući da $x^{m\alpha-1} \rightarrow x^{\theta-1}$ kada $m \rightarrow \infty$ preostaje vidjeti što se događa s integralima. Neka je zato $0 < \theta_0 < \theta$, a budući da je $\lim_{m \rightarrow \infty} (m-k)\alpha = \theta$ za dovoljno veliki m je $(m-k)\alpha > \theta_0$. Zato vrijedi

$$\begin{aligned} & \left| \frac{(1 - \sum_{i=1}^k u_i)^{(m-k)\alpha-1}}{u_1 \cdots u_k} (u_1 \cdots u_k)^\alpha - \frac{(1 - \sum_{i=1}^k u_i)^{\theta-1}}{u_1 \cdots u_k} \right|_{u_1, \dots, u_k > \frac{1}{x}} \leq \\ & x^k \left| \left(1 - \sum_{i=1}^k u_i\right)^{(m-k)\alpha-1} (u_1 \cdots u_k)^\alpha - \left(1 - \sum_{i=1}^k u_i\right)^{\theta-1} \right|_{u_1, \dots, u_k < 1} \leq \\ & x^k \left| \left(1 - \sum_{i=1}^k u_i\right)^{(m-k)\alpha-1} \right| + \left| \left(1 - \sum_{i=1}^k u_i\right)^{\theta-1} \right|_{0 < 1 - \sum_{i=1}^k u_i < 1} 2x^k \left(1 - \sum_{i=1}^k u_i\right)^{\theta_0-1} . \\ & \int_{D_k(x)} 2x^k \left(1 - \sum_{i=1}^k u_i\right)^{\theta_0-1} du_1 \dots du_k \leq 2x^k \int_{B_k(1)} \left(1 - \sum_{i=1}^k u_i\right)^{\theta_0-1} du_1 \dots du_k \stackrel{\text{korolar 1.4}}{=} \\ & 2x^k \frac{1}{\Gamma(k)} \int_0^1 (1-t)^{\theta_0-1} t^{k-1} dt = 2x^k \frac{\Gamma(\theta_0)\Gamma(k)}{\Gamma(k)\Gamma(\theta_0+k)} = 2x^k \frac{\Gamma(\theta_0)}{\Gamma(\theta_0+k)} < \infty. \end{aligned}$$

Budući da

$$\frac{(1 - \sum_{i=1}^k u_i)^{(m-k)\alpha-1}}{u_1 \cdots u_k} (u_1 \cdots u_k)^\alpha - \frac{(1 - \sum_{i=1}^k u_i)^{\theta-1}}{u_1 \cdots u_k} \rightarrow 0 \text{ kada } m \rightarrow \infty$$

po Lebesgueovom teoremu o dominiranoj konvergenciji slijedi

$$\int_{D_k(x)} \frac{(1 - \sum_{i=1}^k u_i)^{(m-k)\alpha-1}}{u_1 \cdots u_k} (u_1 \cdots u_k)^\alpha - \frac{(1 - \sum_{i=1}^k u_i)^{\theta-1}}{u_1 \cdots u_k} du_1 \dots du_k \rightarrow 0 \text{ kada } m \rightarrow \infty.$$

Ovime smo pokazali da $g_m(x; \alpha) \rightarrow g_\theta(x)$ kada $m \rightarrow \infty$ i $m\alpha \rightarrow \theta$ za svaki pozitivni x .

Preostaje pokazati da je g_θ vjerojatnosna funkcija gustoće. Budući da je $g_m(x; \alpha) \geq 0$ za svaki $x > 0$ pa isto vrijedi i za g_θ . Kako bi smo pokazali da je $\int_0^\infty g_\theta(x) dx = 1$ koristimo sljedeću lemu.

Lema 3.13. Laplaceova transformacija funkcije g_θ jest

$$\int_0^\infty e^{-tx} g_\theta(x) dx = e^{-\gamma\theta} t^{-\theta} \exp \left[-\theta \int_t^\infty e^{-u} u^{-1} du \right] = \exp \left[-\theta \int_0^t \frac{1 - e^{-u}}{u} du \right]. \quad (3.8)$$

Dokaz. Označimo sa $\sigma := \sum_{i=1}^k s_i$.

$$\int_0^{\infty} e^{-tx} g_{\theta}(x) dx = \int_0^{\infty} \sum_{0 \leq k < x} e^{-tx} \frac{(-1)^k \theta^k}{\Gamma(\theta) e^{\gamma\theta} k!} \int_{C_k(x)} \frac{(x - \sum_{i=1}^k s_i)^{\theta-1}}{s_1 \cdot \dots \cdot s_k} ds_1 \dots ds_k dx \stackrel{C_k(x)=\emptyset \text{ za } k \geq x}{=} \\ \frac{1}{\Gamma(\theta) e^{\gamma\theta}} \int_0^{\infty} \sum_{k=0}^{\infty} e^{-tx} \frac{(-\theta)^k}{k!} \int_{C_k(x)} \frac{(x - \sum_{i=1}^k s_i)^{\theta-1}}{s_1 \cdot \dots \cdot s_k} ds_1 \dots ds_k dx = \quad (3.9)$$

$$\frac{1}{\Gamma(\theta) e^{\gamma\theta}} \sum_{k=0}^{\infty} \frac{(-\theta)^k}{k!} \int_0^{\infty} \int_{C_k(x)} \frac{e^{-tx} (x - \sum_{i=1}^k s_i)^{\theta-1}}{s_1 \cdot \dots \cdot s_k} ds_1 \dots ds_k dx =$$

$$\frac{1}{\Gamma(\theta) e^{\gamma\theta}} \sum_{k=0}^{\infty} \frac{(-\theta)^k}{k!} \int_{s_1, \dots, s_k > 1} \frac{1}{s_1 \cdot \dots \cdot s_k} \int_{\sigma}^{\infty} e^{-tx} (x - \sigma)^{\theta-1} dx ds_1 \dots ds_k =$$

$$[y = t(x - \sigma), dy = t dx] =$$

$$\frac{1}{\Gamma(\theta) e^{\gamma\theta}} \sum_{k=0}^{\infty} \frac{(-\theta)^k}{k!} \int_{s_1, \dots, s_k > 1} \frac{1}{s_1 \cdot \dots \cdot s_k} \int_0^{\infty} e^{-t\sigma} e^{-y} t^{-\theta} (y)^{\theta-1} dy ds_1 \dots ds_k =$$

$$\frac{1}{\Gamma(\theta) e^{\gamma\theta}} \sum_{k=0}^{\infty} \frac{(-\theta)^k}{k!} \int_{s_1, \dots, s_k > 1} \frac{1}{s_1 \cdot \dots \cdot s_k} e^{-t \sum_{i=1}^k s_i} t^{-\theta} \Gamma(\theta) ds_1 \dots ds_k =$$

