UNIVERZA V LJUBLJANI
FAKULTETA ZA MATEMATIKO IN FIZIKO

Matematika – 2. stopnja

Erik Langerholc

KOPULE S PREDPISANIMI SINGULARNOSTMI

Magistrsko delo

Mentor: prof. dr. Matjaž Omladič

Ljubljana, 2019

Kazalo

Program dela

Kopule so postale v zadnjih desetletjih ključno orodje za študij odvisnosti slučajnih spremenljivk, saj nam Sklarov izrek zagotavlja univerzalno uporabnost, namreč, da je možno z njimi modelirati vsako (!) stohastično odvisnost. Uporabe se širijo v različna področja naravoslovja in tehnike, pa tudi v bio vede, ekonomijo in kvantitativne finance. Še posebej je postalo pomembno modeliranje tveganj s pomočjo kopul. Brez tega orodja postane skoraj edina možnost kvantitativne ocene odvisnosti "koreliranost", kar pomeni zgolj linearno odvisnost. Tak pristop je sicer bolj enostaven, vendar le redko lahko zadovolji vse zapletene praktične scenarije. Tudi modeliranje s kopulami je lahko bolj ali manj zahtevno. Med primeri, ki se jih v aplikacijah najbolj "bojijo", so modeli singularnih kopul. Glavni cilj magistrskega dela naj bo konstrukcija kopul s singularnimi komponentami na predpisanih množicah in izreki v zvezi s tem. Dobro bi bilo delo uvodoma opremiti z osnovnimi dejstvi o kopulah, kot je npr. Sklarov izrek. Za dobro razumevanje dela bi bilo pred tem dobro pripraviti tudi kratek uvod v teorijo mere.

Osnovna literatura

[2] F. Durante, J. Fernandez-Sanchez in W. Trutschnig, *On the singular components of a copula*, Journal of Applied Probability **52**(4) (2015) 1175–1182

Podpis mentorja:

Kopule s predpisanimi singularnostmi

POVZETEK

Delo vsebuje pregled teorije mere za potrebe verjetnosti in singularnih kopul, osnove teorije kopul ter podroben dokaz Sklarovega izreka. Predstavljena so tudi Markovska jedra in dve konstrukciji kopul s predpisanimi singularnostmi.

Copulas with prescribed singularities

ABSTRACT

Contents of the present thesis are an overview of measure theory for the purposes of working with probability theory and singular copulas, fundamentals of copula theory and a detailed proof of Sklar's theorem, followed by an introduction to Markov kernels and two constructions of copulas with prescribed singularities.

Math. Subj. Class. (2010): 60E05

Ključne besede: kopula, Sklarov izrek, singularna mera

Keywords: copula, Sklar's theorem, singular measure

1 Uvod

V verjetnosti in statistiki so osrednji predmet študija slučajne spremenljivke. Mnogokrat se zaradi računskih potreb različne spremenljivke vzame kot neodvisne, a običajno je to kvečjemu dober približek pojava iz resničnega sveta, kjer imajo različni dejavniki lahko znaten medsebojni vpliv. Študij odvisnosti je zato še kako pomemben.

Odvisnost slučajnih spremenljivk pa vodi do pojma kopul – standardne oblike za opis medsebojne odvisnosti. Njihova obravnava se je dobro začela šele v drugi polovici prejšnjega stoletja, kljub temu pa so danes zelo uporabljano orodje v številnih disciplinah, predvsem pa v finančni matematiki, katere zanimanje za kopule je močno pospešilo razvoj teorije od konca tisočletja dalje.

Kljub temu so v svetu kopul marsikatera vprašanja še nerazrešena. Na primer, kopule s singularnimi deli (tj. brez pripadajoče gostote) so nujne, a bolj redko uporabljane, pri obravnavi določenih pojavov iz narave, razlog pa je preprosto to, da niso tako enostavne za uporabo. Celo do te mere, da ni čisto jasno, kakšni singularni deli sploh lahko porodijo kopulo in kakšna je v tem primeru konstrukcija.

V tem delu si bomo pogledali osnove teorije kopul in kako skonstruirati kopule z določenimi singularnimi deli. Obravnava bo natančna in primerno splošna, zato je v drugem poglavju nujna obravnava teorije mere, na kateri današnja verjetnost sloni. V tretjem poglavju sledi ponovitev in posplošitev znanih verjetnostnih pojmov ter osnove kopul. V polni splošnosti je dokazan Sklarov izrek, ki zagotavlja, da so kopule res standardno orodje za obravnavo vseh primerov medsebojne odvisnosti slučajnih spremenljivk. V zadnjem poglavju obravnavamo Markovska jedra kot posplošitev enostavnih pogojnih porazdelitev in nato zaključimo z dvema konstrukcijama kopul s predpisanimi singularnimi deli.

2 Verjetnost z mero

2.1 Oznake

Uporabljene oznake bodo običajne oznake teorije verjetnosti.

- Minimum dveh števil ali funkcij a, b se namesto z $\min(a, b)$ označi krajše z
$$a \wedge b,$$
minimum več objektov $a_1, \ldots, a_n$ pa z
$$\bigwedge_{i=1}^{n} a_i.$$
Podobno se uporablja znak $\vee$ za maksimum.

- Pozitivni in negativni del funkcije se označi z
$$f^+ = f \vee 0 \qquad f^- = (-f) \vee 0.$$
Velja torej $f = f^+ - f^-$.

- Indikatorska funkcija $\mathbb{1}_A$ množice A je preslikava, ki s številom pove, ali je njen argument element A. Natančneje:

$$\mathbb{1}_A(x) = \begin{cases} 1 & \text{če } x \in A \\ 0 & \text{sicer} \end{cases}$$

  Indikatorske funkcije so zelo uporabne za opis funkcij, ki so podane z različnim predpisom na različnih množicah, brez potrebe po ločevanju primerov.

- Za preslikavo $f : A \to B$ z $f(x)$ označimo sliko elementa $x \in A$ in z $f[M]$ sliko množice $M \subseteq A$

$$f[M] = \{f(x) \mid x \in M\},$$

  z $f^{-1}[N]$ pa prasliko množice $N \subseteq B$

$$f^{-1}[N] = \{x \in A \mid f(x) \in N\}.$$

- Za funkciji f in g definiramo relacije in operacije po komponentah, na primer $f < g$ pomeni, da je $f(x) < g(x)$ za vsak x, $(f + g)$ pa je funkcija

$$x \longmapsto f(x) + g(x).$$

- Uporabljena bo tudi oznaka "s pikco", tj. funkcijo f lahko označimo tudi z

$$f(\cdot) \qquad \text{namesto z} \qquad x \longmapsto f(x).$$

  To je uporabno predvsem pri kompozicijah funkcij ali pa, ko imajo funkcije več argumentov in želimo nekatere izmed njih fiksirati.

2.2 Motivacija

Osnove matematične verjetnosti obravnavajo enostavne verjetnostne porazdelitve:

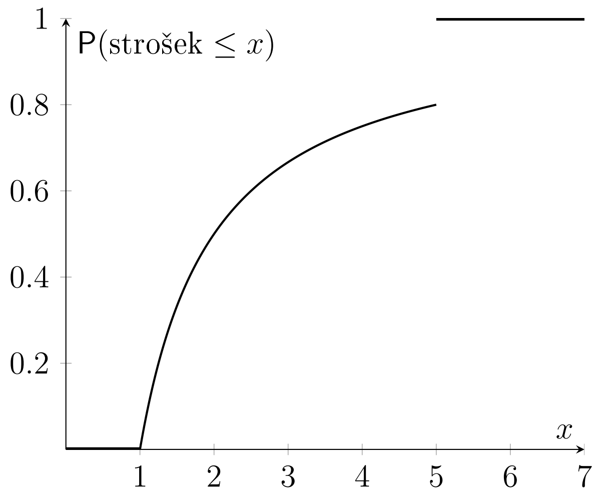
- Zvezne: to so porazdelitve, kot je npr. normalna, ki dani množici izidov dodeli verjetnost enako integralu določene funkcije po tej množici. Porazdelitvena funkcija zvezne porazdelitve je zvezna.

- Diskretne: to so porazdelitve, kjer je možnih izidov števno mnogo, npr. binomska. Porazdelitvena funkcija diskretne porazdelitve je odsekoma konstantna.

Obravnavanje le teh dveh tipov porazdelitev pa ni dovolj niti iz praktičnega niti iz teoretičnega vidika.

Primer 2.1. Zavarovalnica dobi škodni zahtevek v višini Pareto$(1, 1)$-porazdeljene slučajne spremenljivke. V pogodbi s stranko je določeno, da zavarovalnica krije škodo največ do zneska 5, preostanek plača stranka. Kako modeliramo strošek zavarovalnice?

Po definiciji je slučajna spremenljivka Pareto$(1, 1)$-porazdeljena, če je z verjetnostjo $\frac{1}{x} \wedge 1$ večja od x za poljubno pozitivno število x. To pomeni, da ima z



Slika 1: Porazdelitvena funkcija stroška zavarovalnice.

verjetnostjo $\frac{1}{5}$ zavarovalnica strošek 5, sicer pa je strošek zvezno porazdeljen na intervalu $[1, 5]$. Porazdelitvena funkcija stroška zavarovalnice je prikazana na sliki 1. Očitno ne spada niti pod diskretne niti pod zvezne porazdelitve – ker ima skok, ni integral nobene funkcije in, ker ni odsekoma konstantna, ima neštevno mnogo možnih vrednosti. Je kombinacija zvezne in diskretne porazdelitve.

Ena možnost obravnavanja porazdelitve iz primera 2.1 je, da za njeno gostoto proglasimo

$$\mathbf{1}_{[1,5)} f + \frac{1}{5}\delta_5.$$

Tu f označuje gostoto Pareto$(1, 1)$ porazdelitve in δ_5 Diracovo delta "funkcijo" na točki 5. Ta pristop bi pravzaprav dobro deloval v tem primeru in lahko bi ga uporabili vedno, ko obravnavamo porazdelitve, ki so kombinacija zveznih in diskretnih. Ampak niti to ni dovolj – poglejmo si še en primer.

Primer 2.2. Pri igri na srečo mečemo pošten kovanec v nedogled. Če pri n-tem metu pade cifra, zaslužimo 4^{-n} denarne enote. Potem porazdelitev skupnega izkupička ni ne diskretna, ne zvezna in niti kombinacija obeh [5, str. 71].

Res je, da v praksi nikoli ne moremo vreči kovanca neskončnokrat. Ampak velikokrat je končne pojave iz narave lažje modelirati z njihovo neskončno verzijo. Poleg tega pa so, kot bomo videli v primeru 4.3, v več razsežnostih tudi zelo naravne porazdelitve take, ki niso kombinacija zveznih in diskretnih. Zato je pomembno, da zna teorija verjetnosti natančno opisati tudi tovrstne porazdelitve. Prevladujoča teorija verjetnosti, ki to zmore, pa je Kolmogorova verjetnost, ki je osnovana na teoriji mere.

2.3 Merljivi prostori in preslikave

Začnimo z osnovami teorije mere. Obravnavali bomo le definicije in rezultate, ki jih bomo potrebovali pri študiju singularnih kopul in tudi mnogi dokazi bodo izpuščeni. Za vse manjkajoče dele priporočam ogled knjige [5], kjer so obdelani z vidika verjetnosti.

Z Ω bomo vedno označevali množico *izidov*, tj. množico vseh možnih rezultatov poizkusa, ki ga modeliramo. Komplemente množic bomo vedno gledali glede na Ω.

Definicija 2.3. *Sigma algebra (σ-algebra)* na Ω je družina $\mathcal{A}$ množic izidov, ki vsebuje množico vseh izidov Ω in je:

- Zaprta za komplemente: če je $A \in \mathcal{A}$, je tudi $A^{\mathsf{c}} \in \mathcal{A}$.
- Zaprta za števne unije: če so $A_1, A_2, \ldots$ množice iz $\mathcal{A}$, je tudi njihova unija.

Elementom σ-algebre pravimo *merljive množice*, z vidika verjetnosti pa *dogodki*. Paru $(\Omega, \mathcal{A})$ pravimo *merljiv prostor*.

S σ-algebro povemo, za katere množice izidov sploh lahko poznamo verjetnost. Z različnimi σ-algebrami na isti množici izidov Ω lahko modeliramo različne količine informacij o poizkusu pod različnimi pogoji (npr. z vidika različnih časov). Omejevanje merljivosti množic pa je pogosto tudi nujno – če bi vedno lahko določili verjetnost vsake množce izidov, bi prišlo do protislovij. Več o tem v razdelku 2.4.

Iz definicije σ-algebre sledijo tudi zaželene osnovne lastnosti: vedno poznamo verjetnost polnega dogodka (vseh izidov skupaj, tj. Ω), ki je 1, in, če lahko poznamo verjetnost nekega dogodka, potem lahko poznamo tudi verjetnost komplementarnega dogodka.

Poglejmo si nekaj naravnih σ-algeber. Lahko je videti, da je presek poljubne zbirke σ-algeber spet sigma algebra, zato definiramo

Definicija 2.4. Naj bo $\mathcal{D}$ družina podmnožic množice izidov Ω. Najmanjša σ-algebra na Ω (glede na relacijo vsebovanja $\subseteq$), ki vsebuje vse množice iz $\mathcal{D}$, je presek vseh σ-algeber na Ω, ki vsebujejo vse množice iz $\mathcal{D}$. Pravimo, da je *generirana* z $\mathcal{D}$ in jo označimo s $\sigma(\mathcal{D})$.

V topološkemu prostoru (X, τ) je naravna σ-algebra tista, ki je generirana z vsemi odprtimi množicami τ. Pravimo, da je *Borelova* on jo označimo z $\mathcal{B}_X$.
Ker so σ-algebre zaprte za komplemente, Borelova σ-algebra poleg odprtih vsebuje tudi vse zaprte množice. V primeru realnih števil je to že zelo velik in uporaben razred množic. Zato se, če ni rečeno drugače, realna števila (oz. $\mathbb{R}^n$) vedno opremi s pripadajočo Borelovo σ-algebro. V nadaljevanju se bomo držali tudi dogovora, da se splošni množici reče Borelova oz. n-razsežna Borelova, če je iz $\mathcal{B}_{\mathbb{R}}$ oz. $\mathcal{B}_{\mathbb{R}^n}$, in splošni σ-algebri reče Borelova oz. n-razsežna Borelova, če je $\mathcal{B}_{\mathbb{R}}$ oz. $\mathcal{B}_{\mathbb{R}^n}$.

Še en naraven način tvorjenja σ-algeber pa je produkt. Naj bosta $(\Omega_1, \mathcal{A}_1)$ in $(\Omega_2, \mathcal{A}_2)$ merljiva prostora. Potem lahko na produktu $\Omega_1 \times \Omega_2$ definiramo produktno σ-algebro $\mathcal{A}_1 \times \mathcal{A}_2$ kot najmanjšo σ-algebro, ki vsebuje vse merljive pravokotnike, tj.

$$\mathcal{A}_1 \times \mathcal{A}_2 = \sigma\big(\{A_1 \times A_2 \mid A_1 \in \mathcal{A}_1, a_2 \in \mathcal{A}_2\}\big).$$

Pokazati se da, da je produkt Borelovih σ-algeber na $\mathbb{R}^n$ in $\mathbb{R}^m$ ravno Borelova σ-algebra na $\mathbb{R}^{n+m}$.

Po določitvi, katerim množicam sploh lahko pripišemo verjetnost, pa je čas, da to dejansko storimo.

Definicija 2.5. *Mera* na σ-algebri je preslikava iz σ-algebre v $[0, \infty]$, ki prazni množici dodeli vrednost 0 in je števno aditivna, tj. mera števne unije disjunktnih množic je enaka vsoti mer vsake od teh množic posebej.
Mera je *σ-končna*, če se množica izidov Ω lahko zapiše kot števno unijo množic s končno mero, *končna*, če mera slika v $[0, \infty)$ namesto $[0, \infty]$ in *verjetnostna*, če množici Ω dodeli vrednost 1. Merljivemu prostoru skupaj z mero na njem pravimo *prostor z mero*, če je mera verjetnostna, pa *verjetnostni prostor*.

Splošno mero običajno označimo z μ, verjetnostno mero pa s P. Od zdaj naprej je verjetnostna mera sopomenka za porazdelitev. To je matematični objekt, ki določa, kako verjeten je posamezni dogodek. Motivacija za tako definicijo je jasna: vsaka porazdelitev dodeli verjetnost 0 praznemu dogodku (prazni množici izidov), 1 polnemu in je števno aditivna.