$$e^{-\gamma\theta} t^{-\theta} \sum_{k=0}^{\infty} \frac{(-\theta)^k}{k!} \left(\int_1^{\infty} \frac{e^{-ts}}{s} ds \right)^k = [ts = u, tds = du] = e^{-\gamma\theta} t^{-\theta} \sum_{k=0}^{\infty} \frac{(-\theta)^k}{k!} \left(\int_t^{\infty} \frac{e^{-u}}{u} du \right)^k =$$

$$e^{-\gamma\theta} t^{-\theta} \exp \left[-\theta \int_t^{\infty} e^{-u} u^{-1} du \right] = \exp \left[-\theta \left(\gamma + \ln(t) + \int_t^{\infty} e^{-u} u^{-1} du \right) \right] = \quad (3.10)$$

$$\exp \left[-\theta \left(- \int_0^{\infty} e^{-u} \ln(u) du + \ln(t) + e^{-u} \ln(u) \Big|_t^{\infty} + \int_t^{\infty} e^{-u} \ln(u) du \right) \right] =$$

$$\exp \left[-\theta \left(- \int_0^t e^{-u} \ln(u) du + (1 + e^{-t}) \ln(t) \right) \right] = \exp \left[-\theta \int_0^t \frac{1 - e^{-u}}{u} du \right]. \quad (3.11)$$

Zamjena integrala i sume u (3.9) je opravdana po Lebesgueovom teoremu o dominiranoj

konvergenciji, jer je daljnji račun isti bez člana $(-1)^k$, a

$$e^{-\gamma\theta} t^{-\theta} \exp \left[\theta \int_t^{\infty} e^{-u} u^{-1} du \right] < \infty.$$

Jednakost u (3.10) vrijedi jer je po (1.4) $\gamma = -\Gamma'(1) = -\int_0^{\infty} e^{-u} \ln(u) du$, a (3.11) zbog parcijalne integracije, jer je $\lim_{u \searrow 0} (1 + e^{-u}) \ln(u) = 0$. \square

Budući da je

$$\lim_{t \searrow 0} \exp \left[-\theta \int_0^t \frac{1 - e^{-u}}{u} du \right] = e^0 = 1 \text{ po (3.8) dobivamo } \lim_{t \searrow 0} \int_0^{\infty} e^{-tx} g_{\theta}(x) dx = 1.$$

Kako za $t_1 > t_2$ vrijedi $0 \leq e^{-t_1 x} g_{\theta}(x) \leq e^{-t_2 x} g_{\theta}(x) \leq g_{\theta}(x)$ po Lebesgueovom teoremu o monotonij konvergenciji dobivamo

$$1 = \lim_{t \searrow 0} \int_0^{\infty} e^{-tx} g_{\theta}(x) dx = \int_0^{\infty} \lim_{t \searrow 0} e^{-tx} g_{\theta}(x) dx = \int_0^{\infty} g_{\theta}(x) dx.$$

Ovime smo pokazali da je g_{θ} vjerojatnosna funkcija gustoće i dovršili dokaz korolara. \square

Zaključno s ovim korolarom raspoložemo svim rezultatima potrebnima da odredimo funkciju gustoće slučajnog vektora $\mathbb{G}_{(\cdot)}$ na prostoru Δ koji ima Poisson-Dirichletovu razdiobu s parametrom θ .

3.3 Svojstva Poisson-Dirichletove razdiobe

Teorem 3.14. *Za fiksiran r kada $n \rightarrow \infty$ i $n\alpha \rightarrow \theta$ funkcija gustoće slučajnog vektora $(Z_{(1)}^n, Z_{(2)}^n, \dots, Z_{(r)}^n)$ konvergira na skupu M_r , definiranom u (3.5), prema funkciji*

$$\theta^r \Gamma(\theta) e^{\gamma\theta} z_1^{-1} \cdot \dots \cdot z_{r-1}^{-1} z_r^{\theta-2} g_{\theta} \left(\frac{1 - z_1 - \dots - z_r}{z_r} \right). \quad (3.12)$$

Dokaz. Funkcija gustoće slučajnog vektora $(Z_{(1)}^n, Z_{(2)}^n, \dots, Z_{(r)}^n)$ u točki $(z_1, \dots, z_r) \in M_r$ po lemi 3.10 je

$$n(n-1)\dots(n-r+1) \frac{\Gamma(n\alpha)}{\Gamma^n(\alpha)} \alpha^{-(n-r)} z_1^{\alpha-1} \dots z_{r-1}^{\alpha-1} z_r^{(n-r+1)\alpha-2} g_{n-r} \left(\frac{1-z_1-\dots-z_r}{z_r}; \alpha \right) =$$

$$n\alpha \cdot (n-1)\alpha \cdot \dots \cdot (n-r+1)\alpha \frac{\Gamma(n\alpha)}{(\alpha\Gamma(\alpha))^n} z_1^{\alpha-1} \dots z_{r-1}^{\alpha-1} z_r^{(n-r+1)\alpha-2} g_{n-r} \left(\frac{1-z_1-\dots-z_r}{z_r}; \alpha \right).$$

$$\lim_{\substack{n \rightarrow \infty \\ n\alpha \rightarrow \theta}} (n-i)\alpha = \theta, \text{ za } i = 0, 1, \dots, r-1.$$

$$\lim_{\substack{n \rightarrow \infty \\ n\alpha \rightarrow \theta}} \Gamma(n\alpha) = \Gamma(\theta), \text{ jer je funkcija } \Gamma \text{ neprekidna.}$$

$$\lim_{\substack{n \rightarrow \infty \\ n\alpha \rightarrow \theta}} (\alpha\Gamma(\alpha))^n = e^{-\gamma\theta}, \text{ što je pokazano u dokazu korolara 3.12.}$$

$$\lim_{\substack{n \rightarrow \infty \\ n\alpha \rightarrow \theta}} z_i^{\alpha-1} = z_i^{-1}, \text{ za } i = 1, \dots, r-1.$$

$$\lim_{\substack{n \rightarrow \infty \\ n\alpha \rightarrow \theta}} z_r^{(n-r+1)\alpha-2} = z_r^{\theta-2}.$$

$$\lim_{\substack{n \rightarrow \infty \\ n\alpha \rightarrow \theta}} g_{n-r} \left(\frac{1-z_1-\dots-z_r}{z_r}; \alpha \right) = g_\theta \left(\frac{1-z_1-\dots-z_r}{z_r} \right), \text{ po tvrdnji korolara 3.12, jer je}$$

$$\frac{1-z_1-\dots-z_r}{z_r} > 0 \text{ za } (z_1, \dots, z_r) \in M_r.$$

Ovime je pokazano da funkcija gustoće slučajnog vektora $(Z_{(1)}^n, Z_{(2)}^n, \dots, Z_{(r)}^n)$ konvergira na skupu M_r prema funkciji

$$f(z_1, \dots, z_r) = \theta^r \Gamma(\theta) e^{\gamma\theta} z_1^{-1} \cdot \dots \cdot z_{r-1}^{-1} z_r^{\theta-2} g_\theta \left(\frac{1-z_1-\dots-z_r}{z_r} \right)$$

kada $n \rightarrow \infty$ i $n\alpha \rightarrow \theta$. □

Kada uzmemo u obzir zaključke s kraja sekcije 3.1, uočavamo važnost pokazivanja da je funkcija f definirana u prethodnom teoremu vjerojatnosna funkcija gustoće. To ćemo i pokazati, ali prvo se okrenimo proučavanju funkcije koju dobivamo iz f nakon što se integriranjem po skupu na kojem to ima smisla riješimo svih osim zadnje varijable. Tu funkciju možemo smatrati r -tom marginalnom gustoćom od funkcije f i njenu eksplisnitnu formulu daje nam sljedeći teorem.

Teorem 3.15. *Za $0 < x < r^{-1}$ r -ta marginalna gustoća od funkcije dane formulom (3.12) u točki x jest*

$$d_r(x; \theta) = x^{\theta-2} \sum_{0 \leq k < x^{-1}-r} (-1)^k \frac{\theta^{r+k}}{k!(r-1)!} J_{k+r-1} \left(\frac{1-x}{x}; \theta \right), \quad (3.13)$$

gdje je funkcija $J_k(\cdot; \theta)$ definirana u (3.6).

Napomena 3.16. Prije nego krenemo na dokaz obrazložimo zašto se u iskazu teorema uzima $0 < x < r^{-1}$. Nas zanima x takav da postoji vektor $(z_1, \dots, z_{r-1}, x) \in M_r$. Iz uvjeta $z_1 > \dots > z_r > 0$ i $\sum_{i=1}^r z_i < 1$ iz definicije prostora M_r slijedi $x > 0$, ali nameće se i uvjet $1 > \sum_{i=1}^{r-1} z_i + x > r \cdot x$ odnosno $x < r^{-1}$. Okrenimo se sada dokazu teorema 3.15.

Dokaz. Uvedimo oznaku

$$M_r(x) := \left\{ (z_1, \dots, z_{r-1}) \in \mathbb{R}^{r-1} : 1 > z_1 > \dots > z_{r-1} > x > 0, \sum_{i=1}^{r-1} z_i < 1 - x \right\},$$

jer se upravo po tom skupu vrši parcijalna integracija kojom dobivamo

$$\begin{aligned} d_r(x; \theta) &= \int_{M_r(x)} \theta^r \Gamma(\theta) e^{\gamma\theta} z_1^{-1} \cdot \dots \cdot z_{r-1}^{-1} x^{\theta-2} g_\theta \left(\frac{1 - z_1 - \dots - x}{x} \right) dz_1 \dots dz_{r-1} = \\ &= \int_{M_r(x)} \theta^r \Gamma(\theta) e^{\gamma\theta} z_1^{-1} \cdot \dots \cdot z_{r-1}^{-1} x^{\theta-2} \sum_{k=0}^{\infty} \frac{(-1)^k \theta^k}{\Gamma(\theta) e^{\gamma\theta} k!} J_k \left(\frac{1 - z_1 - \dots - x}{x}; \theta \right) dz_1 \dots dz_{r-1}. \end{aligned}$$

Budući da funkcija $J_k(\cdot; \theta)$ ima faktor u kojem se integrira po području $C_k(\cdot)$, pribrojnik za koje je $k \geq (1 - z_1 - \dots - x) x^{-1}$ su jednaki 0. Zato su pogotovo pribrojnici za koje je $k \geq x^{-1} - r = (1 - x - \dots - x) x^{-1} > (1 - z_1 - \dots - x) x^{-1}$ jednaki 0. Iz ovoga zaključujemo da se vrijednost integrala ne mijenja ukoliko se sumacija vršiti samo po $k < x^{-1} - r$.

$$d_r(x; \theta) = \sum_{0 \leq k < (\frac{1}{x} - r)} \int_{M_r(x)} \theta^{r+k} z_1^{-1} \cdot \dots \cdot z_{r-1}^{-1} x^{\theta-2} \frac{(-1)^k}{k!} J_k \left(\frac{1 - z_1 - \dots - x}{x}; \theta \right) dz_1 \dots dz_{r-1}.$$

Zbog simetričnosti izraza koji integriramo uvjet $z_1 > \dots > z_{r-1}$ možemo maknuti ukoliko tako dobiveni integral podijelimo s $(r-1)!$. Ukoliko još napravimo i zamjenu varijabli $xs_i = z_i$, $xs_i = dz_i$ za $i = 1, 2, \dots, r-1$ dobivamo

$$x^{\theta-2} \sum_{0 \leq k < (\frac{1}{x} - r)} \frac{(-1)^k \theta^{r+k}}{k! (r-1)!} \int_{C_{r-1}(\frac{1-x}{x})} s_1^{-1} \cdot \dots \cdot s_{r-1}^{-1} J_k \left(\frac{1-x}{x} - s_1 - \dots - s_{r-1}; \theta \right) ds_1 \dots ds_{r-1}.$$

Uvrstimo li u dobiveni integral čemu je po (3.6) jednaka vrijednost funkcije $J_k(\cdot; \theta)$ dobivamo

$$\begin{aligned} & \int_{C_{r-1}\left(\frac{1-x}{x}\right)} s_1^{-1} \cdots s_{r-1}^{-1} J_k\left(\frac{1-x}{x} - s_1 - \dots - s_{r-1}; \theta\right) ds_1 \dots ds_{r-1} = \\ & \int_{C_{r-1}\left(\frac{1-x}{x}\right)} s_1^{-1} \cdots s_{r-1}^{-1} \int_{C_k\left(\frac{1-x}{x} - s_1 - \dots - s_{r-1}\right)} \frac{\left(\frac{1-x}{x} - \sum_{i=1}^{r-1} s_i - \sum_{j=1}^k t_j\right)^{\theta-1}}{t_1 \cdots t_k} dt_1 \dots dt_k ds_1 \dots ds_{r-1} = \\ & \int_{C_{k+r-1}\left(\frac{1-x}{x}\right)} s_1^{-1} \cdots s_{r-1}^{-1} t_1^{-1} \cdots t_k^{-1} \left(\frac{1-x}{x} - \sum_{i=1}^{r-1} s_i - \sum_{j=1}^k t_j\right)^{\theta-1} ds_1 \dots ds_{r-1} dt_1 \dots dt_k = \\ & J_{k+r-1}\left(\frac{1-x}{x}; \theta\right). \end{aligned}$$

Ovime smo pokazali tvrdnju teorema, odnosno da je

$$d_r(x; \theta) = x^{\theta-2} \sum_{0 \leq k < x^{-1} - r} (-1)^k \frac{\theta^{r+k}}{k! (r-1)!} J_{k+r-1}\left(\frac{1-x}{x}; \theta\right).$$

□

Od pomoći za dokazivanje da je f definirana u teoremu 3.14 vjerojatnosna funkcija gustoće, ali i od važnosti kao takav, bit će nam sljedeći teorem.