Primer 2.6. Naj bo $(\Omega, \mathcal{A}, \mathsf{P})$ verjetnostni prostor in A, B disjunktna dogodka. Po definiciji mere je $\mathsf{P}(A) + \mathsf{P}(B) = \mathsf{P}(A \cup B)$ in posledično

$$\mathsf{P}(A) = \mathsf{P}(A \cup B) - \mathsf{P}(B).$$

Ta enakost ne bi nujno veljala, če prostor ne bi bil verjetnostni, saj bi lahko bilo $\mathsf{P}(A \cup B) = \mathsf{P}(B) = \infty$ in bi prišlo do odštevanja $\infty - \infty$, česar pa ni možno smiselno definirati.

Najprej najštejmo tri uporabne osnovne lastnosti vsake mere.

Trditev 2.7. *Naj bo μ mera. Potem velja:*

- (Monotonost mere). *Če sta $A \subseteq B$ merljivi množici, je $\mu(A) \leq \mu(B)$.*
- (Subaditivnost mere). *Če so $A_1, A_2, \ldots$ (ne nujno disjunktne) merljive množice, je*

$$\mu\left(\bigcup_{n \in \mathbb{N}} A_n\right) \leq \sum_{n \in \mathbb{N}} \mu(A_n).$$

- (Zveznost mere). *Če je $A_1, A_2, \ldots$ naraščujoče (glede na relacijo vsebovanja $\subseteq$) zaporedje merljivih množic, je*

$$\mu\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \lim_{n \to \infty} \mu(A_n).$$

  Podobno, če je zaporedje padajoče in je $\mu(A_1) < \infty$,[1] *velja*

$$\mu\left(\bigcap_{n \in \mathbb{N}} A_n\right) = \lim_{n \to \infty} \mu(A_n).$$

[1] S tem pogojem se izognemo odštevanju $\infty - \infty$ v dokazu trditve.

Primer 2.8. Z δ_ω označimo mero *delta* na izidu $\omega \in \Omega$, podano z

$$\delta_\omega(A) = \mathbf{1}_A(\omega).$$

Mera delta opiše verjetnostno porazdelitev, ki dodeli verjetnost 1 vsakemu dogodku, ki vsebuje izbrano točko, in 0 vsem, ki je ne. Torej mera delta predstavlja poizkus, pri katerem vedno dobimo isti izid. Temu pravimo tudi, da je vsa verjetnost skoncentrirana v eni točki.

S sestavljanjem delta mer pa lahko opišemo poljubno diskretno porazdelitev. Npr. za množico izidov $\Omega = \{\text{cifra}, \text{grb}\}$ mera $\frac{1}{2}\delta_{\text{cifra}} + \frac{1}{2}\delta_{\text{grb}}$ predstavlja met poštenega kovanca. Bolj splošno: če je $\Omega = \{\omega_1, \omega_2, \dots\}$ množica vseh možnih izidov in so $p_1, p_2, \dots$ njihove verjetnosti, potem poizkus opišemo z verjetnostnim prostorom $\{\Omega, \mathcal{P}\Omega, \sum_{n\in\mathbb{N}} p_n \delta_{\omega_n}\}$.[2]

Na produktih prostorov z mero vedno obstaja produktna mera, kot pravi naslednja trditev.

Trditev 2.9. *Naj bosta $(\Omega_1, \mathcal{A}_1, \mu_1)$ in $(\Omega_2, \mathcal{A}_2, \mu_2)$ prostora z mero. Potem obstaja enolična mera $\mu_1 \times \mu_2$ na produktnii σ-algebri $\mathcal{A}_1 \times \mathcal{A}_2$, ki vsakemu merljivemu pravokotniku dodeli vrednost enako produktu mer stranic, tj. velja*

$$(\mu_1 \times \mu_2)(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2)$$

za vse merljive množice $A_1 \in \mathcal{A}_1$ in $A_2 \in \mathcal{A}_2$.

Zdaj, ko poznamo dva najbolj temeljna pojma teorije mere, se lahko lotimo obravnave slučajnih spremenljivk v splošnejši obliki.

Definicija 2.10. Naj bosta $(\Omega, \mathcal{A})$ in $(\Psi, \mathcal{B})$ merljiva prostora. Preslikava $f : \Omega \to \Psi$ je *merljiva*, če je $X^{-1}[B] \in \mathcal{A}$ za vsak $B \in \mathcal{B}$, tj. praslika vsake merljive množice je merljiva.
V smislu verjetnosti je *slučajna spremenljivka* (okrajšava *s.s.*) merljiva preslikava iz verjetnostnega prostora v prostor realnih števil.

Slučajna spremenljivka torej vsak izid opiše z realnim številom. Število je velikokrat bolj uporabno ali pa omogoča lažjo predstavo o problemu, kot pa izid sam.

Naj bo X slučajna spremenljivka[3] in B Borelova množica. Množico $X^{-1}[B]$ oziroma

$$\{\omega \in \Omega \,|\, X(\omega) \in B\}$$

na krajše označimo z

$$\{X \in B\}.$$

Če je X slučajna spremenljivka, to pomeni, da je $\{X \in B\}$ dogodek (tj. tej množici izidov lahko pripišemo verjetnost) za vsako Borelovo množico B (tj. za vsako praktično uporabno podmnožico realnih števil). Zato v verjetnosti zahtevamo, da so vse preslikave slučajne spremenljivke. Nadalje okrajšamo še $\{X \in B\} \cap \{X \in C\}$ v

$$\{X \in B, X \in C\}$$

[2] $\mathcal{P}$ je simbol za potenčno množico.

[3] Če ni rečeno drugače, je vsaka slučajna spremenljivka definirana na poljubnem verjetnostnem prostoru $(\Omega, \mathcal{A}, \mathsf{P})$.

in $\mathsf{P}\big(\{X \in B\}\big)$ v

$$\mathsf{P}(X \in B).$$

Enake oznake lahko uporabljamo tudi za splošne merljive preslikave.

Sedaj pokažimo, da so mnoge preslikave merljive. V resnici se je potrebno kar potruditi, če želimo najti nemerljivo preslikavo $\mathbb{R} \to \mathbb{R}$.

Trditev 2.11. *Kompozicija merljivih preslikav je merljiva. Simbolično:*

$$(\Omega, \mathcal{A}) \xrightarrow[\text{merljiva}]{g} (\Psi, \mathcal{B}) \xrightarrow[\text{merljiva}]{f} (\Phi, \mathcal{C}) \quad \Longrightarrow \quad (\Omega, \mathcal{A}) \xrightarrow[\text{merljiva}]{f \circ g} (\Phi, \mathcal{C})$$

Naj bo $f : \mathbb{R} \to \mathbb{R}$ merljiva funkcija in X slučajna spremenljivka. Potem je $f \circ X$ po trditvi 2.11 tudi slučajna spremenljivka. Za večjo predstavljivost se $f \circ X$ lahko označi z $f(X)$. Pisali bomo tudi $\mathsf{P}(X \leq a)$ namesto $\mathsf{P}\big(X \in (-\infty, a)\big)$ in podobno.

Če je A dogodek, je indikatorska funkcija $\mathbb{1}_A$ slučajna spremenljivka, saj je $\mathbb{1}_A^{-1}[B] \in \{\emptyset, A, A^{\mathsf{c}}, \Omega\}$ za poljubno Borelovo množico B. Pri preverjanju merljivosti manj enostavnih preslikav pa nam je v pomoč naslednja trditev.

Trditev 2.12.

- *Naj bo f preslikava s kodomeno $\mathbb{R}$. Če je za vsak $a \in \mathbb{R}$ množica $\{f \leq a\}$ merljiva, je f merljiva.*
- *Vsaka zvezna funkcija $\mathbb{R} \to \mathbb{R}$ je merljiva.*
- *Vsota in produkt dveh merljivih funkcij*[4] *sta merljiva.*
- *Absolutna vrednost, pozitivni in negativni del merljive funkcije so merljivi.*
- *Infimum, supremum, liminf, limsup (in limita, če obstaja) zaporedja merljivih funkcij so merljivi.*

Po trditvi 2.12 je vsaka preslikava $\Omega \to \mathbb{R}$ oblike

$$\sum_{n=1}^{\infty} a_n \mathbb{1}_{A_n},$$

kjer so a_n realna števila in A_n dogodki, slučajna spremenljivka. Pravimo, da je *diskretna.*

[4] Če ni rečeno drugače, bo merljiva funkcija vedno pomenila merljivo preslikavo iz prostora z mero $(\Omega, \mathcal{A}, \mu)$ v prostor realnih števil.

2.4 Lebesguova mera in integral

Zelo uporabno je imeti translacijsko invariantno mero na realnih številih, tj. tako, ki bi vsaki merljivi množici dodelila isto mero kot njenim translacijam (zamikom). Žal pa se celo te na videz enostavne zahteve ne da izpolniti, če zahtevamo merljivost vseh podmnožic realnih števil in upoštevamo aksiom izbire (zadostuje aksiom števne izbire). Protiprimer so npr. Vitalijeve množice [7].

Obravnavajmo najprej aksiom izbire. Če ga upoštevamo, potem veljajo morda neželjeni rezultati, kot je Banach-Tarski ali pa dobra urejenost realnih števil. Če ga zavržemo, pa lahko veljajo trditve, kot ta, da množica lahko premore več ekvivalenčnih razredov kot elementov ali pa, da so realna števila števna unija števnih množic. Slednje je sploh pogubno za teorijo mer s števno aditivnostjo, ki jo pravkar obravnavamo.

Zato moramo aksiom izbire sprejeti in se sprijazniti, da niso vse podmnožice realnih števil merljive glede na translacijsko invariantno mero. Med drugim to pomeni tudi, da za nekatere množice $A \subseteq [0,1]$ ne moremo opredeliti verjetnosti, da $[0,1]$-enakomerno porazdeljena slučajna spremenljivka zavzame vrednost v A. K sreči pa, ker konstrukcija vsake take A potrebuje aksiom izbire (ki zagotavlja obstoj brez konstrukcije), se v praksi na A nikoli ne bo naletelo. Le malo moteče je pri formalnem razvijanju teorije mere.

Izrek 2.13. *Za vsako naravno število n obstaja enolična translacijsko invariantna mera λ_n na $\mathcal{B}_{\mathbb{R}^n}$, ki izmeri volumen večrazsežnega kvadra:*

$$\lambda_n\big([a_1, b_1] \times \cdots \times [a_n, b_n]\big) = \prod_{i=1}^{n} (b_i - a_i).$$

(To velja tudi, če so poljubni intervali odprti ali polodprti.[5]) Imenujemo jo Lebesguova mera na *$\mathbb{R}^n$.*

Sicer se običajno Lebesguovo mero definira na Lebesguovi σ-algebri. To je večja zbirka podmnožic realnih števil, kot je Borelova σ-algebra, a za naše potrebe zadošča Borelova. Lebesguovo mero se, kot bomo še videli, nenehno uporablja.

Verjetno najpogostejši podatek, ki nas zanima o dani slučajni spremenljivki, je njena povprečna vrednost. Vemo že, da je povprečje nenegativne diskretne slučajne spremenljivke $X = \sum_{n=1}^{\infty} a_n \mathbb{1}_{A_n}$ smiselno definirati kot

$$\mathsf{E}X = \mathsf{E}_{\mathsf{P}}X = \sum_{n=1}^{\infty} a_n \mathsf{P}(A_n). \tag{2.1}$$

Neintuitivno dejstvo je, da je povprečna vrednost slučajne spremenljivke lahko neskončna, čeprav spremenljivka zavzame le končne vrednosti. Namreč, če nek poizkus ponavljamo, beležimo rezultat in sproti računamo povprečje vseh dosedanjih rezultatov, bomo vedno dobili končno število. Ampak, če je teoretična povprečna vrednost eksperimenta neskončna, to pomeni, da bodo sicer končna sprotna povprečja naraščala čez vse meje. Ozadje vprašanja je isto kot pri npr. funkciji $\frac{1}{\cdot}$ na $(0,1)$ –

[5] Torej velja npr. $\lambda(\{a\}) = 0$ za $a \in \mathbb{R}$.

funkcija zavzame le končne vrednosti na končnem intervalu (končni interval lahko primerjamo s tem, da je verjetnost polnega dogodka 1, tj. končna), a njen integral je neskončen.

Poglejmo si (2.1) bolj podrobno. X si lahko predstavljamo kot odsekoma konstantno funkcijo, $\mathsf{P}(A_n)$ pa kot velikost množice A_n. Potem (2.1) predstavlja ploščino pod funkcijo X, torej neke vrste integral!
Posplošimo zdaj definicijo na vse merljive preslikave.

Definicija 2.14. Merljiva funkcija f je *enostavna*, če je oblike

$$f = \sum_{n=1}^{N} a_n \mathbf{1}_{A_n},$$

kjer so A_n disjunktne merljive množice in $a_n \geq 0$ za vse n.

Definicija 2.15. *Lebesguov integral* merljive funkcije f po μ je število iz $[-\infty, \infty]$, ki ga označimo z $\int f \, d\mu$ in definiramo po vrsti:

1. Če je $f = \sum_{n=1}^{N} a_n \mathbf{1}_{A_n}$ enostavna, definiramo

$$\int f \, d\mu = \sum_{n=1}^{N} a_n \mu(A_n).$$

2. Če je f nenegativna, označimo z $\mathcal{E}_f$ množico enostavnih preslikav e, za katere je $e \leq f$ in definiramo

$$\int f \, d\mu = \sup_{e \in \mathcal{E}_f} \int e \, d\mu.$$

3. Za splošne preslikave f, če je $\int f^+ \, d\mu = \int f^- \, d\mu = \infty$, pravimo, da Lebesguov integral *ne obstaja*. Sicer pa definiramo

$$\int f \, d\mu = \int f^+ \, d\mu - \int f^- \, d\mu.$$

Če je $(\Omega, \mathcal{A}, \mu) = (\Omega, \mathcal{A}, \mathsf{P})$ verjetnostni prostor, Lebesguovemu integralu pravimo tudi *povprečna vrednost*, *matematično upanje* ali *pričakovana vrednost*[6] in ga označimo z $\mathsf{E}_{\mathsf{P}} f$ ali pa $\mathsf{E} f$, če je verjetnostna mera P znana.
Integral po merljivi množici A definiramo z

$$\int_A f \, d\mu = \int f \mathbf{1}_A \, d\mu.$$

Za boljšo berljivost ob delu s konkretnimi funkcijami (ali pa, če želimo integrirati le po eni komponenti večrazsežne funkcije) lahko v integral vstavimo še umetno spremenljivko (npr. x). Definiramo torej

$$\int f(x) \, d\mu(x) = \int f \, d\mu.$$

[6] Sam se temu imenu izogibam. Če vržemo kocko, v resnici nikoli ne pričakujemo, da bo padlo 3.5 pik. Boljše ime je povprečna vrednost.

Opomba 2.16. Vedno, ko bomo govorili o integralih, bomo v ozadju privzeli, da so obravnavane merljive preslikave take, da integrali obstajajo. Za ostale preslikave pa v zvezi z integrali ni kaj povedati.

Za enostavne preslikave je očitno, da je integral monoton in linearen. To pa velja tudi v splošnem.

Trditev 2.17 (Monotonost integrala). *Za merljivi funkciji $f \geq g$ velja*[7]

$$\int f\,d\mu \geq \int g\,d\mu.$$

Trditev 2.18 (Linearnost integrala). *Za merljivi funkciji f, g in realno število a velja*

$$\int af\,d\mu = a\int f\,d\mu \qquad \text{in} \qquad \int (f+g)\,d\mu = \int f\,d\mu + \int g\,d\mu.$$

Omenimo še, da na konstanto $c \in \mathbb{R}$ lahko gledamo kot preslikavo $c\mathbb{1}_\Omega$ in zato definiramo $\int c\,d\mu = \int c\mathbb{1}_\Omega\,d\mu = c\mu(\Omega) = c$.

Množice z mero 0 v teoriji mere pogosto nimajo nobenega vpliva na ostalo dogajanje – pravimo jim *ničelne.* Zato se uporablja okrajševanje s standardnimi izrazi.

Definicija 2.19. Trditev v zvezi s prostorom z mero velja *μ-skoraj povsod (s.p.)*, če je množica izidov, za katere ne velja, ničelna glede na mero μ. Če je prostor verjetnostni, pravimo, da velja *skoraj gotovo (s.g.)*.