Teorem 3.17. Za $m = 0, 1, 2, \dots$ m -ti momenti s obzirom na funkciju $d_r(\cdot; \theta)$ danu u (3.13) jesu

$$\int_0^{\frac{1}{r}} x^m d_r(x; \theta) dx = \frac{\Gamma(\theta+1)}{\Gamma(\theta+m)} \int_0^\infty y^{m-1} \frac{\theta^{r-1}}{(r-1)!} \left(\int_y^\infty e^{-u} u^{-1} du \right)^{r-1} \exp\left[-y - \theta \int_y^\infty e^{-u} u^{-1} du\right] dy.$$

Dokaz. Budući da je $d_r(x; \theta) = 0$ za $x \geq r^{-1}$ vrijedi

$$\begin{aligned} \int_0^{\frac{1}{r}} x^m d_r(x; \theta) dx &= \int_0^1 x^m d_r(x; \theta) dx = \left[y = \frac{1}{x}, dy = -\frac{dx}{x^2} \right] = \int_1^\infty \frac{d_r(y^{-1}; \theta)}{y^{m+2}} dy \stackrel{(3.13)}{=} \\ &= \int_1^\infty \frac{1}{y^{m+2}} y^{2-\theta} \sum_{0 \leq k < y-r} (-1)^k \frac{\theta^{r+k}}{k! (r-1)!} J_{k+r-1}(y-1; \theta) dy = \end{aligned}$$

$$\begin{aligned} & \sum_{k=0}^{\infty} (-1)^k \frac{\theta^{r+k}}{k! (r-1)!} \int_1^{\infty} y^{-(\theta+m)} J_{k+r-1}(y-1; \theta) dy = \text{(neka je } n := k+r-1) \\ & \sum_{k=0}^{\infty} (-1)^k \frac{\theta^{r+k}}{k! (r-1)!} \int_1^{\infty} y^{-(\theta+m)} \int_{C_n(y-1)} \frac{(y-1 - \sum_{i=1}^n s_i)^{\theta-1}}{s_1 \cdot \dots \cdot s_n} ds_1 \dots ds_n dy = \\ & \sum_{k=0}^{\infty} (-1)^k \frac{\theta^{r+k}}{k! (r-1)!} \int_{s_1, \dots, s_n > 1} (s_1 \cdot \dots \cdot s_n)^{-1} \int_{\sum_{i=1}^n s_i + 1}^{\infty} y^{-(\theta+m)} \left(y - 1 - \sum_{i=1}^n s_i \right)^{\theta-1} dy ds_1 \dots ds_n. \end{aligned}$$

Primijenimo li supstituciju $s = (1 + \sum_{i=1}^n s_i) \cdot y^{-1}$, $ds = -(1 + \sum_{i=1}^n s_i) \cdot y^{-2} dy$ dobivamo

$$\begin{aligned} & \int_{\sum_{i=1}^n s_i + 1}^{\infty} y^{-(\theta+m)} \left(y - 1 - \sum_{i=1}^n s_i \right)^{\theta-1} dy = \\ & \int_0^1 \frac{s^{\theta+m-2}}{(1 + \sum_{i=1}^n s_i)^{\theta+m-2}} \left(1 + \sum_{i=1}^n s_i \right)^{\theta-1} \left(\frac{1}{s} - 1 \right)^{\theta-1} \frac{ds}{1 + \sum_{i=1}^n s_i} = \\ & \frac{1}{(1 + \sum_{i=1}^n s_i)^m} \int_0^1 s^{m-1} (1-s)^{\theta-1} ds = \frac{1}{(1 + \sum_{i=1}^n s_i)^m} \frac{\Gamma(\theta) \Gamma(m)}{\Gamma(\theta+m)}, \end{aligned}$$

iz čega slijedi

$$\begin{aligned} & \int_0^{\frac{1}{r}} x^m d_r(x; \theta) dx = \\ & \frac{\Gamma(\theta) \Gamma(m)}{\Gamma(\theta+m)} \sum_{k=0}^{\infty} (-1)^k \frac{\theta^{r+k}}{k! (r-1)!} \int_{s_1, \dots, s_n > 1} (s_1 \cdot \dots \cdot s_n)^{-1} \left(1 + \sum_{i=1}^n s_i \right)^{-m} ds_1 \dots ds_n = \\ & \frac{\Gamma(\theta+1) \theta^{r-1}}{\Gamma(\theta+m) (r-1)!} \sum_{k=0}^{\infty} (-1)^k \frac{\theta^k}{k!} \int_{s_1, \dots, s_n > 1} (s_1 \cdot \dots \cdot s_n)^{-1} \Gamma(m) \left(1 + \sum_{i=1}^n s_i \right)^{-m} ds_1 \dots ds_n. \end{aligned}$$

Označimo sa $\sigma := \sum_{i=1}^n s_i$. Tada je

$$\begin{aligned} \Gamma(m) \left(1 + \sum_{i=1}^n s_i\right)^{-m} &= (1 + \sigma)^{-m} \int_0^{\infty} x^{m-1} e^{-x} dx = \left[y = \frac{x}{1 + \sigma}, dy = \frac{dx}{1 + \sigma} \right] \\ &= (1 + \sigma)^{-m} \int_0^{\infty} (1 + \sigma)^{m-1} y^{m-1} e^{-y(1+\sigma)} (1 + \sigma) dy = \int_0^{\infty} y^{m-1} e^{-y(1+\sigma)} dy. \end{aligned}$$

Zato vrijedi

$$\begin{aligned} \int_0^{\frac{1}{r}} x^m d_r(x; \theta) dx &= \\ \frac{\Gamma(\theta + 1) \theta^{r-1}}{\Gamma(\theta + m) (r-1)!} \sum_{k=0}^{\infty} (-1)^k \frac{\theta^k}{k!} \int_{s_1, \dots, s_n > 1} (s_1 \cdot \dots \cdot s_n)^{-1} \int_0^{\infty} y^{m-1} e^{-y(1+\sum_{i=1}^n s_i)} dy ds_1 \dots ds_n &= \\ \frac{\Gamma(\theta + 1) \theta^{r-1}}{\Gamma(\theta + m) (r-1)!} \sum_{k=0}^{\infty} (-1)^k \frac{\theta^k}{k!} \int_0^{\infty} y^{m-1} e^{-y} \int_{s_1, \dots, s_n > 1} \prod_{i=1}^n \frac{e^{-ys_i}}{s_i} ds_1 \dots ds_n dy &= \\ \frac{\Gamma(\theta + 1) \theta^{r-1}}{\Gamma(\theta + m) (r-1)!} \sum_{k=0}^{\infty} (-1)^k \frac{\theta^k}{k!} \int_0^{\infty} y^{m-1} e^{-y} \left(\int_1^{\infty} \frac{e^{-ys}}{s} ds \right)^n dy &= [ys = u, yds = du] \\ \frac{\Gamma(\theta + 1) \theta^{r-1}}{\Gamma(\theta + m) (r-1)!} \int_0^{\infty} y^{m-1} e^{-y} \left(\int_y^{\infty} \frac{e^{-u}}{u} du \right)^{r-1} \sum_{k=0}^{\infty} (-1)^k \frac{\theta^k}{k!} \left(\int_y^{\infty} \frac{e^{-u}}{u} du \right)^k dy &= \\ \frac{\Gamma(\theta + 1) \theta^{r-1}}{\Gamma(\theta + m) (r-1)!} \int_0^{\infty} y^{m-1} e^{-y} \left(\int_y^{\infty} \frac{e^{-u}}{u} du \right)^{r-1} \exp \left[-\theta \int_y^{\infty} \frac{e^{-u}}{u} du \right] dy &= \\ \frac{\Gamma(\theta + 1)}{\Gamma(\theta + m)} \int_0^{\infty} y^{m-1} \frac{\theta^{r-1}}{(r-1)!} \left(\int_y^{\infty} e^{-u} u^{-1} du \right)^{r-1} \exp \left[-y - \theta \int_y^{\infty} e^{-u} u^{-1} du \right] dy. \end{aligned}$$

□

Sada smo u mogućnosti pokazati sljedeći korolar.

Korolar 3.18. Funkcija $f : \mathbb{R}^r \rightarrow \mathbb{R}$,

$$f(z_1, \dots, z_r) = \theta^r \Gamma(\theta) e^{\gamma\theta} z_1^{-1} \cdot \dots \cdot z_{r-1}^{-1} z_r^{\theta-2} g_\theta \left(\frac{1 - z_1 - \dots - z_r}{z_r} \right) \mathbb{1}_{M_r}(z_1, \dots, z_r)$$

jest vjerojatnosna funkcija gustoće.

Dokaz. U teoremu 3.14 je pokazano da kada $n \rightarrow \infty$ i $n\alpha \rightarrow \theta$ funkcije gustoće slučajnog vektora $(Z_{(1)}^n, Z_{(2)}^n, \dots, Z_{(r)}^n)$ konvergiraju prema f . Zbog činjenice da je funkcija gustoće slučajnog vektora $(Z_{(1)}^n, Z_{(2)}^n, \dots, Z_{(r)}^n)$ pozitivna slijedi da je $f \geq 0$. Preostaje pokazati da je

$$\int_{\mathbb{R}^r} f(z_1, \dots, z_r) dz_1 \dots dz_r = 1.$$

Pokažimo prvo da je za svaki $r \in \mathbb{N}$

$$\int_0^{\frac{1}{r}} d_r(x; \theta) = 1.$$

U teoremu 3.17 smo za $m = 0$ dokazali da je

$$\begin{aligned} \int_0^{\frac{1}{r}} d_r(x; \theta) &= \frac{\Gamma(\theta + 1)}{\Gamma(\theta)} \int_0^\infty y^{-1} \frac{\theta^{r-1}}{(r-1)!} \left(\int_y^\infty e^{-u} u^{-1} du \right)^{r-1} \exp \left[-y - \theta \int_y^\infty e^{-u} u^{-1} du \right] dy = \\ &= \theta \int_0^\infty y^{-1} e^{-y} \frac{\left(\theta \int_y^\infty e^{-u} u^{-1} du \right)^{r-1}}{(r-1)!} \exp \left[-\theta \int_y^\infty e^{-u} u^{-1} du \right] dy. \end{aligned}$$

U slučaju $r = 1$ imamo

$$\begin{aligned} \int_0^1 d_1(x; \theta) &= \int_0^\infty \theta y^{-1} e^{-y} \exp \left[-\theta \int_y^\infty e^{-u} u^{-1} du \right] dy = \int_0^\infty \frac{\partial}{\partial y} \exp \left[-\theta \int_y^\infty e^{-u} u^{-1} du \right] dy = \\ &= e^0 - \lim_{y \rightarrow 0} \exp \left[-\theta \int_y^\infty e^{-u} u^{-1} du \right] = e^0 - e^{-\infty} = 1. \end{aligned}$$

U slučaju $r > 1$ imamo

$$\begin{aligned}
 & \int_0^\infty \theta y^{-1} e^{-y} \exp \left[-\theta \int_y^\infty e^{-u} u^{-1} du \right] dy = 1 \Rightarrow \\
 & \frac{\partial^{r-1}}{\partial \theta^{r-1}} \left(\int_0^\infty y^{-1} e^{-y} \exp \left[-\theta \int_y^\infty e^{-u} u^{-1} du \right] dy \right) = \frac{\partial^{r-1}}{\partial \theta^{r-1}} \left(\frac{1}{\theta} \right) \Rightarrow \\
 & \int_0^\infty y^{-1} e^{-y} \exp \left[-\theta \int_y^\infty e^{-u} u^{-1} du \right] (-1)^{r-1} \left(\int_y^\infty e^{-u} u^{-1} du \right)^{r-1} dy = (-1)^{r-1} (r-1)! \frac{1}{\theta^r} \Rightarrow \\
 & \int_0^{\frac{1}{\theta}} d_r(x; \theta) = \theta \int_0^\infty y^{-1} e^{-y} \frac{\left(\theta \int_y^\infty e^{-u} u^{-1} du \right)^{r-1}}{(r-1)!} \exp \left[-\theta \int_y^\infty e^{-u} u^{-1} du \right] dy = 1. \quad (3.14)
 \end{aligned}$$

U teoremu 3.15 je dokazano da je $d_r(z_r; \theta) = \int_{\mathbb{R}^{r-1}} f(z_1, \dots, z_r) dz_1 \dots dz_{r-1}$, a to zajedno s (3.14) daje

$$\int_{\mathbb{R}^r} f(z_1, \dots, z_r) dz_1 \dots dz_r = \int_0^\infty d_r(z_r; \theta) dz_r = 1.$$

□

Zaključno sa korolarom koji slijedi dokazali smo tražena svojstva Poisson-Dirichletove razdiobe.