Na primer, za zaporedje merljivih funkcij pravimo, da konvergira skoraj povsod, če je množica izidov, za katere ne konvergira, ničelna. Jasno je vsako konvergentno zaporedje tudi konvergentno s.p.

Brez dokaza navedimo še povezavo med Lebesguovim in bolj poznanim Riemannovim integralom, da vidimo, da sta si pri realnih funkcijah zelo podobna.

Izrek 2.20. *Za funkcijo $f : [a, b] \to \mathbb{R}$ velja:*

- *f je Riemannovo integrabilna natanko tedaj, ko je omejena in zvezna λ-s.p.*
- *Če je f Riemannovo integrabilna, je tudi Lebesguovo in integrala*[8] *sta enaka.*

2.5 Limitni izreki

V teoriji mere je izjemnega pomena vprašanje, kdaj lahko zamenjamo vrstni red integrala in limite. Poglejmo si klasične odgovore na to vprašanje. Njihova uporaba se mnogokrat izkaže pri dokazovanju izrekov – tudi takih, ki jih bomo potrebovali v zvezi s kopulami.

[7] Če ni rečeno drugače, so vse merljive funkcije (oz. slučajne spremenljivke) vedno definirane na istem prostoru z mero (oz. verjetnostnem prostoru) $(\Omega, \mathcal{A}, \mu)$ (oz. $(\Omega, \mathcal{A}, \mathsf{P})$).

[8] Riemannov integral in Lebesguov integral glede na Lebesguovo mero.

Izrek 2.21 (Jensenova neenakost)**.** *Za merljivo funkcijo f in konveksno funkcijo $\psi : \mathbb{R} \to \mathbb{R}$ velja:*

$$\int \psi \circ f \, d\mu \;\geq\; \psi\Big(\int f \, d\mu\Big).$$

Če je ψ celo strogo konveksna in f ni s.p. enaka konstantni funkciji, je neenakost stroga.

Pri Jensenovi enakosti velja analogna trditev z obrnjenim neenačajem, če je funkcija ψ konkavna.

Primer 2.22. Funkcija $|\cdot|$ je konveksna, zato je $\int |f| \, d\mu \geq \left|\int f \, d\mu\right|$ za vsako merljivo funkcijo f.

Preden obravnavamo prvi limitni izrek, pa potrebujemo še drugačen pojem konvergence.

Definicija 2.23. Pravimo, da zaporedje merljivih $(f_n)_{n\in\mathbb{N}}$ konvergira k merljivi funkciji f *v meri* (oz. *v verjetnosti*, če je prostor verjetnostni) in pišemo $f_n \xrightarrow[n\to\infty]{\mu} f$ oz. $f_n \xrightarrow{\mu} f$, če je

$$\lim_{n\to\infty} \mu\big(|f_n - f| > \varepsilon\big) = 0$$

za vsak $\varepsilon > 0$.

To je šibkejši pojem od običajne (po točkah) konvergence – vsako (s.g.) konvergentno zaporedje konvergira tudi v verjetnosti, obratna trditev pa ne velja.

Sledijo klasični izreki o zamenjavi limite in integrala.

Izrek 2.24 (Omejena konvergenca)**.** *Naj merljive funkcije $(f_n)_{n\in\mathbb{N}}$ konvergirajo v meri k merljivi funkciji f. Če je supremum funkcij $(f_n)_{n\in\mathbb{N}}$ omejen in je mera μ končna, potem velja zamenjava limite in integrala*

$$\lim_{n\to\infty} \int f_n \, d\mu = \int f \, d\mu.$$

Izrek 2.25 (Fatoujeva lema)**.** *Za nenegativne merljive funkcije $(f_n)_{n\in\mathbb{N}}$ velja*

$$\int \liminf_{n\in\mathbb{N}} f_n \, d\mu \;\leq\; \liminf_{n\in\mathbb{N}} \int f_n \, d\mu.$$

Izrek 2.26 (Monotona in dominirana konvergenca)**.** *Naj bo $(f_n)_{n\in\mathbb{N}}$ zaporedje merljivih funkcij, ki konvergira proti funkciji f. Naj velja še vsaj eden izmed naslednjih pogojev:*

- (Monotona konvergenca). *Zaporedje funkcij $(f_n)_{n\in\mathbb{N}}$ je nepadajoče in nenegativno.*

- (Dominirana konvergenca). *Integral supremuma absolutnih vrednosti funkcij $(f_n)_{n\in\mathbb{N}}$ je končen: $\int \sup_{n\in\mathbb{N}} |f_n| \, d\mu < \infty$.*

Potem velja zamenjava integrala in limite

$$\lim_{n\to\infty} \int f_n \, d\mu = \int f \, d\mu.$$

Kot rečeno, se limitne izreke pogosto uporablja za dokazovanje drugih izrekov. Sledita dva primera.

Primer 2.27. Naj bodo $(f_n)_{n\in\mathbb{N}}$ nenegativne merljive funkcije. Potem neskončna vsota

$$\sum_{n\in\mathbb{N}} f_n = \lim_{N\to\infty} \sum_{n=1}^{N} f_n$$

obstaja ter zaporedje $\left(\sum_{n=1}^{N} f_n\right)_{N\in\mathbb{N}}$ je nenegativno in naraščujoče, zato je po linearnosti integrala ter monotoni konvergenci

$$\sum_{n\in\mathbb{N}} \int f_n \, d\mu = \lim_{N\to\infty} \sum_{n=1}^{N} \int f_n \, d\mu = \lim_{N\to\infty} \int \sum_{n=1}^{N} f_n \, d\mu = \int \sum_{n\in\mathbb{N}} f_n.$$

Primer 2.28. Naj bo f merljiva funkcija. Pokažimo, da za mero

$$\mu\big(f^{-1}[\cdot]\big),$$

definirano na prostoru realnih števil, velja enakost

$$\int g \, d\mu\big(f^{-1}[\cdot]\big) \;=\; \int g \circ f \, d\mu$$

za poljubno merljivo funkcijo $g : \mathbb{R} \to \mathbb{R}$.

Naj bo najprej $g = \mathbf{1}_B$ indikatorska funkcija Borelove množice B. Potem je po definiciji integrala

$$\int \mathbf{1}_B \, d\mu\big(f^{-1}[\cdot]\big) = \mu\big(f^{-1}[B]\big) = \int \mathbf{1}_{f^{-1}[B]} \, d\mu = \int \mathbf{1}_B \circ f \, d\mu,$$

torej izrek velja za indikatorske funkcije. Iz linearnosti integrala sledi, da mora veljati tudi za vse enostavne funkcije. Naj bo zdaj g poljubna nenegativna merljiva funkcija. Potem obstaja nepadajoče zaporedje enostavnih funkcij $(e_n)_{n\in\mathbb{N}}$, ki konvergira k g (res, vzamemo lahko npr. $e_n = 2^{-n}\lfloor 2^n g \rfloor \wedge n$). Ker izrek že velja za enostavne funkcije, je po dvakratni uporabi monotone konvergence

$$\int g \, d\mu\big(f^{-1}[\cdot]\big) = \lim_{n\to\infty} \int e_n \, d\mu\big(f^{-1}[\cdot]\big) = \lim_{n\to\infty} \int e_n \circ f \, d\mu = \int g \circ f \, d\mu.$$

Na splošne (ne nujno nenegativne) merljive funkcije g pa izrek spet lahko posplošimo z linearnostjo integrala, če upoštevamo razcep funkcije $g = g^+ - g^-$ na dve nenegativni merljivi funkciji.

Dokazi, kot je ta iz primera 2.28, lepo prikažejo moč monotone konvergence. Mnogo izrekov o integralih se namreč lahko dokaže na isti način: pokažemo, da izrek velja za indikatorske funkcije (kar je običajno neprimerljivo lažje, kot za splošne funkcije) ter za linearne kombinacije funkcij, nato pa ga z monotono konvergenco z lahkoto posplošimo na poljubne merljive funkcije.

Opomba 2.29. Dovolj je, da predpostavke izrekov iz tega razdelka veljajo le skoraj gotovo. Npr. za monotnono konvergenco je dovolj, da je $0 \leq f_1 \leq f_2 \leq \cdots$ s.g. in da zaporedje funkcij $(f_n)_{n\in\mathbb{N}}$ konvergira s.g. Podobno bi tudi katera (ali pa vse) izmed f_n lahko bila definirana le skoraj gotovo. Več predpostavk hkrati lahko velja le-skoraj-gotovo, saj je števna unija ničelnih množic še vedno ničelna.

Razlog, da lahko omilimo predpostavke je, da je integral neobčutljiv za spremembe na ničelnih množicah. Za ilustracijo poglejmo prostor z mero $\big([0,1], \mathcal{B}_{[0,1]}, \lambda\big)$ in merljivi funkciji $f = 1$, $g = \mathbf{1}_{\mathbb{R}\setminus\mathbb{Q}} + 10 \cdot \mathbf{1}_{\mathbb{Q}}$. Potem je

$$\int g\,d\lambda = \lambda\big(\mathbb{R}\setminus\mathbb{Q}\big) + 10\lambda(\mathbb{Q}) = 1 + 10\cdot 0 = 1 = \int f\,d\lambda,$$

saj je $\mathbb{Q} = \bigcup_{q\in\mathbb{Q}}\{q\}$ števna unija ničelnih množic in zato ničelna.

Tudi računanje mer množic se ne pokvari, saj je

$$\mu(A) = \mu(A\cap N^{\mathsf{c}}) + \mu(A\cap N) = \mu(A\cap N^{\mathsf{c}}) + 0 = \mu(A\cap N^{\mathsf{c}})$$

za vsako merljivo množico A in ničelno množico N, ker je po monotonosti mere $\mu(A\cap N) \leq \mu(N) = 0$. Zato definiramo

$$\mu(f\in A) = \mu(f\in A\cap N^{\mathsf{c}}),$$

če je f na ničelni množici N nedefinirana merljiva funkcija.

2.6 Absolutna zveznost in razcep mere

Spomnimo se, da je porazdelitev zvezna, če vsaki množici dodeli verjetnost enako integralu določene funkcije po tej množici. Zdaj si bomo ta pojem bolj natančno pogledali in ga razširili v kontekstu teorije mere. Najprej preverimo, da z integriranjem določene funkcije po množici tudi v splošnem dobimo mero.

Trditev 2.30. *Naj bo f nenegativna merljiva funkcija. Potem je preslikava $\nu : \mathcal{A} \to [0,\infty]$, podana z*

$$\nu(A) = \int_A f\,d\mu, \tag{2.2}$$

tudi mera in za vsako merljivo funkcijo h velja

$$\int h\,d\nu = \int hf\,d\mu,$$

zato izrazu (2.2) *pravimo tudi $d\nu = f\,d\mu$.*

Definicija 2.31. Pravimo, da je mera ν *absolutno zvezna* glede na mero[9] μ in pišemo $\nu \ll \mu$, če je vsaka μ-ničelna množica tudi ν-ničelna.

Primer 2.32. Naj bo $d\nu = f\,d\mu$. Za vsako μ-ničelno množico A je

$$\nu(A) = \int_A f\,d\mu \leq \int \infty \mathbf{1}_A\,d\mu = \infty\mu(A) = \infty\cdot 0 = 0,$$

torej je ν absolutno zvezna glede na μ. Iz izreka 2.35 bo sledila tudi obratna trditev, tj. če je ν absolutno zvezna glede na μ, potem obstaja taka nenegativna merljiva funkcija f, da je $d\nu = f\,d\mu$.

[9] Če ni rečeno drugače, so vse mere definirane na istem merljivem prostoru $(\Omega, \mathcal{A})$.

Definicija 2.33. Mera μ je *skoncentrirana* na merljivi množici A, če njenemu komplementu dodeli vrednost 0.
Meri μ in ν *sta si singularni* (oznaka $\mu \perp \nu$), če obstajata taki disjunktni merljivi množici M in N, da je μ skoncentrirana na M in ν na N.
Mera na n-razsežni Borelovi σ-algebri je *singularna*, če si je singularna z n-razsežno Lebesguovo mero λ_n.

Primer 2.34. Mera δ_r na Borelovi σ-algebri je singularna za vsako realno število r, saj je skoncentrirana na množici $\{r\}$, Lebesguova mera pa na množici $\mathbb{R} \setminus \{r\}$. Podobno je vsaka diskretna mera (tj. vsota delta mer) singularna.

Izrek 2.35 (Lebesgue-Radon-Nikodym)**.** *Naj bosta μ in ν σ-končni meri. Potem obstajata enolični meri μ_a in μ_s, ki tvorita* Lebesguov razcep

$$\mu = \mu_a + \mu_s, \qquad \mu_a \ll \nu, \qquad \mu_s \perp \nu,$$

ter taka ν-s.p.-enolična[10] *nenegativna merljiva funkcija g, ki ji pravimo* gostota *oz.* Radon-Nikodymov odvod*, da je*

$$d\mu_a = g\, d\nu.$$

Gostoto g označimo z $\frac{d\mu_a}{d\nu}$.[11]

Opomba 2.36. Uporabne mere so običajno σ-končne. Npr. n-razsežna Lebesguova mera je σ-končna, saj prostor $\mathbb{R}^n$ lahko zapišemo kot števno unijo večrazsežnih kock $\bigcup_{a \in \mathbb{N}} [-a, a]^n$, vseh s končno n-razsežno Lebesguovo mero.

Opomba 2.37. Če je μ σ-končna mera, obstajajo merljive množice $(A_n)_{n\in\mathbb{N}}$ s končno mero, katerih unija je cel prostor Ω. Brez škode lahko privzamemo, da so množice $(A_n)_{n\in\mathbb{N}}$ disjunktne, saj bi sicer lahko obravnavali zaporedje množic

$$\left(A_n \setminus \bigcup_{i=1}^{n-1} A_i \right)_{n\in\mathbb{N}}.$$

Z zožitvijo mere μ na vsako izmed množic A_n dobimo mere $(\mu_n)_{n\in\mathbb{N}}$, podane z

$$\mu_n = \mu(\cdot \cap A_n).$$

Te mere so očitno končne in njihova vsota je μ. Mnogokrat to zadostuje, da trditev, ki velja o končnih merah $(\mu_n)_{n\in\mathbb{N}}$, posplošimo na σ-končno mero μ.

Zdaj končno lahko enotno opišemo verjetnostne porazdelitve:

Primer 2.38. Iz osnov verjetnosti poznamo zvezno porazdelitev na $\mathbb{R}$ kot pravilo, ki za vsak $a \in \mathbb{R}$ intervalu $(-\infty, a]$ dodeli verjetnost

$$\mathsf{P}\big((-\infty, a]\big) = \int_{-\infty}^{a} f(x)\, dx,$$

kjer je f neka nenegativna funkcija, za katero ta (posplošeni Riemannov) integral obstaja. Zdaj lahko preidemo na nov sistem:

[10] To pomeni, da je vsaka druga takšna funkcija $\tilde{g}$ enaka g skoraj povsod.
[11] Oznaka je takšna, da izgleda, kot da se stvari pokrajšajo v izrazu $d\mu_a = \frac{d\mu_a}{d\nu}\, d\nu$.

- Porazdelitev obravnavamo kot verjetnostno mero.
- Namesto samo končnih in neskončnih intervalov zdaj lahko dodelimo verjetnost vsaki Borelovi množici.
- Integral bo zdaj Lebesguov glede na Lebesguovo mero in f naj bo merljiva funkcija.

Tako dobljena mera je absolutno zvezna glede na Lebesguovo, saj je $d\mathsf{P} = f\,d\lambda$.

Primer 2.39. Iz osnov verjetnosti poznamo tudi diskretne porazdelitve. Zanje smo že na koncu razdelka 2.3 ugotovili, da so oblike $\mathsf{P} = \sum_{n\in\mathbb{N}} p_n\delta_{\omega_n}$. Zanimivo pa je, da za dogodek A lahko zapišemo

$$\mathsf{P}(A) = \sum_{n\in\mathbb{N}} p_n\delta_{\omega_n}(A) = \int_A g\,d\varsigma,$$

$$\text{kjer je}\quad g = \sum_{n\in\mathbb{N}} p_n \mathbb{1}_{\{\omega_n\}}$$

in ς mera, ki množici priredi št. njenih elementov.