Korolar 3.19. Neka je dan slučajni vektor $\mathbb{G}_{(r)} := (G_{(1)}, G_{(2)}, \dots)$ na prostoru Δ s Poisson-Dirichletovom razdiobom s parametrom θ . Neka je $r \in \mathbb{N}$ fiksiran. Tada je funkcija gustoće slučajnog vektora $(G_{(1)}, G_{(2)}, \dots, G_{(r)})$ funkcija $f: \mathbb{R}^r \rightarrow \mathbb{R}$,

$$f(z_1, \dots, z_r) = \theta^r \Gamma(\theta) e^{\gamma \theta} z_1^{-1} \cdot \dots \cdot z_{r-1}^{-1} z_r^{\theta-2} g_\theta \left(\frac{1 - z_1 - \dots - z_r}{z_r} \right) \mathbb{1}_{M_r}(z_1, \dots, z_r),$$

gdje je funkcija g_θ dana u (3.7), a prostor M_r u (3.5).

Funkcija gustoće slučajne varijable $G_{(r)}$ funkcija $d_r(\cdot; \theta): \mathbb{R} \rightarrow \mathbb{R}$,

$$d_r(x; \theta) = x^{\theta-2} \sum_{0 \leq k < x^{1-r}} (-1)^k \frac{\theta^{r+k}}{k! (r-1)!} J_{k+r-1} \left(\frac{1-x}{x}; \theta \right) \mathbb{1}_{\langle 0, r^{-1} \rangle}(x),$$

gdje je funkcija J_k dana u (3.6).

Za $m \in \mathbb{N}$ m -ti moment slučajne varijable $G_{(r)}$ iznosi

$$\mathbb{E}[(G_{(r)})^m] = \frac{\Gamma(\theta + 1)}{\Gamma(\theta + m)} \int_0^\infty y^{m-1} \frac{\theta^{r-1}}{(r-1)!} \left(\int_y^\infty e^{-u} u^{-1} du \right)^{r-1} \exp \left[-y - \theta \int_y^\infty e^{-u} u^{-1} du \right] dy. \quad (3.15)$$

Dokaz. U korolaru 3.9 je pokazano da $Z_{(i)}^n \implies \mathbb{G}_{(i)}$ kada $n \rightarrow \infty$ i $n \rightarrow \infty$. Iz toga slijedi da za fiksni r funkcija gustoće slučajnog vektora $(Z_{(1)}^n, Z_{(2)}^n, \dots, Z_{(r)}^n)$ konvergira k funkciji gustoće slučajnog vektora $(G_{(1)}, G_{(2)}, \dots, G_{(r)})$ kada $n \rightarrow \infty$. U teoremu 3.14 je dokazano da funkcija gustoće slučajnog vektora $(Z_{(1)}^n, Z_{(2)}^n, \dots, Z_{(r)}^n)$ konvergira k funkciji f . Budući da je u korolaru 3.18 pokazano da je f vjerojatnosna funkcija gustoće zaključujemo da je f funkcija gustoće slučajnog vektora $(G_{(1)}, G_{(2)}, \dots, G_{(r)})$.

Budući da je funkcija f funkcija gustoće slučajnog vektora $(G_{(1)}, G_{(2)}, \dots, G_{(r)})$, a funkcija $d_r(\cdot; \theta)$ po teoremu 3.15 r -ta marginalna gustoća od f i po dokazu korolara 3.18 vjerojatnosna funkcija gustoće slijedi da je $d_r(\cdot; \theta)$ funkcija gustoće slučajne varijable $G_{(r)}$.

Budući da je $d_r(\cdot; \theta)$ funkcija gustoće slučajne varijable $G_{(r)}$, a po teoremu 3.17 za proizvoljni $m \in \mathbb{N}$ vrijedi

$$\int_0^{\frac{1}{r}} x^m d_r(x; \theta) dx = \frac{\Gamma(\theta + 1)}{\Gamma(\theta + m)} \int_0^\infty y^{m-1} \frac{\theta^{r-1}}{(r-1)!} \left(\int_y^\infty e^{-u} u^{-1} du \right)^{r-1} \exp \left[-y - \theta \int_y^\infty e^{-u} u^{-1} du \right] dy$$

dobivamo traženu tvrdnju za $\mathbb{E}[(G_{(r)})^m]$. \square

Poglavlje 4

Zaključak

U poglavlju 2 proučavali smo dva primjera slučajnih vektora za koje smo pokazali da konvergiraju prema distribuciji koju nazivamo Poisson-Dirichletova. U poglavlju 3 dokazali smo svojstva te distribucije, tako da smo sada u mogućnosti reći nešto više o graničnom ponašanju dugih ciklusa i velikih djelitelja.

Korolar 4.1. Neka je $|C_{(r)}|$ slučajna varijabla koja odgovara duljini r -tog po veličini ciklusa permutacije n -članog skupa. Kada $n \rightarrow \infty$ slučajna varijabla $\frac{|C_{(r)}|}{n}$ konvergira k slučajnoj varijabli s funkcijom gustoće $d_r(\cdot; 1)$ danoj u (3.13) i m -tim momentima danima u (3.15) za $\theta = 1$.

Dokaz. Po napomeni 3.3 slučajni vektor $\mathbb{L}_{(\cdot)}^n$ konvergira prema Poisson-Dirichletovoj razdiobi s parametrom 1. Budući da je $\frac{|C_{(r)}|}{n}$ r -ta komponenta vektora $\mathbb{L}_{(\cdot)}^n$ po korolaru 3.19 vrijedi tražena tvrdnja uz parametar $\theta = 1$. \square

Korolar 4.2. Neka je P_r slučajna varijabla koja odgovara r -tom po veličini prostom faktoru prirodnog broja $N_n \leq n$. Kada $n \rightarrow \infty$ slučajna varijabla $\frac{\log(P_r)}{\log(t(N_n))}$, gdje je funkcija t dana u (2.3), konvergira k slučajnoj varijabli s funkcijom gustoće $d_r(\cdot; 1)$ danoj u (3.13) i m -tim momentima danima u (3.15) za $\theta = 1$.

Dokaz. Slučajni vektor $\left(\frac{\log(P_1)}{\log(T_n)}, \frac{\log(P_2)}{\log(T_n)}, \dots\right)$ konvergira prema Poisson-Dirichletovoj razdiobi s parametrom 1, ponovno po napomeni 3.3. Budući da je $t(N_n) = T_n$ i $\frac{\log(P_r)}{\log(t(N_n))}$ r -ta komponenta vektora $\left(\frac{\log(P_1)}{\log(T_n)}, \frac{\log(P_2)}{\log(T_n)}, \dots\right)$ po korolaru 3.19 vrijedi tražena tvrdnja uz parametar $\theta = 1$. \square

Promotrimo detaljnije što se dobiva iz prethodna dva korolara u slučaju $r = 1$ odnosno za najdulji ciklus i najveći prosti faktor. Kada $n \rightarrow \infty$ slučajna varijabla $\frac{|C_{(1)}|}{n}$ i slučajna

varijabla $\frac{\log(P_1)}{\log(t(N_n))}$ konvergiraju k slučajnoj varijabli s funkcijom gustoće

$$d_1(x; 1) = x^{-1} \sum_{0 \leq k < x^{-1} - 1} \frac{(-1)^k}{k!} \int_{\substack{s_1, \dots, s_k > 1 \\ \sum_{i=1}^k s_i < \frac{1-x}{x}}} \frac{ds_1 \dots ds_k}{s_1 \cdot \dots \cdot s_k}, \quad (4.1)$$

gdje se u slučaju $k = 0$ za vrijednost integrala uzima 1.