Da to drži ne glede na zapis g (če je npr. $p_1 = p_2$, je tudi $g = p_1\mathbb{1}_{\{\omega_1,\omega_2\}} + \sum_{n\geq 3} p_n\mathbb{1}_{\{\omega_n\}}$), sledi iz definicije ς, ki ji pravimo *mera štetja.* Pravzaprav lahko definiramo, da je porazdelitev *diskretna*, če je absolutno zvezna glede na mero štetja.
Mero štetja si lahko predstavljamo kot vsoto vseh delta mer:

$$\varsigma = \sum_{\omega\in\Omega} \delta_\omega.$$

To lahko trdimo tudi, če je Ω neštevna, saj je vsota sestavljena le iz enk in ničel.

Primer 2.40. Zdaj lahko dobro opišemo tudi porazdelitev P iz primera 2.1:

$$\mathsf{P} = \mathsf{P}_a + P_d, \qquad d\mathsf{P}_a = \big(\mathbb{1}_{[0,5)}f\big)\,d\lambda, \qquad P_d = \frac{1}{5}\delta_5,$$

kjer je f gostota Pareto$(1,1)$ porazdelitve. Ni pa to edini možen opis. Prav tako drži tudi

$$d\mathsf{P} = \left(\mathbb{1}_{[0,5)}f + \frac{1}{5}\mathbb{1}_{\{5\}}\right) d(\lambda + \delta_5).$$

Je pa res, da običajno želimo mere na Borelovih σ-algebrah preučevati glede na Lebesguovo mero, ne pa glede na $\lambda + \delta_5$.

Poglejmo si zdaj natančno razvrstitev vseh σ-končnih mer na n-razsežni Borelovi σ-algebri.
Naj bo μ poljubna σ-končna mera na $\mathcal{B}_{\mathbb{R}^n}$. Če jo razcepimo glede na Lebesguovo mero λ_n, dobimo taki meri μ_a in μ_s, da je

$$\mu = \mu_a + \mu_s, \qquad \mu_a \ll \lambda_n, \qquad \mu_s \perp \lambda_n.$$

Nadalje naj bo N nosilec diskretnega dela mere μ_s, tj.

$$N = \big\{x \in \mathbb{R}^n \mid \mu_s(\{x\}) > 0\big\}.$$

Lema 2.41. *Množica N je števna.*

Definirajmo meri μ_d in μ_{sz} z

$$\mu_d = \mu_s(\cdot \cap N)$$
$$\mu_{sz} = \mu_s(\cdot \cap N^{\mathsf{c}})$$

Očitno je μ_d diskretna, μ_{sz} pa singularna in vsaki enoelementni množici priredi vrednost 0, tako kot zvezne mere. Zato ji pravimo *singularno zvezna*.
Pokazali smo:

Trditev 2.42. *Vsaka mera na n-razsežni Borelovi σ-algebri je enolična vsota absolutno zvezne*[12]*, diskretne in singularno zvezne mere.*

Primer 2.43. Porazdelitev iz primera 2.2 je singularno zvezna (njen absolutno zvezni in diskretni del sta oba enaka 0). Prav tako so singularno zvezne porazdelitve, katerih porazdelitvena funkcija je teoretična pošast (npr. Cantorjeva funkcija). V več razsežnostih pa je singularno zvezni del lahko neničeln tudi pri zelo praktičnih primerih porazdelitev.

Dokaz leme 2.41. Z vidika opombe 2.37 zadostuje dokaz v primeru, da je μ_s končna mera, saj se trditev posploši na σ-končne, ker je števna unija števnih množic tudi števna.
Naj bo torej μ_s končna. Ker je

$$N = \bigcup_{n \in \mathbb{N}} \{x \in \mathbb{R}^n \mid \mu_s(x) > 1/n\}$$

števna unija končnih množic (za vsak n množica $\left\{x \in \mathbb{R}^n \mid \mu_s(x) > \frac{1}{n}\right\}$ ne more biti neskončna, saj je μ_s končna mera), je števna. □

3 Kopule

Med najpomembnejšimi temami verjetnosti in statistike so slučajni vektorji. Velika moč in uporaba je v razumevanju medsebojne odvisnosti komponent vektorjev, ko pa začnemo preučevati odvisnost, pa hitro naletimo na kopule. Sklarov izrek 3.20 pravzaprav pove, da so kopule edina stvar, ki jo sploh potrebujemo za študij odvisnosti.

V tem poglavju si bomo pogledali nekatere osnovne rezultate, s posebnim poudarkom na Sklarovem izreku in njegovem dokazu. Za podroben in obsežen pregled kopul se obrnemo na knjigo [3].

3.1 Osnovni pojmi

Zdaj, ko poznamo vse potrebno predznanje iz teorije mere, lahko ponovno definiramo nekatere osnovne verjetnostne pojme.

Merljivim preslikavam iz verjetnostnega prostora v $\mathbb{R}^n$ pravimo *slučajni vektorji*. Vsaka komponenta slučajnega vektorja je tako slučajna spremenljivka.

[12] Če ni rečeno drugače, bo to vedno pomenilo glede na Lebesguovo mero primerne razsežnosti.

Definicija 3.1. *Porazdelitev slučajnega vektorja* $\mathbf{X}$ je mera $\mathsf{P}_{\mathbf{X}}$ na $\mathbb{R}^n$, podana z

$$\mathsf{P}_{\mathbf{X}} = \mathsf{P}(\mathbf{X} \in \cdot) = \mathsf{P}\big(\{\omega \in \Omega \mid \mathbf{X}(\omega) \in \cdot\}\big).$$

Da je $\mathsf{P}_{\mathbf{X}}$ res mera, je očitno: velja $\mathsf{P}_{\mathbf{X}}(\emptyset) = \mathsf{P}\big(\mathbf{X} \in \emptyset\big) = 0$ in

$$\begin{aligned}\mathsf{P}_{\mathbf{X}}\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \mathsf{P}\left(\mathbf{X} \in \bigcup_{n\in\mathbb{N}} A_n\right) = \mathsf{P}\left(\bigcup_{n\in\mathbb{N}}\{\mathbf{X} \in A_n\}\right) &= \sum_{n\in\mathbb{N}} \mathsf{P}(\mathbf{X} \in A_n) \\ &= \sum_{n\in\mathbb{N}} \mathsf{P}_{\mathbf{X}}(A_n)\end{aligned}$$

za disjunktne dogodke $(A_n)_{n\in\mathbb{N}}$, saj je P mera.
Vsakemu slučajnemu vektorju torej lahko pripišemo mero na $\mathbb{R}^n$, velja pa tudi obratna trditev: če je P mera na $\mathbb{R}^n$, potem obstaja nek slučajni vektor na nekem verjetnostnem prostoru, katerega porazdelitev je P. Še več:

Izrek 3.2. *Naj bo* P_i *verjetnostna mera na* $\mathbb{R}^{n_i}$ *za* $i = 1, \ldots, N$. *Potem obstaja verjetnostni prostor* $(\Omega, \mathcal{A}, \mathsf{P})$ *in neodvisni slučajni vektorji* $\mathbf{X}_i : \Omega \to \mathbb{R}^{n_i}$ *s porazdelitvami* P_i *za vsak* i.

Dokaz. Vzamemo lahko npr.

$$\Omega = \times_{i=1}^{N} \mathbb{R}^{n_i}, \qquad \mathcal{A} = \times_{i=1}^{N} \mathcal{B}_{\mathbb{R}^{n_i}}, \qquad \mathsf{P} = \times_{i=1}^{N} \mathsf{P}_i, \qquad \mathbf{X}_i(\omega_1, \ldots, \omega_N) = \omega_i$$

Potem je

$$\begin{aligned}\mathsf{P}(\mathbf{X}_1 \in A_1, \ldots, \mathbf{X}_N \in A_N) &= \mathsf{P}\big((\omega_1, \ldots, \omega_N) \in \Omega \mid \omega_1 \in A_1, \ldots, \omega_N \in A_N\big) \\ &= \mathsf{P}(A_1 \times \cdots \times A_N) = \prod_{i=1}^{N} \mathsf{P}_i(A_i) = \prod_{i=1}^{N} \mathsf{P}(\mathbf{X}_i \in A_i)\end{aligned}$$

za vse Borelove množice $A_1 \in \mathcal{B}_{\mathbb{R}^{n_1}}, \ldots, A_N \in \mathcal{B}_{\mathbb{R}^{n_N}}$, kar pomeni, da so vektorji $\mathbf{X}_1, \ldots, \mathbf{X}_N$ res neodvisni in so $\mathsf{P}_1, \ldots, \mathsf{P}_N$ njihove porazdelitve. □

Opomba 3.3. Izrek se da posplošiti na števno neskončne zbirke slučajnih vektorjev, glej npr. [5, str. 70].

Opomba 3.4. Zato se pogosto reče "naj bo zdaj s.s. X z neko lastnostjo" tudi v primerih, ko imamo verjetnostni prostor že določen in na njem morda željene X ni možno definirati, saj vsak verjetnostni prostor lahko razširimo do takega, ki vsebuje vse že definirane s.s. (oz. vektorje) in na katerem je možno definirati še X z željenimi lastnostmi. Čeprav se te zamenjave verjetnostnega prostora običajno ne omenja, se v ozadju vseeno zgodi.

Definicija 3.5. *Porazdelitvena funkcija* mere P na $\mathbb{R}^n$ (oz. samo *porazdelitvena funkcija*) je funkcija $F_{\mathsf{P}} : \mathbb{R}^n \to [0, 1]$, podana z

$$F_{\mathsf{P}}(x_1, \ldots, x_n) = \mathsf{P}\big((-\infty, x_1] \times \cdots \times (-\infty, x_n]\big).$$

Porazdelitvena funkcija slučajnega vektorja $\mathbf{X}$ je $F_{\mathsf{P}_{\mathbf{X}}}$ (tj. porazdelitvena funkcija porazdelitve vektorja $\mathbf{X}$), ki jo bomo označevali krajše z $F_{\mathbf{X}}$.
Oznaka

$$\mathbf{X} \sim F$$

pomeni, da je F porazdelitvena funkcija slučajnega vektorja $\mathbf{X}$.

Opomba 3.6. Za slučajni vektor $\mathbf{X} = (X_1, \dots, X_n)$ torej velja

$$F_{\mathbf{X}}(x_1, \dots, x_n) = \mathsf{P}_{\mathbf{X}}\big((-\infty, x_1] \times \cdots \times (-\infty, x_n]\big) = \mathsf{P}(X_1 \leq x_1, \dots, X_n \leq x_n)$$

za $(x_1, \dots, x_n) \in \mathbb{R}^n$.

Če je $F : \mathbb{R} \to [0,1]$ porazdelitvena funkcija, iz monotonosti in zveznosti mere sledijo naslednje štiri trditve:

- F je nepadajoča.
- F je desno zvezna.
- $\lim_{x\to\infty} F(x) = 1$
- $\lim_{x\to-\infty} F(x) = 0$.

Pokazati pa se da tudi obrat, tj. če je $F : \mathbb{R} \to [0,1]$ funkcija, ki zadošča tem štirim pogojem, potem obstaja taka mera na prostoru realnih števil, da je F njena porazdelitvena funkcija. V naslednjem izreku to posplošimo na slučajne vektorje.

Izrek 3.7. *Za funkcijo $F : \mathbb{R}^n \to [0,1]$ je ekvivalentno:*

- *F je n-razsežna porazdelitvena funkcija.*
- *F zadošča naslednjim pogojem:*
  1. *Funkcija $F(x_1, \dots, x_j, \cdot, x_{j+1}, \dots, x_n)$ je desno zvezna za vsak $j = 1, \dots, n$ in $x_1, \dots, x_j, x_{j+1}, \dots, x_n \in \mathbb{R}$.*
  2. *Vrednost $F(x_1, \dots, x_n)$ gre proti 1, če gredo vsi argumenti proti ∞.*
  3. *Vrednost $F(x_1, \dots, x_n)$ gre proti 0, če gre vsaj en argument proti $-\infty$.*
  4. *Funkcija F je n-*razsežno nepadajoča*, tj. vsak n-razsežni kvader*

     $$K = (a_1, b_1] \times \cdots \times (a_n, b_n]$$

     *ima nenegativen F-*volumen

     $$\sum_{\mathbf{v}} (-1)^{\left|\{j=0,\dots,n \,|\, v_j = a_j\}\right|} F(\mathbf{v}), \tag{3.1}$$

     kjer vsota teče po vseh ogljiščih $\mathbf{v} = (v_1, \dots, v_n)$ kvadra K.

Dokaz je malo težji. Najdemo ga v [1, str. 187].

n-razsežno nepadajočo funkcijo si je lažje predstavljati na primeru $n = 2$, takrat je namreč

$$(3.1) = F(b_1, b_2) - F(a_1, b_2) - F(a_2, b_1) + F(a_1, a_2),$$

kar pa je ravno enako $\mathsf{P}\big((X,Y) \in K\big)$, če je F porazdelitvena funkcija slučajnega vektorja (X,Y). Pogoj torej pravi, da mora porazdelitev vsakemu kvadru dodeliti nenegativno vrednost.

Definicija 3.8. Naj bo P porazdelitev na R^n. *Robne porazdelitve* so porazdelitve P_i na prostoru realnih števil, podane z

$$\mathsf{P}_i = \mathsf{P}(\mathbb{R} \times \cdots \times \mathbb{R} \times \cdot \times \mathbb{R} \times \cdots \times \mathbb{R}),$$

kjer je $\cdot$ na i-tem mestu.
Robna porazdelitev slučajnega vektorja $\mathbf{X}$ je robna porazdelitev porazdelitve $\mathsf{P}_{\mathbf{X}}$.
Robna porazdelitvena funkcija je porazdelitvena funkcija robne porazdelitve.

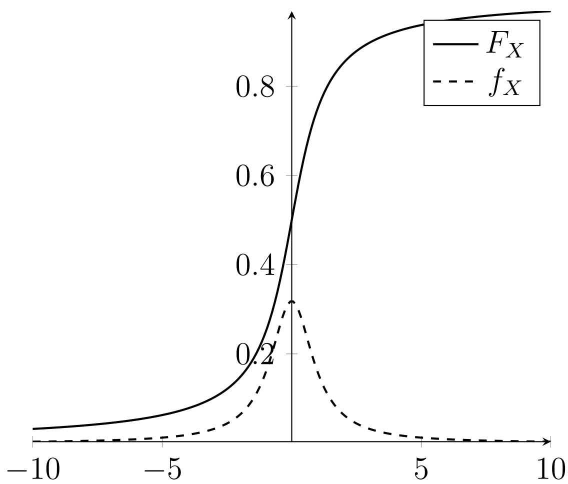
Robna porazdelitev torej opiše porazdelitev ene komponente mere oz. vektorja, če na ostalih komponentah ni omejitev.
Očitno je i-ta robna porazdelitvena funkcija slučajnega vektorja $(X_1, \dots, X_n)$ ravno F_{X_i}.

3.2 Sklarov izrek

Primer 3.9. Naj bo X standardno Cauchyjevo porazdeljena slučajna spremenljivka, tj. naj bo njena porazdelitvena funkcija F_X podana z

$$F_X(x) = \frac{1}{2} + \frac{\arctan(x)}{\pi}.$$

Jasno je, da je porazdelitev X drugačna od porazdelitve s.s. $h(X)$ za mnogo funkcij $h : \mathbb{R} \to \mathbb{R}$. Zanimivo bi bilo pogledati, katero porazdelitev bi dobili, če bi vzeli $h = F_X$. Poizkusimo najprej uganiti rezultat.
Situacijo si lahko predstavljamo kot, da najprej izberemo naključno točko na abscisi



Slika 2: Standardna Cauchyjeva porazdelitev.