Za niz slučajnih varijabli $\frac{|C_{(1)}|}{n}$ vrijedi

$$\sup_n \mathbb{E} \left[\left| \frac{|C_{(1)}|}{n} \right|^{1+\varepsilon} \right] \leq 1 < \infty$$

čime je zadovoljen uvjet korolara 1.18. Iz čega možemo zaključiti da je za dovoljno veliki n očekivana relativna duljina najduljeg ciklusa slučajne permutacije približno dana formulom

$$\begin{aligned} \mathbb{E} \left[\frac{|C_{(1)}|}{n} \right] &\approx \frac{\Gamma(1+1)}{\Gamma(1+1)} \int_0^\infty y^{1-1} \frac{1^{1-1}}{(1-1)!} \left(\int_y^\infty e^{-u} u^{-1} du \right)^{1-1} \exp \left[-y - \int_y^\infty e^{-u} u^{-1} du \right] dy = \\ &\int_0^\infty e^{-y} \exp \left[- \int_y^\infty e^{-u} u^{-1} du \right] dy \approx 0.624. \end{aligned}$$

Vjerojatnosti $\mathbb{P}[G_{(r)} > x]$ mogu se numerički izračunati i tabelirati za različite parametre θ . Slučaj koji nas najviše zanima za $\theta = 1$ i $r = 1$ po Griffiths [3] izračunat na četiri decimale dan je u tablici 4.1.

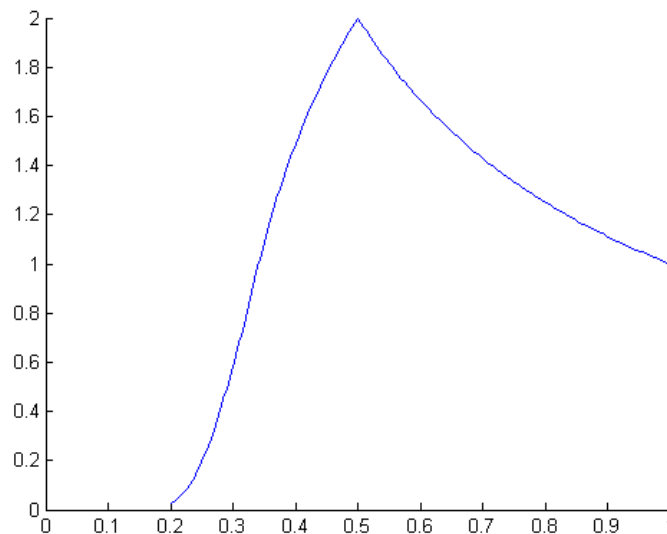
x	0.10	0.15	0.20	0.25	0.30	0.35
$\mathbb{P}[G_{(1)} > x]$	1.0000	1.0000	0.9996	0.9951	0.9763	0.9346
x	0.40	0.45	0.50	0.55	0.60	0.65
$\mathbb{P}[G_{(1)} > x]$	0.8697	0.7878	0.6931	0.5978	0.5108	0.4308
x	0.70	0.75	0.80	0.85	0.90	0.95
$\mathbb{P}[G_{(1)} > x]$	0.3567	0.2877	0.2231	0.1625	0.1054	0.0513

Tablica 4.1: Distribucija slučajne varijable $G_{(1)}$.

Ponovno za dovoljno veliki n možemo po podacima iz tablice 4.1 zaključiti da je

$$\begin{aligned} \mathbb{P} \left[\frac{\log(P_1)}{\log(T_n)} > 0.8 \right] &\approx 0.2231 \Rightarrow \mathbb{P} [P_1 > T_n^{0.8}] \approx 0.2231, \\ \mathbb{P} \left[\frac{\log(P_1)}{\log(T_n)} \leq 0.2 \right] &\approx 1 - 0.9996 \Rightarrow \mathbb{P} [P_1 \leq T_n^{0.2}] \approx 0.0004. \end{aligned}$$

Prirodno se postavlja pitanje što znači dovoljno veliki n . Na to pitanje nećemo dati formalan odgovor, ali ga intuitivno možemo naslutiti vizualnim pristupom. Na slici 4.1 prikazan je graf funkcije dane u (4.1) prema kojoj po korolaru 4.2 konvergira slučajna varijabla $\frac{\log(p_1)}{\log(T_n)}$. Budući da je za određivanje vrijednosti funkcije gustoće za $x < 0.2$ potrebno računati peterostruke i više integrale što je numerički vrlo zahtjevno graf je prikazan za $x \in [0.2, 1]$.

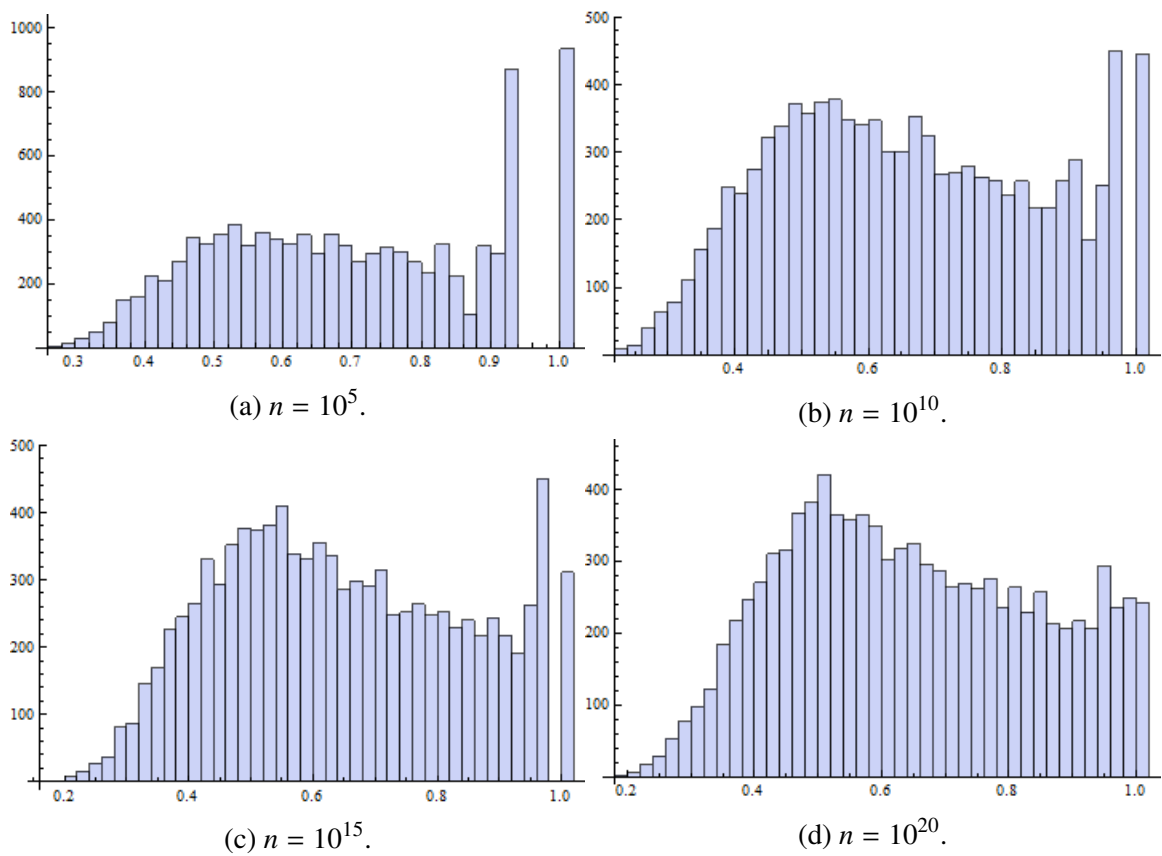


Slika 4.1: Graf funkcije $d_1(x; 1)$, za $x \in [0.2, 1]$.

Kako bismo odredili ponašanje varijable $\frac{\log(p_1)}{\log(T_n)}$ za konkretan n promatrali smo slučajne uzorke od po $m = 1000$ elemenata i za dobivene realizacije crtali histograme.

Tako su na slici 4.2 prikazani histograni dobiveni u slučaju $n = 10^5$, $n = 10^{10}$, $n = 10^{15}$ i $n = 10^{20}$ pomoću programskog paketa Wolfram Mathematica [5]. Uočimo da se na svim grafovima može uočiti rast otprilike do vrijednosti 0.5 kao što i očekujemo s obzirom na graf sa slike 4.1, i pad nakon 0.5. Ono što na grafovima za $n < 10^{20}$ posebno odskače od očekivanog jesu slučajevi kada je $\frac{\log(p_1)}{\log(t_n)} > 0.9$ što se događa na primjer u slučaju kada je $t_n = p_1$, odnosno kada je nasumično odabran broj prost ili potencija prostog. Uočimo da taj utjecaj postupno slabi kako n raste i da za $n = 10^{20}$ histogram vrlo dobro odgovara grafu sa slike 4.1. Budući da se to postiže tek za $n = 10^{20}$ zaključujemo da je brzina konvergencije vrlo spora.

S ovih nekoliko od mnogobrojnih primjena rezultata dobivenih u ovom radu zaokružena je tema razdiobe dugih ciklusa i velikih djelitelja.



Slika 4.2: Histogrami vrijednosti $\frac{\log(P_1)}{\log(T_n)}$ za različite n i za 10000 realizacija slučajnih varijabli P_1 i T_n .

Bibliografija

- [1] P. Billingsley, *Convergence of Probability Measures*, A Wiley-Interscience Publication, 1999.
- [2] A. Dujella, *Uvod u teoriju brojeva*, Skripta, PMF - Matematički odjel, Sveučilište u Zagrebu, <http://web.math.pmf.unizg.hr/duje/utb/utblink.pdf>.
- [3] R.C. Griffiths, *On the Distribution of Points in a Poisson Dirichlet Process*, Journal of Applied Probability **25**, br. 2, 336–345, 1988.
- [4] N. Sarapa, *Teorija vjerojatnosti*, Školska knjiga, 2002.
- [5] Inc. Wolfram Research, *Mathematica*, verzija 9.0, Champaign, IL, 2012.

Sažetak

Jedan od ciljeva ovog diplomskog rada bio je definirati Poisson-Dirichletovu razdiobu i proučiti njezina glavna svojstva. Uvedeni su pojmovi iz područja matematičke analize, teorije vjerojatnosti i teorije brojeva nužni za razumijevanje kasnijih rezultata, a neka njihova svojstva su i dokazana. Motivacija i opravdanost definicije Poisson-Dirichletove razdiobe ilustrirani su kroz dva primjera u drugom poglavlju. Detaljno su obrazložena i formalno pokazana asimptotska svojstva slučajnih vektora koji odgovaraju relativnim duljinama ciklusa slučajnih permutacija, odnosno logaritamskih vrijednosti prostih djeljitelja slučajnog prirodnog broja manjeg od n . U trećem poglavlju izrečeni su i dokazani glavni rezultati razmatrani u ovom radu. Povezana je funkcija gustoće Poisson-Dirichletove razdiobe, njezine marginalne gustoće i momenti s asimptotskim svojstvima dugih ciklusa i velikih djeljitelja. U zadnjem poglavlju su ti rezultati primijenjeni na konkretnim podacima te je ilustrirana brzina konvergencije distribucija dugih ciklusa i velikih djeljitelja za velik, ali konačan n .

Summary

One of the aims of this thesis was to define Poisson-Dirichlet distribution and study its main properties. Concepts from the fields of Analysis, Probability theory and Number theory are introduced and some of their properties proven in order to fully understand later results. The motivation and justification for the definition of Poisson-Dirichlet distribution are illustrated by two examples in Chapter 2. Asymptotic properties of random vectors corresponding to relative lengths of cycles in random permutations and logarithmic values of prime divisors of random natural numbers smaller than n are formally proven and explained in detail. The main results discussed in this thesis are stated and proven in Chapter 3. The relation between probability density function of Poisson-Dirichlet distribution, its marginal densities and moments and asymptotic properties of long cycles and great divisors is proven. In concluding chapter these results are applied to specific data, we also illustrate the speed of convergence of distributions of long cycles and great divisors for large, but finite n .

Životopis

Rođena sam 17. lipnja 1991. godine u Zagrebu. Nakon završetka Osnovne škole Vladimira Nazora upisala sam XV. gimnaziju u Zagrebu. 2010. godine upisala sam preddiplomski sveučilišni studij Matematika na Prirodoslovno-matematičkom fakultetu u Zagrebu, a 2013. godine diplomski sveučilišni studij Teorijska matematika na istom fakultetu. Tokom studija nagrađena sam trećom nagradom na međunarodnom studentskom natjecanju IMC 2012. godine, Priznanjem za izniman uspjeh u studiju 2013. i 2015. godine te Rektorovom nagradom 2015. godine.