(v pomoč je slika 2) tako, da bodo točke, kjer je f_X večja, bolj verjetne. Nato izbrano točko preslikamo na ordinato z F_X. Postopek velikokrat ponovimo in pogledamo gostoto točk na ordinati – to bo gostota porazdelitve $F_X(X)$.
Opazimo, da večja kot je gostota f_X v določeni točki x na abscisi, večja je verjetnost, da bomo naključno izbrali točko blizu x, a hkrati je ravno zato, ker je tam f_X večja, tudi graf F_X tam bolj strm in posledično se majhen interval okoli točke x na abscisi

razpotegne v večji interval na ordinati. Pri intervalih na abscisi, ki so manj verjetni, pa je ravno obraten učinek: naključno bo izbranih manj točk iz njih, ampak, ker je F_X tam položna, se bodo vse te točke zgostile v zelo majhen interval na ordinati.
V obeh primerih imamo torej dva dejavnika, od katerih eden povečuje gostoto na ordinati (večja f_X oz. bolj položen graf F_X) in drugi zmanjšuje (bolj strm graf F_X oz. manjša f_X). Ali ne bi bilo zanimivo, če imata oba dejavnika enak vpliv, da se skupno izničita? V tem primeru bi bile točke na ordinati porazdeljene enakomerno! To se dejansko zgodi. Dokažimo: za $x \in [0,1]$ je

$$\begin{aligned} F_{F_X(X)}(x) &= \mathsf{P}\big(F_X(X) \leq x\big) = \mathsf{P}\big(F_X(X) \in (-\infty, x]\big) = \mathsf{P}\Big(X \in F_X^{-1}\big[(-\infty, x]\big]\Big) \\ &= \mathsf{P}\big(X \leq F_X^{-1}(x)\big) \qquad\qquad \text{(Ker je } F_X \text{ naraščujoča in zvezna.)} \\ &= F_X(F_X^{-1}(x)) = x \end{aligned}$$

Za $x < 0$ pa je $\mathsf{P}\big(F_X(X) \leq x\big)$ enaka 0 oz. 1 za $x > 1$, saj F_X zavzame vrednosti le v $[0,1]$. Torej je

$$F_{F_X(X)}(x) = x\mathbb{1}_{[0,1]}(x) + \mathbb{1}_{(1,\infty)}(x),$$

kar pa je ravno porazdelitvena funkcija enakomerne porazdelitve na $[0,1]$.

Pravzaprav sploh ni bilo pomembno, da je X porazdeljena standardno Cauchy-jevo. Isti razultat bi dobili za vsako s.s. X, katere porazdelitvena funkcija je zvezna in naraščujoča (torej taka, ki ima povsod pozitivno gostoto).

S primerom 3.9 smo pokazali, da v veliko primerih iz slučajne spremenljivke lahko dobimo $[0,1]$-enakomerno spremenljivko, če jo preslikamo s svojo porazdelitveno funkcijo.
Ideja: za slučajni vektor to lahko storimo na vsaki komponenti posebej in tako dobimo slučajni vektor, katerega robne porazdelitve so porazdeljene enakomerno na $[0,1]$.
Kaj pa nam to pomaga? Ker smo naredili preslikavo na vsaki komponenti slučajnega vektorja posebej, smo spremenili samo obliko porazdelitve, medsebojne odvisnosti komponent pa ne! To pomeni, da smo dobili nekakšno standardno obliko za preučevanje odvisnosti, namreč študij odvisnosti kateregakoli slučajnega vektorja lahko prevedemo na študij odvisnosti njegove pripadajoče (v smislu opisane transformacije) porazdelitve na $[0,1]^n$.
Ta pristop pa je izjemno uporaben. Ker je z njim možno obliko slučajnega vektorja popolnoma ločiti od medsebojne odvisnosti njegovih komponent, je možno pojave modelirati večkoračno, npr. najprej lahko razmislimo, katere oblike odvisnosti so v določenem primeru smiselne, med njimi poiščemo najboljšo, potem pa samo še določimo obliko vektorja, ki se prilega podatkom, in odvisnost komponent se pri tem ne pokvari.

Da je zgornji razmislek veljaven za vse slučajne vektorje (ne le za zvezne), pa nam zagotavlja Sklarov izrek.

Definicija 3.10. Funkcija $C : \mathbb{R}^n \to [0,1]$ je *kopula*, če je porazdelitvena funkcija neke porazdelitve (verjetnostne mere), katere robne porazdelitve so enakomerne na $[0,1]$.

Opomba 3.11. Kopule lahko brez škode obravnavamo tudi kot funkcije $[0,1]^n \to [0,1]$, saj so njihove vrednosti izven večrazsežne enotske kocke jasne in iste za vse kopule.

Opomba 3.12. i-ta robna porazdelitvena funkcija kopule C je

$$C(1,\ldots,1,\cdot,1,\ldots,1)$$

kjer je $\cdot$ na i-tem mestu. Nanjo lahko gledamo kot funkcijo $[0,1] \to [0,1]$.

Opomba 3.13. Z vidika izreka 3.7 je funkcija $C : [0,1]^n \to [0,1]$ kopula natanko tedaj, ko je desno zvezna v vsaki komponenti, n-razsežno nepadajoča in velja

$$C(1,\ldots,1) = 1 \qquad \text{ter} \qquad C(x_1,\ldots,x_n) = 0,$$

če je vsaj eden izmed $x_1,\ldots,x_n$ enak 0.

Poglejmo si najprej nekaj najosnovnejših kopul.

Primer 3.14 (Neodvisne komponente)**.** Naj bodo $(U_i)_{i=1}^n$ neodvisne slučajne spremenljivke, porazdeljene enakomerno na $[0,1]$. Potem ima vektor $(U_1,\ldots,U_n)$ robne porazdelitve $[0,1]$-enakomerne, torej je $F_{U_1,\ldots,U_n}$ kopula. Poglejmo, kakšna je. Za $(x_1,\ldots,x_n) \in [0,1]^n$ je

$$\begin{aligned} F_{U_1,\ldots,U_n}(x_1,\ldots,x_n) &= \mathsf{P}(U_1 \le x_1,\ldots,U_n \le x_n) = \mathsf{P}(U_1 \le x_1)\cdots\mathsf{P}(U_n \le x_n) \\ &= \prod_{i=1}^n x_i. \end{aligned}$$

Kopuli $F_{U_1,\ldots,U_n}$ pravimo *n-razsežna kopula neodvisnosti* in jo označimo s Π_n.

Primer 3.15 (Popolnoma pozitivno korelirane komponente)**.** Naj bo s.s. U porazdeljena enakomerno na $[0,1]$. Potem so komponente n-razsežnega vektorja $(U,\ldots,U)$ popolnoma iste, jasno pa je tudi, da je njegova porazdelitvena funkcija kopula. Izračunajmo za $(x_1,\ldots,x_n) \in [0,1]^n$:

$$F_{U,\ldots,U}(x_1,\ldots,x_n) = \mathsf{P}(U \le x_1,\ldots,U \le x_n) = \mathsf{P}\left(U \le \bigwedge_{i=1}^n x_i\right) = \bigwedge_{i=1}^n x_i.$$

Kopuli $F_{U,\ldots,U}$ pravimo *n-razsežna komonotonostna kopula* in jo označimo z M_n.

Primer 3.16 (Popolnoma negativno korelirani komponenti)**.** Naj bo s.s. U porazdeljena enakomerno na $[0,1]$. Komponente slučajnega vektorja $(U, 1-U)$ so popolnoma negativno korelirane in enakomerno porazdeljene na $[0,1]$. Tako je za $x, y \in [0,1]$ funkcija

$$F_{U,1-U}(x,y) = \mathsf{P}(U \le x, 1-U \le y) = \mathsf{P}\big(U \in [1-y, x]\big) = (x+y-1) \vee 0$$

kopula, ki ji pravimo *kontramonotonostna kopula* in jo označimo z W_2. Predpis lahko posplošimo na več razsežnosti z

$$W_n(x_1,\ldots x_n) = \left(\sum_{i=1}^n x_i - (n-1)\right) \vee 0,$$

ampak za razliko od Π_n in M_n funkcija W_n ni kopula za $n \ge 3$.

Naslednji izrek pokaže, da je odvisnost poljubnega slučajnega vektorja vedno nekje med popolno pozitivno in popolno negativno koreliranostjo.

Izrek 3.17 (Frechet-Hoeffdingove meje)**.** *Za vsako n-razsežno kopulo C velja*

$$W_n \leq C \leq M_n.$$

Dokaz. Ker je C kopula, obstaja tak slučajni vektor $(U_1, \ldots, U_n)$ z $[0,1]$-enakomerno porazdeljenimi komponntami, da je

$$C(x_1, \ldots, x_n) = \mathsf{P}(U_1 \leq x_1, \ldots, U_n \leq x_n)$$

za vse $(x_1, \ldots, x_n) \in [0,1]^n$. Za vsak $k \in \{1, \ldots, n\}$ je

$$\{U_1 \leq x_1, \ldots, U_n \leq x_n\} = \bigcap_{i=1}^{n} \{U_i \leq x_i\} \subseteq \{U_k \leq x_k\},$$

zato iz monotonosti mere sledi

$$C(x_1, \ldots, x_n) \leq \mathsf{P}(U_k \leq x_k) = x_k$$

za vsak k. Torej mora biti

$$C(x_1, \ldots, x_n) \leq \bigwedge_{k=1}^{n} x_k = W_n(x_1, \ldots, x_n).$$

V drugo smer je po zakonih množic

$$\left(\bigcap_{i=1}^{n} \{U_i \leq x_i\} \right)^{\mathsf{c}} = \bigcup_{i=1}^{n} \{U_i > x_i\},$$

potem pa po subaditivnosti mere

$$\begin{aligned} C(x_1, \ldots, x_n) = 1 - \mathsf{P}\left(\bigcup_{i=1}^{n} \{U_i > x_i\} \right) \geq 1 - \sum_{i=1}^{n} \mathsf{P}(U_i > x_i) &= 1 - \sum_{i=1}^{n} (1 - x_i) \\ &= \sum_{i=1}^{n} x_i - (n-1). \end{aligned}$$

Nazadnje sledi še $C = C \vee 0 \geq W_2$, saj je funkcija C nenegativna. □

Nadaljujmo z razmislekom. Naj bo $\mathbf{X} = (X_1, \ldots, X_n)$ slučajni vektor, katerega robne porazdelitvene funkcije so zvezne in naraščujoče. Potem je kopula $C_{\mathbf{X}}$, ki pripada $\mathbf{X}$, podana z

$$\begin{aligned} C_{\mathbf{X}}(x_1, \ldots, x_n) &= \mathsf{P}\big(F_{X_1}(X_1) \leq x_1, \ldots, F_{X_n}(X_n) \leq x_n\big) \\ &= \mathsf{P}\Big(X_1 \leq F_{X_1}^{-1}(x_1), \ldots, X_n \leq F_{X_n}^{-1}(x_n)\Big) \\ &= F_{\mathbf{X}}\Big(F_{X_1}^{-1}(x_1), \ldots, F_{X_n}^{-1}(x_n)\Big) \end{aligned}$$

Če vstavimo $x_i = F_{X_i}(y_i)$ za vsak i in preimenujemo spremenljivke, dobimo

$$F_{\mathbf{X}}(x_1, \ldots, x_n) = C_{\mathbf{X}}\big(F_{X_1}(x_1), \ldots, F_{X_n}(x_n)\big). \tag{3.2}$$

Enačba (3.2) pove, kako lahko iz dane kopule in robnih porazdelitev dobimo pripadajoči slučajni vektor! To deluje tudi v splošnem (ni treba, da so funkcije F_{X_i}, $i = 1, \ldots, n$ zvezne). Pravzaprav je ta način zelo smiseln – najprej točko $(x_1, \ldots, x_n)$ pretvorimo v standardno obliko, tako, da vsako komponento posebej preslikamo s pripadajočo robno porazdelitveno funkcijo, nato pa pogledamo, kaj o porazdelitvi pravi kopula.
Opisan način je v obe smeri (vektor $\to$ kopula ali kopula $\to$ vektor) konsistenten. V primeru naraščujočih zveznih robnih porazdelitev smo to že videli, splošni primer pa bo razviden iz Sklarovega izreka.

Primer 3.18. Vemo, da sta s.s. X in Y neodvisni natanko tedaj, ko je

$$F_{X,Y}(x, y) = F_X(x)F_Y(y)$$

za vse $x, y \in \mathbb{R}$. Spomnimo se, da je kopula neodvisnosti $\Pi_2(x, y) = xy$. Torej je res

$$F_{X,Y}(x, y) = \Pi_2(F_X(x), F_Y(x)).$$

Primer 3.19. Kakšna je porazdelitvena funkcija dvorazsežnega slučajnega vektorja, katerega komponenti sta standardno Laplaceovo in Cauchyjevo porazdeljeni ter imata korelacijski koeficient -1? S kopulami na preprost način pridemo do odgovora.
Naj bo $X \sim$ Laplace(0,1) in $Y \sim \mathrm{Cauchy}(0, 1)$. Potem je

$$F_X(x) = \frac{1}{2}e^x \mathbf{1}_{(-\infty,0]}(x) + \left(1 - \frac{1}{2}e^{-x}\right)\mathbf{1}_{(0,\infty)}(x) \qquad \text{in} \qquad F_Y(y) = \frac{1}{2} + \frac{\arctan(y)}{\pi}.$$

Vse kar je potrebno narediti je, da ti funkciji povežemo s kontramonotonostno kopulo:

$$\begin{aligned} F_{X,Y}(x, y) &= W_2\big(F_X(x), F_Y(y)\big) \\ &= \left(\frac{\arctan(y)}{\pi} + \frac{1}{2}e^x \mathbf{1}_{(-\infty,0]}(x) + \left(1 - \frac{1}{2}e^{-x}\right)\mathbf{1}_{(0,\infty)}(x) - \frac{1}{2}\right) \vee 0 \end{aligned}$$

Konkretno v tem primeru porazdelitvene funkcije $F_{X,Y}$ ne bi bilo pretežko ugotoviti tudi brez pomoči kopul. Vendar način s kopulami enako preprosto deluje tudi pri mnogo bolj zakompliciranih strukturah odvisnosti kot kontramonotonost W_2.

Dokažimo zdaj vse ideje še v splošnem.

Izrek 3.20 (Sklar). *Za vsako n-razsežno porazdelitveno funkcijo F z robnimi porazdelitvenimi funkcijami $F_1, \ldots, F_n$ obstaja taka n-razsežna kopula C_F, da je*

$$F(x_1, \ldots, x_n) = C_F\big(F_1(x_1), \ldots, F_n(x_n)\big)$$

za vse $(x_1, \ldots, x_n) \in \mathbb{R}^n$.
Obratno, če je C n*-razsežna kopula in* $F_1, \ldots, F_n$ *enorazsežne porazdelitvene funkcije, potem je funkcija* $F : \mathbb{R}^n \to [0, 1]$*, definirana z*

$$F(x_1, \ldots, x_n) = C\big(F_1(x_1), \ldots, F_n(x_n)\big),$$

n*-razsežna porazdelitvena funkcija z robnimi porazdelitvenimi funkcijami* $F_1, \ldots, F_n$.

Obstaja veliko dokazov Sklarovega izreka, po mojem mnenju najlepši pa posploši rezultat primera 3.9 na splošne porazdelitve. Preostanek razdelka je posvečen temu dokazu. Idejo sem vzel iz [3, str. 53] ali pa [6], dokazal pa sam na svoj način.

Definicija 3.21. Kjer obstajata, označimo z f_- *levo zvezno verzijo* in z Δf *funkcijo levih skokov* funkcije $f : \mathbb{R} \to \mathbb{R}$:

$$f_-(x) = \lim_{s \uparrow x} f(s) \qquad \text{in} \qquad \Delta f = f - f_-.$$

Podobno definiramo *desno zvezno verzijo* f_+ in *funkcijo desnih skokov* $f_+ - f$ funkcije f, kjer obstajata.

Definicija 3.22. *Spodnja kvantilna funkcija* enorazsežne porazdelitvene funkcije F je funkcija $F^\wedge : (0, 1) \to \mathbb{R}$, podana z

$$F^\wedge(\alpha) = \min\{x \in \mathbb{R} \mid F(x) \geq \alpha\} = \sup\{x \in \mathbb{R} \mid F(x) < \alpha\}. \tag{3.3}$$

Opomba 3.23. Ker je F nepadajoča, je množica $\{x \in \mathbb{R} \mid F(x) \geq \alpha\}$ interval oblike (c, ∞) ali $[c, \infty)$ za nek $c \in \mathbb{R}$. Ker pa je F še desno zvezna in neenakost ni stroga, je pravilna možnost $[c, \infty)$. Torej minimum v (3.3) vedno obstaja. Potem je tudi $\{x \in \mathbb{R} \mid F(x) < \alpha\} = (-\infty, c)$, kar pojasni enakost v (3.3).

Kvantilna funkcija $F^\wedge$ je posplošitev inverza porazdelitvene funkcije F. Če je F zvezna in naraščujoča, je namreč $F^\wedge = F^{-1}$.

Lema 3.24. *Naj bo* F *porazdelitvena funkcija slučajne spremenljivke* X*. Potem veljajo naslednje trditve:*

1. *Funkcija* F *je zvezna v točki* $x \in \mathbb{R}$ *natanko tedaj, ko je* $\Delta F(x) = 0$.
2. $F_- = \mathsf{P}(X < \cdot)$.
3. $\Delta F = \mathsf{P}(X = \cdot)$.
4. *Funkcija* F *ima števno mnogo skokov.*
5. *Za vsak* $\alpha \in (0, 1)$ *velja*

$$\alpha \in \Big[F_-\big(F^\wedge(\alpha)\big), F\big(F^\wedge(\alpha)\big)\Big].$$

Torej, če je F *zvezna v točki* $F^\wedge(\alpha)$*, je* $F\big(F^\wedge(\alpha)\big) = \alpha$.

6. *Naj bo x realno število. Za vsak*

$$\alpha \in \big(F_-(x), F(x)\big] \cap (0,1)$$

velja $F^\wedge(\alpha) = x$.

7. *Za $\alpha \in (0,1)$ in realno število x velja ekvivalenca*

$$\alpha \leq F(x) \quad \Longleftrightarrow \quad F^\wedge(\alpha) \leq x.$$

Dokaz. Spomnimo se, da F_- povsod obstaja, saj je F nepadajoča.

1. Ker je F porazdelitvena funkcija, je desno zvezna. Torej je zvezna natanko v tistih točkah, kjer je levo zvezna oz. je $F - F_- = \Delta F = 0$.

2. Naj bo $x \in \mathbb{R}$ in $(s_n)_{n\in\mathbb{N}}$ naraščujoče zaporedje realnih števil z limito x. Potem je

$$\lim_{n\to\infty} F(s_n) = \lim_{n\to\infty} \mathsf{P}(X \leq s_n).$$

Upoštevamo zveznost mere, naraščanje zaporedja števil $(s_n)_{n\in\mathbb{N}}$ in posledično tudi zaporedja množic $\big\{\{X \leq s_n\}\big\}_{n\in\mathbb{N}}$, da dobimo

$$\lim_{n\to\infty} \mathsf{P}(X \leq s_n) = \mathsf{P}\left(\bigcup_{n\in\mathbb{N}}\{X \leq s_n\}\right) = \mathsf{P}\left(X < \sup_{n\in\mathbb{N}} s_n\right) = \mathsf{P}(X < x).$$

Ker je bilo zaporedje $(s_n)_{n\in\mathbb{N}}$ poljubno, sledi

$$F_-(x) = \mathsf{P}(X < x).$$

3. Sledi iz prejšnje točke:

$$\mathsf{P}(X = \cdot) = \mathsf{P}(X \leq \cdot) - \mathsf{P}(X < \cdot) = F - F_- = \Delta F.$$

4. Mera P_X je verjetnostna, torej σ-končna, zato je po lemi 2.41 izraz

$$\Delta F(x) = \mathsf{P}(X = x) = \mathsf{P}_X\big(\{x\}\big)$$

neničeln za števno mnogo $x \in \mathbb{R}$. Po prvi točki to pomeni, da je skokov števno mnogo.

5. Naj bo $(s_n)_{n\in\mathbb{N}}$ naraščujoče zaporedje realnih števil z limito $F^\wedge(\alpha)$. Ker je $s_n < F^\wedge(\alpha)$, je $F(s_n) < \alpha$, za vsak n. Torej je

$$\lim_{n\to\infty} F(s_n) \leq \alpha.$$

Zaporedje $(s_n)_{n\in\mathbb{N}}$ je bilo poljubno, to pa pomeni, da je $F_-\big(F^\wedge(\alpha)\big) \leq \alpha$. Druga neenakost $F\big(F^\wedge(\alpha)\big) \geq \alpha$ pa sledi iz definicije $F^\wedge$.

6. Ker je $\alpha > F_-(x)$, je $\alpha > F(s)$ za vsako število $s \in (-\infty, x)$, saj je F nepadajoča. Torej s ni iz množice $\{y \in \mathbb{R} \mid F(y) \geq \alpha\}$, za vsak $s < x$. Sledi, da tudi minimum te množice, tj. $F^\wedge(\alpha)$, ne more biti manjši od x. Po drugi strani pa, ker je $\alpha \leq F(x)$, iz definicije $F^\wedge$ sledi še $F^\wedge(\alpha) \leq x$.

7. Če je $\alpha \leq F(x)$, potem je
$$x \in \{y \in \mathbb{R} \mid F(y) \geq \alpha\},$$
zato je
$$F^\wedge(\alpha) = \min\{y \in \mathbb{R} \mid F(y) \geq \alpha\} \leq x.$$
Obratno, če je $F^\wedge(\alpha) \leq x$, sledi iz točke 5 in dejstva, da je F nepadajoča
$$\alpha \leq F\big(F^\wedge(\alpha)\big) \leq F(x).$$ □

Izrek 3.25 (Posplošena porazdelitvena transformacija). *Naj bosta $X \sim F$ in $V \sim U[0,1]$ neodvisni slučajni spremenljivki na verjetnostnem prostoru $(\Omega, \mathcal{A}, \mathsf{P})$. Potem je slučajna spremenljivka*
$$F_-(X) + V\Delta F(X)$$
enakomerno porazdeljena na $[0,1]$ in velja
$$F^\wedge\big(F_-(X) + V\Delta F(X)\big) = X \qquad \mathsf{P}\text{-skoraj gotovo.} \tag{3.4}$$

Opomba 3.26. Enakost (3.4) pove mnogo več, kot samo, da je $F^\wedge(Z) \sim F$ za $[0,1]$-enakomerno porazdeljeno s.s. U. Pove, da sta spremenljivki na obeh straneh enačbe skoraj vedno isti, to pa nam dovoli, da ju v vseh izračunih verjetnosti lahko prosto zamenjamo (tudi, če nastopata le kot kot komponenti slučajnega vektorja).

Izrek 3.25 nadgradi primer 3.9 s to razliko, da F v točkah nezveznosti t prilagodimo z izbiro enakomerno naključne vrednosti med $F_-(t)$ in $F(t)$. Intuitivno je smiselno, da na ta način dobimo $[0,1]$-enakomerno porazdelitev (ampak morda je smiselno tudi, da dobimo katero drugo porazdelitev – intuicija je velikokrat koristna, ampak brez dokaza ji ne smemo popolnoma zaupati).

Dokaz izreka 3.25. Izračunajmo porazdelitveno funkcijo s.s. $F_-(X) + V\Delta F(X)$! Za $\alpha \in (0,1)$ razdelimo dogodek $\{F_-(X) + V\Delta F(X) \leq \alpha\}$ na disjunktne primere
$$\mathsf{P}\big(F_-(X) + V\Delta F(X) \leq \alpha\big) = \mathsf{P}\big(F_-(X) + V\Delta F(X) \leq \alpha, X < F^\wedge(\alpha)\big) \tag{3.5}$$
$$+ \mathsf{P}\big(F_-(X) + V\Delta F(X) \leq \alpha, X = F^\wedge(\alpha)\big) \tag{3.6}$$
$$+ \mathsf{P}\big(F_-(X) + V\Delta F(X) \leq \alpha, X > F^\wedge(\alpha)\big) \tag{3.7}$$

Obravnavajmo vsakega posebej.

1. Po definiciji $F^\wedge(\alpha)$ je
$$\big\{X < F^\wedge(\alpha)\big\} \subseteq \big\{F(X) < \alpha\big\},$$
ker pa je $F_-(X) + V\Delta F(X) \leq F(X)$, je nadalje
$$\big\{F(X) < \alpha\big\} \subseteq \big\{F_-(X) + V\Delta F(X) \leq \alpha\big\},$$
saj je F nepadajoča. To pomeni
$$(3.5) = \mathsf{P}\big(X < F^\wedge(\alpha)\big) = F_-\big(F^\wedge(\alpha)\big).$$

2. Če je funkcija F zvezna v točki $F^\wedge(\alpha)$, je $\Delta F\big(F^\wedge(\alpha)\big) = 0$. Potem iz monotonosti mere sledi

$$(3.6) \leq \mathsf{P}\big(X = F^\wedge(\alpha)\big) = \Delta F\big(F^\wedge(\alpha)\big) = 0.$$

Naj bo zdaj F nezvezna v $F^\wedge(\alpha)$, tj. naj bo $\Delta F\big(F^\wedge(\alpha)\big) > 0$. Uporabimo podatek, da sta spremenljivki X in V neodvisni, da dobimo

$$\begin{aligned}(3.6) &= \mathsf{P}\Big(F_-\big(F^\wedge(\alpha)\big) + V\Delta F\big(F^\wedge(\alpha)\big) \leq \alpha, X = F^\wedge(\alpha)\Big)\\ &= \mathsf{P}\left(V \leq \frac{\alpha - F_-\big(F^\wedge(\alpha)\big)}{\Delta F\big(F^\wedge(\alpha)\big)}, X = F^\wedge(\alpha)\right)\\ &= \mathsf{P}\left(V \leq \frac{\alpha - F_-\big(F^\wedge(\alpha)\big)}{\Delta F\big(F^\wedge(\alpha)\big)}\right)\mathsf{P}\big(X = F^\wedge(\alpha)\big)\end{aligned}$$

Po lemi 3.24.5 je

$$\frac{\alpha - F_-\big(F^\wedge(\alpha)\big)}{\Delta F\big(F^\wedge(\alpha)\big)} \in [0,1],$$

in, ker je V enakomerno porazdeljena na $[0,1]$, sledi

$$(3.6) = \frac{\alpha - F_-\big(F^\wedge(\alpha)\big)}{\Delta F\big(F^\wedge(\alpha)\big)}\Delta F\big(F^\wedge(\alpha)\big) = \alpha - F_-\big(F^\wedge(\alpha)\big).$$

Nazadnje opazimo še, da je pravzaprav

$$(3.6) = \alpha - F_-\big(F^\wedge(\alpha)\big)$$

tudi v primeru, ko je F zvezna v $F^\wedge(\alpha)$, saj je takrat po lemi 3.24.5

$$F_-\big(F^\wedge(\alpha)\big) = F\big(F^\wedge(\alpha)\big) = \alpha.$$

3. Intuitivno je lahko $(3.7) > 0$ samo, če je F konstantna na $\big[F^\wedge(\alpha), F^\wedge(\alpha) + \varepsilon\big]$ za nek $\varepsilon > 0$ in je X iz tega intervala. Dokažimo.
Ker je F nepadajoča, za $x > y$ velja tako $F(x) \geq F(y)$ kot tudi $F_-(x) \geq F(y)$. Torej je

$$\begin{aligned}\big\{X > F^\wedge(\alpha)\big\} &= \Big\{F_-(X) \geq F\big(F^\wedge(\alpha)\big), X > F^\wedge(\alpha)\Big\}\\ &= \big\{F_-(X) \geq \alpha, X > F^\wedge(\alpha)\big\} \quad \text{(Po 3.24.5 je } F\big(F^\wedge(\alpha)\big) \geq \alpha.\text{)}\end{aligned}$$

Zapišimo

$$\begin{aligned}(3.7) &= \mathsf{P}\big(F_-(X) + V\Delta F(X) \leq \alpha, F_-(X) \geq \alpha, X > F^\wedge(\alpha)\big)\\ &= \mathsf{P}\big(V \leq 0, F_-(X) + V\Delta F(X) \leq \alpha, F_-(X) \geq \alpha, X > F^\wedge(\alpha)\big) \qquad (3.8)\\ &\quad + \mathsf{P}\big(V > 0, F_-(X) + V\Delta F(X) \leq \alpha, F_-(X) \geq \alpha, X > F^\wedge(\alpha)\big) \qquad (3.9)\end{aligned}$$

Ker je F nepadajoča in sta spremenljivki X in $V \sim U[0,1]$ neodvisni, je

$$\begin{aligned}
(3.8) &\leq \mathsf{P}(V \leq 0) = 0 \\
(3.9) &= \mathsf{P}\big(V > 0, F_-(X) = \alpha, \Delta F(X) = 0, X > F^\wedge(\alpha)\big) \\
&= \mathsf{P}(V > 0)\mathsf{P}\big(F_-(X) = \alpha, \Delta F(X) = 0, X > F^\wedge(\alpha)\big) \\
&= \mathsf{P}\big(F(X) = \alpha, \Delta F(X) = 0, X > F^\wedge(\alpha)\big) \\
&= \mathsf{P}(X \in B_\alpha)
\end{aligned}$$

Tu je množica B_α definirana z

$$B_\alpha = \big\{x \in \mathbb{R} \mid F(x) = \alpha, \Delta F(x) = 0, x > F^\wedge(\alpha)\big\}. \tag{3.10}$$

Preučimo jo. Najprej je jasno, da je $B_\alpha \subseteq (F^\wedge(\alpha), \infty)$. Nadalje, če je $x \in (\inf B_\alpha, \sup B_\alpha)$, je tudi x iz B_α, ker je F nepadajoča. To pomeni, da je B_α interval.
Če je $B_\alpha = \emptyset$, potem je

$$(3.7) = \mathsf{P}(X \in B_\alpha) = 0.$$

Sicer pa pokažimo, da je $\inf B_\alpha = F^\wedge(\alpha)$. Če to ne drži, potem je $\inf B_\alpha > F^\wedge(\alpha)$ in zato po lemi 3.24.4 obstaja tako število $c \in \big(F^\wedge(\alpha), \inf B_\alpha\big)$, da je $\Delta F(c) = 0$ (saj je množica $\big(F^\wedge(\alpha), \inf B_\alpha\big)$ neštevna). Ker je $c > F^\wedge(\alpha)$ in $c \notin B_\alpha$, mora biti $F(c) \neq \alpha$. Nadalje, ker je B_α neprazna, obstaja $b \in B_\alpha$ in zanj velja $c < b$ ter $F(b) = \alpha$. To pa je v nasprotju z neenakostjo $F\big(F^\wedge(\alpha)\big) \geq \alpha$ in dejstvom, da je F nepadajoča.
Pokazali smo, da je

$$B_\alpha = \big(F^\wedge(\alpha), \sup B_\alpha\big) \qquad \text{ali} \qquad B_\alpha = \big(F^\wedge(\alpha), \sup B_\alpha\big].$$

Če je $\sup B_\alpha \in B_\alpha$, je F zvezna v točki $\sup B_\alpha$, kar pomeni $F_-(\sup B_\alpha) = F(\sup B_\alpha)$. V vsakem primeru pa je

$$F_-(\sup B_\alpha) = \lim_{s \uparrow x} F(s) = \lim_{s \uparrow x} \alpha = \alpha,$$

saj je s iz B_α za dovolj velike s. Torej ne glede na to, ali je $\sup B_\alpha \in B_\alpha$, dobimo

$$(3.7) = \mathsf{P}(X \in B_\alpha) = F_-(\sup B_\alpha) - F\big(F^\wedge(\alpha)\big) = \alpha - F\big(F^\wedge(\alpha)\big) \leq \alpha - \alpha = 0.$$

Končno lahko seštejemo

$$\begin{aligned}
\mathsf{P}\big(F_-(X) + V\Delta F(X) \leq \alpha\big) &= (3.5) + (3.6) + (3.7) \\
&= F_-\big(F^\wedge(\alpha)\big) + \Big(\alpha - F_-\big(F^\wedge(\alpha)\big)\Big) + 0 = \alpha
\end{aligned}$$

Edina porazdelitvena funkcija, ki ustreza temu pogoju za vsak $\alpha \in (0,1)$, pa je porazdelitvena funkcija enakomerne porazdelitve na $[0,1]$.

Pokažimo še, da je

$$F^\wedge\big(F_-(X) + V\Delta F(X)\big) = X \qquad \text{s.g.}$$

Najprej utemeljimo, da je ta izraz smiseln – funkcija $F^\wedge$ je definirana na $(0,1)$, spremenljivka $F_-(X) + V\Delta F(X)$ pa je je skoraj gotovo element tega intervala, saj je enakomerno porazdeljena na $[0,1]$. To pomeni, da je izraz

$$F^\wedge\big(F_-(X) + V\Delta F(X)\big)$$

skoraj gotovo definirana slučajna spremenljivka.
Naj bo $A \subseteq \mathbb{R} \times (0,1)$ množica vseh parov (x,v), za katere je enakost

$$F^\wedge\big(F_-(x) + v\Delta F(x)\big) = x. \tag{3.11}$$

definirana in resnična. Pokazati torej želimo, da je

$$\mathsf{P}\big((X,V) \in A\big) = 1.$$

Naj bo x realno število. Ker je F desno zvezna in nepadajoča, so tri disjunktne možnosti: ali ima funkcija F levi skok pri x, ali je F zvezna in naraščujoča na $(x-\varepsilon, x]$ za nek $\varepsilon > 0$, ali pa je konstantna na $(x-\varepsilon, x]$ za nek $\varepsilon > 0$. Obravnavajmo vsako posebej.

1. Če ima F skok pri x, potem je

   $$F_-(x) + v\Delta F(x) \in \big(F_-(x), F(x)\big)$$

   za vsak $v \in (0,1)$. Zato je po lemi 3.24.6 enakost (3.11) izpolnjena.

2. Če je F zvezna in naraščujoča na $(x-\varepsilon, x]$ za nek $\varepsilon > 0$, potem je

   $$F^\wedge\big(F_-(x) + v\Delta F(x)\big) = F^\wedge\big(F(x)\big) = x$$

   po definiciji $F^\wedge$, saj je $F(x) > F(y)$ za vse $y < x$. To velja ne glede na v. Enakost 3.11 je tudi v tem primeru izpolnjena.

3. Naj bo zdaj F konstantna na $(x-\varepsilon, x]$ za nek $\varepsilon > 0$. Če je $F(x) \in \{0,1\}$, potem enačba 3.11 ni definirana in posledično $(x,v) \notin A$ za noben v. Naj bo zdaj $F(x) \in (0,1)$. Jasno je

   $$F^\wedge\big(F_-(x) + v\Delta F(x)\big) = F^\wedge\big(F(x)\big) < x$$

   za vse v, kar pomeni, da tokrat enakost (3.11) ne drži, očitno pa je tudi $x \in B_{F(x)}$, kjer je množica $B_{F(x)}$ definirana z enačbo (3.10). Torej je

   $$\big\{x \in \mathbb{R} \mid F(x) \in (0,1), F \text{ konstantna na } (x-\varepsilon, x] \text{ za nek } \varepsilon > 0\big\} \subseteq \bigcup_{\alpha \in (0,1)} B_\alpha.$$

   Pokažimo, da je ta unija števna. Ideja dokaza je, da so območja konstantnosti F v očitni bijekciji s skoki funkcije $F^\wedge$, poleg tega pa $F^\wedge$ ni tako zelo različna od neke porazdelitvene funkcije slučajne spremenljvike, za katere pa vemo (lema 3.24.4), da ima števno mnogo skokov.

Naj bo $\alpha \in (0,1)$. Vemo že, da je B_α interval. Če je B_α prazna množica, k uniji nič ne prispeva in jo lahko izločimo. Sicer pa vsebuje vsaj dva elementa $a < b$, za katera je $F(a) = F(b)$. To pomeni, da ima funkcija $F^\wedge$ skok v točki $F(a) \in (0,1)$. Očitno je $F^\wedge$ tudi nepadajoča funkcija, zato obstaja njena desno zvezna verzija $F_+^\wedge$, za katero potem velja $\Delta F_+^\wedge\big(F(a)\big) > 0$. Nadalje lahko, podobno kot v dokazu izreka 3.7, definiramo mero $\mu_{F_+^\wedge}$ na merljivem prostoru $\big((0,1), \mathcal{B}_{(0,1)}\big)$, katere porazdelitvena funkcija je $F_+^\wedge$. Ta mera ni verjetnostna, je pa še vedno σ-končna, saj je interval $(0,1)$ števna unija $\bigcup_{n\in\mathbb{N}}(1/n, 1-1/n]$ množic s končno mero

$$\mu_{F_+^\wedge}\left(\left(\frac{1}{n}, 1-\frac{1}{n}\right]\right) = F_+^\wedge\left(1-\frac{1}{n}\right) - F_+^\wedge\left(\frac{1}{n}\right) < \infty.$$

Velja

$$\mu_{F_+^\wedge}\Big(\big\{F(a)\big\}\Big) = \Delta F_+^\wedge\big(F(a)\big) > 0,$$

po lemi 2.41 pa je to možno le za števno mnogo vrednosti $F(a)$, kar pomeni, da je le števno mnogo različnih množic B_α.

Pokazali smo, da je

$$\mathsf{P}\big((X,V) \in A\big) \geq \mathsf{P}\left((X,V) \in \left(\mathbb{R} \setminus \bigcup_{\alpha\in S} B_\alpha\right) \times (0,1)\right),$$

kjer je S neka števna množica. Vemo že, da je $\mathsf{P}(X \in B_\alpha) = 0$ za vse α, nato pa upoštevamo še neodvisnost spremenljivk X in $V \sim U[0,1]$ ter subaditivnost mere, da dobimo

$$\mathsf{P}\big((X,V) \in A\big) \geq \left(1 - \mathsf{P}\left(X \in \bigcup_{\alpha\in S} B_\alpha\right)\right)\mathsf{P}\big(V \in (0,1)\big) \geq 1 - \sum_{\alpha\in S}\mathsf{P}(X \in B_\alpha) = 1.$$

□

Dokaz Sklarovega izreka. Dokaz je analog primera 3.9 z uporabo posplošene porazdelitvene transformacije (izreka 3.25).

Iz predpostavk Sklarovega izreka sledi, da obstaja slučajni vektor $\mathbf{X} = (X_1, \ldots, X_n)$ s porazdelitveno funkcijo F. Naj bo s.s. V enakomerno porazdeljena na $[0,1]$ in neodvisna od $\mathbf{X}$ (npr. tukaj se uporabi opombo 3.4). Po pravkar dokazanem izreku 3.25 ima slučajni vektor

$$\big(F_i(X_i) + V\Delta F_i(X_i)\big)_{i=1}^n$$

komponente porazdeljene enakomerno na $[0,1]$. Torej je njegova porazdelitvena funkcija kopula; označimo jo s C_F in z upoštevanjem enakosti

$$F_i^\wedge\big(F_i(X_i) + V\Delta F_i(X_i)\big) = X_i \qquad \text{s.g.}$$

za $i = 1, \ldots, n$ ter leme 3.24.7 preverimo

$$\begin{aligned}
&F(x_1, \ldots, x_n) \\
&= \mathsf{P}(X_1 \le x_1, \ldots, X_n \le x_n) \\
&= \mathsf{P}\Big(F_1^{\wedge}\big(F_1(X_1) + V\Delta F_1(X_1)\big) \le x_1, \ldots, F_n^{\wedge}\big(F_n(X_n) + V\Delta F_n(X_n)\big) \le x_n\Big) \\
&= \mathsf{P}\big(F_1(X_1) + V\Delta F_1(X_1) \le F_1(x_1), \ldots, F_n(X_n) + V\Delta F_n(X_n) \le F_n(x_n)\big) \\
&= C_F\big(F_1(x_1), \ldots, F_n(x_n)\big)
\end{aligned}$$

Obratna trditev je z vidika izreka 3.7 in opomb pri definiciji kopule 3.10 trivialna. □

4 Singularni deli kopul

Za modeliranje pojavov iz resničnega sveta bi nam vsekakor prav prišle kopule, ki imajo tako nesingularen kot singularen del. Na primer, lahko bi gledali škodne zahtevke, kot v primeru 2.1, a še bolj splošno: imamo več virov zahtevkov, ki so medsebojno odvisni. Drug primer je modeliranje življenskih časov (npr. internetnih strežnikov). V tem modelu bi zvezni del kopule opisoval običajno porazdelitev žiljenskih časov, medtem ko bi singularni del poskrbel, da imajo izidi z globalnim učinokm (npr. splošen izpad elektrike), kjer se vsi življenski časi hkrati iztečejo, pozitivno verjetnost.
Pri singularnih komponentah je treba bolj paziti kot sicer, saj morajo robne porazdelitve vsake kopule biti popolnoma nesingularne – enakomerno porazdeljene na $[0, 1]$.

Pogledali si bomo konstrukcije kopul s predpisanimi singularnimi deli – v našem primeru si bomo dovolili predpisati ali porazdelitev velikosti singularnega dela, ali pa krivuljo, na kateri naj bo singularni del skoncentriran. Glavni vir za ta razdelek je članek [2].

Iz splošnih n dimenzij se bomo zdaj spustili na 2.

4.1 Markovska jedra

Spomnimo se pogojnih porazdelitev iz osnov verjetnosti.
Slučajnemu vektorju (X, Y) na verjetnostnem prostoru $(\Omega, \mathcal{A}, \mathsf{P})$ pripada dvorazsežna porazdelitev $\mathsf{P}_{X,Y} = \mathsf{P}\big((X, Y) \in \cdot\big)$ z enorazsežnima robnima porazdelitvama $P_X = \mathsf{P}_{X,Y}(\cdot \times \mathbb{R})$ in $\mathsf{P}_Y = \mathsf{P}_{X,Y}(\mathbb{R} \times \cdot)$. Če je vektor diskreten (tj. mera $\mathsf{P}_{X,Y}$ ima z vidika trditve 2.42 samo diskretni del neničeln), potem pogojno porazdelitev $\mathsf{P}_{Y|X=x} = \mathsf{P}\big(Y \in \cdot \mid X = x\big)$ slučajne spremenljivke Y pri pogoju $X = x$ poznamo kot enorazsežno porazdelitev, podano z

$$\mathsf{P}_{Y|X=x} = \mathsf{P}\big(Y \in \cdot \mid X = x\big) = \frac{\mathsf{P}(Y \in \cdot, X = x)}{\mathsf{P}(X = x)} = \frac{\mathsf{P}_{X,Y}\big(\{x\} \times \cdot\big)}{\mathsf{P}_X\big(\{x\}\big)}.$$

Seveda je definirana le za tiste x, za katere je $\mathsf{P}(X = x) > 0$. Vemo tudi, da potem velja zakon popolne verjetnosti:

$$\mathsf{P}_{X,Y}(A) = \sum_x \mathsf{P}_{Y|X=x}(A_x)\mathsf{P}_X(\{x\}). \tag{4.1}$$

Tu je A poljubna dvorazsežna Borelova množica in A_x njena *x-rezina*

$$A_x = \{y \in \mathbb{R} \mid (x, y) \in A\},$$

vsota pa teče po tistih x, za katere je $\mathsf{P}_X(\{x\}) = \mathsf{P}(X = x) > 0$.
V primeru, ko je vektor (X, Y) zvezen, oz. bolje rečeno, če je njegova porazdelitev $\mathsf{P}_{X,Y}$ absolutno zvezna (glede na dvorazsežno Lebesguovo mero λ_2) z gostoto $f_{X,Y}$, postopamo podobno. Robni porazdelitvi P_X in P_Y sta v tem primeru tudi absolutno zvezni (glede na enorazsežno Lebesguovo mero λ) z gostotama f_X in f_Y, pogojno porazdelitev $\mathsf{P}_{Y|X=x}$ pa poznamo kot enorazsežno porazdelitev, ki je absolutno zvezna z gostoto

$$f_{Y|X=x} = \frac{f_{X,Y}(x, \cdot)}{f_X(x)}.$$

Tudi tokrat je definicija veljavna le za tiste x, za katere je $f_X(x) > 0$ in zakon popolne verjetnosti za absolutno zvezne porazdelitve se glasi

$$\mathsf{P}_{X,Y}(A) = \int_{\{x \in \mathbb{R} \,|\, f_X(x) > 0\}} \mathsf{P}_{Y|X=x}(A_x) f_X(x)\, d\lambda(x). \tag{4.2}$$

Pogojne porazdelitve se da posplošiti tudi na primer, ko je porazdelitev $\mathsf{P}_{X,Y}$ poljubna. Žal je definicija precej bolj abstraktna in težje predstavljiva, kot v primerih diskretnih ali absolutno zveznih porazdelitev. Obstaja pa lep način, ki je malo manj (a vseeno dovolj) splošen in se tem težavam izogne. Ideja je, da tudi v splošnem pogojne porazdelitve definiramo kot tiste mere, za katere velja zakon popolne verjetnosti.
Ponovno si oglejmo izraza (4.1) in (4.2). V zveznem primeru je f_X gostota porazdelitve P_X glede na λ, tj. $f_X = \frac{d\mathsf{P}_X}{d\lambda}$, zato je

$$(4.2) = \int_{\{x \in \mathbb{R} \,|\, f_X(x) > 0\}} \mathsf{P}_{Y|X=x}(A_x)\, d\mathsf{P}_X(x) = \int \mathsf{P}_{Y|X=x}(A_x)\, d\mathsf{P}_X(x),$$

saj je množica

$$\{x \in \mathbb{R} \mid f_X(x) > 0\}^{\mathsf{c}} = \{x \in \mathbb{R} \mid f_X(x) = 0\}$$

P_X-ničelna (merljiva preslikava $x \mapsto \mathsf{P}_{Y|X=x}(A_x)$ je torej definirana le skoraj gotovo).
Zanimivo (a ne presenetljivo) je, da je tudi vsota v (4.1) pravzaprav integral. Res, ker je mera P_X diskretna, je absolutno zvezna glede na mero štetja ς in gostota $\frac{d\mathsf{P}_X}{d\varsigma}$ je ravno $\mathsf{P}_X(\{\cdot\})$. Torej je

$$(4.1) = \int \mathsf{P}_{Y|X=x}(A_x) \mathsf{P}_X(\{x\})\, d\varsigma = \int \mathsf{P}_{Y|X=x}(A_x)\, d\mathsf{P}_X(x),$$

kot v zveznem primeru! Zato definiramo:

Definicija 4.1. Naj bo (X, Y) slučajni vektor. *Sistem pogojnih verjetnosti* oz. *Markovsko jedro* Y glede na X je preslikava

$$K : \mathbb{R} \times \mathcal{B}_{\mathbb{R}} \longrightarrow [0, 1]$$

z lastnostmi:

- $K(\cdot, A) : \mathbb{R} \to [0,1]$ je slučajna spremenljivka za vsako Borelovo množico A.
- $K(x, \cdot) : \mathcal{B}_{\mathbb{R}} \to [0,1]$ je porazdelitev za P_X-skoraj vse x.
- Velja zakon popolne verjetnosti: za vsako dvorazsežno Borelovo množico A je

$$\mathsf{P}_{X,Y}(A) = \int K(x, A_x)\, d\mathsf{P}_X(x).$$

Opomba 4.2. Če je vektor (X, Y) porazdeljen po kopuli C, se pripadajoče Markovsko jedro označi s K_C. Nanj lahko gledamo kot na preslikavo $[0,1] \times \mathcal{B}_{[0,1]} \to [0,1]$.

Markovsko jedro si lahko predstavljamo kot razrez dvorazsežne porazdelitve v mnogo enorazsežnih, podobno, kot razrez ravnine v neskončno vzporednih premic. Videli smo že, da za diskretne in absolutno zvezne vektorje vedno obstaja Markovsko jedro. To pa velja tudi v splošnem, tj. za poljubno porazdeljene slučajne vektorje. Nadalje je P_X-skoraj gotovo enoličen. Dokaz tega dejstva ni enostaven, najdemo pa ga lahko v [4, str. 108].

4.2 Kopule s predpisanimi singularnimi deli

Poglejmo si, kako so singularne mere povezane s kopulami. Ali sploh lahko obstajajo? Konec koncev morajo robne porazdelitve kopul biti enakomerne, te pa so zelo daleč od singularnih. Najprej si ustvarimo nekaj predstave o problemu.
Kot prikazano v primeru 2.1, si enorazsežno porazdelitev, ki v neki točki vsebuje diskretno komponento, lahko neformalno predstavjamo kot zvezno porazdelitev, katere gostota v tej točki je neskončna. V dveh razsežnostih to še vedno velja – gostota je dvorazsežna funkcija in za modeliranje diskretne vrednosti v določeni točki si še vedno lahko predstavljamo, da je gostota v tej točki neskončna. Ampak tokrat imamo še eno možnost: gostota je lahko neskončna na poljubni enorazsežni podmnožici ravnine!

Primer 4.3. Naj bo porazdelitev P na $\mathcal{B}_{[0,1]^2}$ podana z[13]

$$\mathsf{P}(A) = \lambda\big(\pi_1(A \cap \diagup)\big),$$

kjer je $\diagup$ diagonala enotskega kvadrata in π_1 projekcija na prvo komponento. Porazdelitev P torej vsaki množici dodeli dolžino diagonale, ki jo seka. Očitno ne vsebuje diskretnih delov, saj ni nobene enoelementne množice $\{(x, y)\}$, ki bi ji P dodelila pozitivno vrednost, niti ni absolutno zvezna, saj bi bil integral gostote po diagonali enak 0, ker je diagonala λ_2-ničelna množica. Torej je ta mera singularno zvezna. Neformalno si jo lahko predstavljamo kot porazdeitev z neskončno gostoto na diagonali. Ampak, ta neskončnost je v nekem smislu šibkejša, kot bi bila neskončnost gostote diskretne porazdelitve, saj ima pri diskretni porazdelitvi vsaka točka z neskončno gostoto tudi pozitivno verjetnost, tukaj pa je verjetnost enoelementne množice 0, čeprav ima njen element lahko enskončno gostoto.

[13]Porazdelitvena funkcija te mere je komonotonostna kopula M_2.

Vse to je seveda samo za predstavo, v resnici bomo delali s singularnimi merami namesto z neskončnimi gostotami.

Ta primer pokaže, da so v več kot eni razsežnosti zvezno singularne porazdelitve zelo naravne.

Naj bo C kopula. Pripadajočo mero P_C lahko, kot vemo iz trditve 2.42, razcepimo na tri dele

$$\mathsf{P}_C = \mathsf{P}_C^a + \mathsf{P}_C^s + \mathsf{P}_C^d.$$

Tu je P_C^a absolutno zvezna, P_C^s singularno zvezna in P_C^d diskretna mera, med njimi pa v splošnem nobena ni verjetnostna.

Trditev 4.4. *Diskretni del vsake kopule (oz. pripadajoče mere) je* 0.

Dokaz. Porazdelitvena funkcija vsote mer je vsota porazdelitvenih funkcij posamezne mere. Enako velja za robne porazdelitvene funkcije. Robni porazdelitveni funkciji absolutno zveznega dela sta zvezni, prav tako pa so zvezne robne porazdelitvene funkcije vsake kopule (saj so enake $\mathbb{1}_{(0,1)}\mathrm{Id}+\mathbb{1}_{[1,\infty)})$. Torej, če pokažemo, da so robne porazdelitvene funkcije singularno zveznega dela tudi zvezne, bo to moralo veljati tudi za diskretni del, kar pa je možno le, če je enak 0.
Naj bo C kopula, P_C pripadajoča porazdelitev in P_C^s njen singularno zvezni del. Za robno porazdelitev singularno zveznega dela $\mathsf{P}_C^s\big(\cdot \times [0,1]\big)$ potem velja

$$\mathsf{P}_C^s\big(A \times [0,1]\big) \leq \mathsf{P}_C\big(A \times [0,1]\big) = \lambda(A)$$

za vsako Borelovo množico A, saj so robne porazdelitve mere P_C enakomerne na $[0,1]$. Torej mora vsaka λ-ničelna množica biti tudi $\mathsf{P}_C^s\big(\cdot \times [0,1]\big)$-ničelna oz. mera $\mathsf{P}_C^s\big(\cdot \times [0,1]\big)$ je absolutno zvezna, iz česar pa sledi, da je porazdelitvena funkcija mere $\mathsf{P}_C^s\big(\cdot \times [0,1]\big)$ zvezna. □

Vidimo, da so edini singularni deli kopul singularno zvezni, torej taki, kot npr. v primeru 4.3. Zdaj se pojavi vprašanje: za katere podmnožice enotskega kvadrata obstaja taka kopula, da je predpisana množica najmanjša, na kateri je singularni del kopule skoncentriran? Vsekakor ne za vse – niti ne za vse krivulje. Protipimer je katerakoli daljica, vzporedna koordinatnima osema, saj bi v tem primeru ena od robnih porazdelitvenih funkcij imela skok. Rezultat iz članka [2] predstavi zadostni pogoj za konstrukcijo kopule s singularnim delom na dani množici.

Izrek 4.5. *Naj bo merljiva funkcija $f : [0,1] \to [0,1]$ taka, da je mera*

$$\lambda\big(f^{-1}[\cdot]\big) : \mathcal{B}_{[0,1]} \longrightarrow [0,1]$$

absolutno zvezna glede na Lebesguovo mero λ. Potem obstaja kopula C, katere singularni del je skoncentriran na grafu funkcije f in njeno Markovsko jedro vsaki točki grafa dodeli pozitivno verjetnost, tj.

$$K_C\Big(x, \big\{f(x)\big\}\Big) > 0$$

za vse $x \in [0,1]$.

Opomba 4.6. Lastnost, da Markovsko jedro vsaki točki grafa funkcije f dodeli pozitivno verjetnost pomeni, da ne obstaja prava podmnožica grafa funkcije f, na kateri bi bil skoncentriran singularni del kopule C, tj. graf funkcije f je najmanjša množica, na kateri je singularni del kopule skoncentriran.
To sledi iz dejstva, da je na območjih, kjer je kopula absolutno zvezna, tudi njeno Markovsko jedro absolutno zvezno (tako, kot so pri absolutno zveznih slučajnih vektorjih pogojne porazdelitve tudi absolutno zvezne) in zato enoelementni množici $\{f(x)\}$ dodeli vrednost 0.

Da si olajšamo dokaz, pa prej pokažimo še drugačen izrek.
Recimo, da bi zdaj namesto določitve množice, na kateri naj bo kopula singularna, rajši določili njen singularni profil. Pojasnimo, kaj to je. Naj bo C kopula in K_C pripadajoče Markovsko jedro. Potem lahko za vsak $x \in [0,1]$ mero $K_C(x,\cdot)$ razcepimo na absolutno zvezni, singularno zvezni in diskretni del (označimo ga s $K_C^d(x,\cdot)$). Predstavljamo si lahko, da diskretni deli vsake mere v Markovskem jedru sovpadajo s singularno zveznimi deli kopule. Zato definiramo *singularni profil* kot funkcijo

$$K_C^d(\cdot,[0,1]) : [0,1] \longrightarrow [0,1].$$

Singularni profil torej za vsak $x \in [0,1]$ pove, kolikšna je skupna verjetnost vseh diskretnih delov Markovskega jedra v točki x. Neformalno si lahko predstavljamo, da singularni profil predstavlja skupno gostoto singularno zveznega dela kopule na vsaki daljici $\{x\} \times [0,1]$.

Izrek 4.7. *Za vsako merljivo funkcijo $f : [0,1] \to [0,1]$ obstaja kopula s singularnim profilom f.*

Opomba 4.8. Merljive funkcije so zelo velik razred funkcij. Izrek torej zagotavlja, da obstajajo kopule s praktično kakršnimkoli singularnim profilom, če le ni (skoraj) nikoli večji od 1. Zakaj je omejitev pri 1? Intuitivno si lahko razložimo, da v primeru, ko je singularni profil večji od 1 na neki neničelni množici, bi ena izmed robnih porazdelitvenih funkcij naraščala hitreje kot z odvodom 1, to pa je hitreje, kot sploh lahko, da bi porazdelitev s takšnimi robnimi porazdelitvenimi funkcijami sploh lahko bila kopula, saj robne porazdelitve kopule naraščajo z odvodom 1. Če pa je profil kje (ali pa povsod) manjši od 1, pa podobnega problema ni – namreč vedno lahko dobljeni singularni meri prištejemo še tako absolutno zvezno mero, da se robne porazdelitve seštejejo ravno v enakomerne porazdelitve na $[0,1]$. Več o tem v dokazu izreka.

Dokaz izreka 4.7. Ideja dokaza je podobna primeru 4.3, le da tokrat skonstruiramo mero, ki ima na diagonali "gostoto" f. Ker je $\int f\,d\lambda$ lahko manjši od 1 in mora končna mera biti verjetnostna (da dobimo kopulo), pa bomo dodali še primerno absolutno zvezno mero.
Najprej, če je $\int f\,d\lambda = 0$, je $f = 0$ s.g. in katerakoli absolutno zvezna kopula bo imela tak singularni profil (ker je Markovsko jedro določeno le skoraj gotovo, je tudi singularni profil), npr. kopula neodvisnosti Π_2. Sicer, če je $\int f\,d\lambda = 1$, ima tak singularni profil komonotonostna kopula M_2 iz primera 4.3.
Naj bo zdaj $\int f\,d\lambda \in (0,1)$. Definirajmo mero μ na $\mathcal{B}_{[0,1]^2}$ z

$$\mu(A) = \int_{\pi_1[A\cap\diagup]} f\,d\lambda.$$

Tu (kot prej) π_1 označuje projekcijo na prvo komponento in $\diagup$ diagonalo enotskega kvadrata. Očitno je mera μ singularna (μ je dvorazsežna, zato se singularnost oz. absolutno zveznost obravnava glede na dvorazsežno Lebesguovo mero λ_2) in je njen singularni profil enak f. Vse kar nam ostane je, da meri μ prištejemo tako absolutno zvezno mero, da bodo robne porazdelitve vsote mer enakomerne. Po simetriji sta robni porazdelitveni funkciji mere μ obe enaki funkciji F, podani z

$$F(x) = \mu\big([0,x] \times [0,1]\big) = \int_{\pi_1\left[[0,x]\times[0,1]\cap\diagup\right]} f\,d\lambda = \int_{[0,x]} f\,d\lambda \leq x.$$

Zato lahko definiramo funkcijo

$$G = \mathrm{Id} - F = \int_{[0,\cdot]} (1-f)\,d\lambda,$$

ki bo robna porazdelitvena funkcija absolutno zvezne mere ν, dobljene iz dvorazsežne porazdelitvene funkcije

$$(x,y) \longmapsto \frac{G(x)G(y)}{G(1)}.$$

Tukaj je namen deljenja z $G(1)$ enakost

$$\nu\big([0,1]^2\big) = \frac{G(1)G(1)}{G(1)} = G(1) = 1 - F(1),$$

saj je potem mera $\mu+\nu$ verjetnostna, ker je $\mu\big([0,1]^2\big) = \int_{[0,1]} f\,d\lambda = F(1)$. Očitno sta tudi robni porazdelitveni funkciji mere $\mu+\nu$ enaki $F+G = \mathrm{Id}$, mera ν pa je res absolutno zvezna z gostoto

$$\frac{d\nu}{d\lambda_2}(x,y) = \frac{\big(1-f(x)\big)\big(1-f(y)\big)}{G(1)}.$$

To pomeni, da je porazdelitvena funkcija mere $\mu+\nu$ kopula, katere singularni profil je enak singularnemu profilu mere μ, to pa je f. □

Dokaz izreka 4.5. Ideja dokaza je ista kot pri prejšnjem: skonstruiramo primerno singularno mero, katere robne porazdelitvene funkcije so manjše od porazdelitvene funkcije enakomerne porazdelitve, nato pa ji prištejemo primerno absolutno zvezno mero, da dobimo kopulo.

Razlika med izrekom 4.7 in 4.5 pa je, da pri enem ustvarjamo kopule s predpisanim singularnim profilom, pri drugem pa s singularnim delom na predpisanem grafu funkcije. Treba pa se je zavedati, da se v splošnem obojega ne da predpisati. Namreč, če je predpisan graf funkcije zelo položen in ima hkrati predpisan velik singularni profil, potem bo robna porazdelitvena funkcija v smeri ordinate prehitro naraščala, da bi lahko bila enakomerna (ali pa manjša od enkomerne). Pravzaprav med položnostjo grafa in največjim možnim singularnim profilom obstaja določeno obratno sorazmerje, če želimo, da je konstrukcija predpisane kopule možna. Ugotovili smo že, da je pri "položnosti 1" (tj. naklon diagonale) največji možen predpisan singularni profil enak 1. Ali to lahko posplošimo na manj enostavne grafe funkcij?

Ključno je opažanje, da je za poljubno merljivo funkcijo f, ki ustreza predpostavki izreka, neposredna zveza med njenim naklonom in gostoto $\frac{d\lambda(f^{-1}[\cdot])}{d\lambda}$. Res, če je f položna v točki $x \in (0,1)$ potem je za majhen $\varepsilon > 0$ praslika $f^{-1}\big[(x-\varepsilon, x+\varepsilon)\big]$ relativno velika množica, kar pomeni, da je gostota $\frac{d\lambda(f^{-1}[\cdot])}{d\lambda}$ velika v $f(x)$. Učinek je obraten, če je graf f strm. Seveda ni nobenega razloga, da je f sploh kjerkoli zvezna, zato je tovrstno razmišljanje morda brezpomensko. Vseeno poizkusimo definirati funkcijo $s : [0,1] \to [0,1]$ z

$$s(x) = \left(\frac{d\lambda\big(f^{-1}[\cdot]\big)}{d\lambda}\big(f(x)\big)\right)^{-1} \wedge 1$$

in definirati kopulo s singularnim profilom s in singularnim delom skoncentriranim na grafu funkcije f. Pri singularnemu profilu smo se navzgor omejili z 1, saj bi sicer ena od robnih porazdelitvenih funkcij naraščala prehitro ne glede na to, kakšen je graf f.

Zdaj po zgledu dokaza izreka 4.7 definiramo dvorazsežno mero μ z

$$\mu(A) = \int_{\pi_1[A \cap G_f]} s\, d\lambda,$$

kjer G_f označuje graf funkcije f. Potem je

$$\mu\big([0,1]^2\big) = \int_{[0,1]} s\, d\lambda \leq 1$$

in, če z $F_1, F_2 : [0,1] \to [0,1]$ označimo robni porazdelitveni funkciji mere μ, je

$$\begin{aligned}
F_1(x) &= \mu\big([0,x] \times [0,1]\big) = \int_{[0,x]} s\, d\lambda \leq x \\
F_2(y) &= \mu\big([0,1] \times [0,y]\big) = \int_{f^{-1}[[0,y]]} s\, d\lambda \\
&= \int 1_{[0,y]}\big(f(x)\big)\left(\left(\frac{d\lambda\big(f^{-1}[\cdot]\big)}{d\lambda}\big(f(x)\big)\right)^{-1} \wedge 1\right) d\lambda(a) \\
&= \int 1_{[0,y]}\left(\left(\frac{d\lambda\big(f^{-1}[\cdot]\big)}{d\lambda}\right)^{-1} \wedge 1\right) d\lambda\big(f^{-1}[\cdot]\big) && \text{(Primer 2.28.)} \\
&= \int_{[0,y]}\left(\left(\frac{d\lambda\big(f^{-1}[\cdot]\big)}{d\lambda}\right)^{-1} \wedge 1\right)\frac{d\lambda\big(f^{-1}[\cdot]\big)}{d\lambda}\, d\lambda \\
&= \int_{[0,y]}\left(\frac{d\lambda\big(f^{-1}[\cdot]\big)}{d\lambda} \wedge 1\right) d\lambda \leq y
\end{aligned}$$

Ker sta obe robni porazdeliteveni funkciji manjši od identitete (robne porazdelitvene funkcije kopul), lahko uporabimo željeni trik in kot prej definiramo dvorazsežno porazdelitveno funkcijo

$$(x,y) \longmapsto \frac{\big(x - F_1(x)\big)\big(y - F_2(y)\big)}{1 - F_1(1)}$$

in pripadajočo mero označimo z ν. Tako sta robni porazdelitveni funkciji mere ν ravno $\mathrm{Id} - F_1$ in $\mathrm{Id} - F_2$, poleg tega pa je

$$(x,y) \longmapsto \frac{(1 - s(x))\left(1 - \frac{d\lambda\left(f^{-1}[\cdot]\right)}{d\lambda}(y) \wedge 1\right)}{1 - F(1)}$$

gostota mere ν (to je razvidno iz gostot funkcij F_1 in F_2), ki je posledično absolutno zvezna. Očitno mera $\mu + \nu$ izpolnjuje vse željene pogoje: njena porazdelitvena funkcija je kopula, njen singularni del je skoncentriran na grafu funkcije f in njen singularni profil je s. $\square$

Literatura

[1] P. Billingsley, *Probability and Measure*, John Wiley & Sons, 2012.

[2] F. Durante, J. Fernandez-Sanchez in W. Trutschnig, *On the singular components of a copula*, Journal of Applied Probability **52**(4) (2015) 1175–1182.

[3] F. Durante in C. Sempi, *Principles of copula theory*, Chapman and Hall/CRC, 2015.

[4] O. Kallenberg, *Foundations of Modern Probability, 2nd edn.(2002)*, Springer, New York, 1997.

[5] D. Khoshnevisan, *Probability (Graduate Studies in Mathematics)*, American Mathematical Society, 2007.

[6] F. Oertel, *An analysis of the rüschendorf transform – with a view towards sklar's theorem*, Dependence Modeling **3**(1) (2015).

[7] *Vitali set*, [ogled 19. 7. 2019], dostopno na `https://en.wikipedia.org/wiki/Vitali_set`.