Univerzitet u Beogradu
Matematički fakultet

MASTER RAD

Zavisnost slučajnih veličina pomoću kopula

Student: Ana Brajović
Br. indeksa: 1049/2012.

Beograd, decembar 2013.

Mentor: dr. Slobodanka Janković

Komisija: dr. Pavle Mladenović
dr. Jovan Vukmirović

Sadržaj

1 Uvod 3

2 Definicije i osnovna svojstva 5

2.1 Dvodimenzione kopule 5
2.2 Sklarova teorema 11
2.3 Kopule i slučajne promenljive 15
2.4 Freše-Hefdingove granice za zajedničke funkcije raspodela 18
2.5 Kopule preživljavanja 21
2.6 Višedimenzione kopule 23

3 Zavisnost 28

3.1 Mere saglasnosti, osnovne definicije 28
3.2 Kendalovo tau 32
3.3 Spirmanovo ro 34
3.4 Veza između Kendalovog tau i Spirmanovog ro 37
3.5 Ostale mere saglasnosti 41
3.6 Zavisnost, svojstva, mere zavisnosti 43
3.7 Švajcer-Volfovo sigma 45
3.8 Hefdingov indeks zavisnosti 46
3.9 Repna zavisnost 46
3.10 Empirijske kopule 48
3.11 Zavisnost više promenljivih 49

4 Zaključak 50

Reference 51

1 Uvod

Izučavanje kopula, sam njihov pojam i primene u statistici predstavljaju jedan moderan, savremen fenomen koji se sve više razvija i dobija na popularnosti. Do skoro samu reč „kopula“ bilo je teško, gotovo nemoguće pronaći u literaturi vezanoj za statistiku. Krajem sedamdesetih i početkom osamdesetih godina mogle su se pronaći prve reference i naznake pojma kopula. Naravno, bilo ih je zaista malo. U poslednjih petnaest godina počinje rast interesovanja za pojam kopula kao i za njihovu primenu u verovatnoći i statistici. U tom periodu održano je pet međunarodnih konferencija koje su bile posvećene idejama vezanih za pojam kopula:

- „Symposium on distribution with given marginals“ u Rimu 1990.
- „Distribution with fixed marginals, Doubly stochastic measures, and Markov operators“ u Sijetlu 1993.
- „Distributions with given marginals and Moment problems“ u Pragu 1996.
- „Distributions with given marginals and Statistical modelling“ u Barseloni 2000.
- „Dependence modelling: Statistical theory and applications in finance and insurance“ u Kvebeku 2004.

Može se primetiti iz naziva ovih konferencija da je pojam kopula usko povezan sa pojmom fiksiranih i zadatih marginalnih raspodela. Hačinson (Hutchinson) i Lai (Lai) su prvi autori koji su popularizovali teoriju kopula (početkom devedesetih). Nakon njih i drugi istraživači su počeli da objavljuju svoje radove koji su se bazirali na kopulama. Na primer Nelsen (Nelsen [16]) je predstavio sveobuhvatnu ulogu dvodimenzionalnih kopula, dok je Džo (Joe [9]) posvetio čitav odeljak u svojoj knjizi o višedimenzionim kopulama. Kasnije je Nelsen u svojim radovima proširivao teoriju kopula i jedan je od najznačajnijih naučnika kada su kopule u pitanju.

Reč kopula je imenica latinskog porekla koja se u gramatici opisuje kao deo rečenice koji povezuje subjekat i predikat. Ovaj pojam može se dovesti i u vezu sa engleskom reči „couple“, što znači par ili spoj, a ako gledamo na tu reč kao glagol onda ona znači vezati, spariti, spojiti. U matematičkom smislu, reč kopula prvi put se pojavljuje 1959. godine u Sklarovoj (Sklar [20]) teoremi u kojoj se opisuju funkcije koje povezuju jednodimenzione funkcije raspodela sa funkcijom raspodele više promenljivih.

Postavlja se pitanje šta su to kopule i zašto su one važne za ljude koji izučavaju verovatnoću i statistiku. Na samo pitanje šta su kopule ne može se dati jedinstven odgovor. Možemo ih posmatrati na više načina. Na primer, možemo na njih gledati kao funkcije koje spajaju zajedničku funkciju raspodele više promenjivih sa njihovim jednodimenzionim marginalnim funkcijama raspodele. Ili ih možemo smatrati raspodelama više promenljivih, čiji marginali su uniformni na intervalu (0, 1).
Fišer (Fisher [6]) je 1997. odgovorio na pitanje zašto su kopule bitne za statističare i naveo je dva razloga. Prvi razlog je taj što ih možemo koristiti kao način izučavanja mera

zavisnosti, a drugi razlog, odnosno drugi način na koji ih možemo koristiti je kao početnu tačku za konstrukciju familija raspodela dve slučajne promenljive. Specijalno, kopule predstavljaju vrlo važan deo u proučavanju zavisnosti dve slučajne promenljive, jer nam dopuštaju da odvojimo uticaj zavisnosti od uticaja marginalnih raspodela. Ovo svojstvo se može primetiti kod dvodimenzionalne normalne raspodele gde su vektori sredine „otkačeni“ od kovarijacione matrice i zajedno determinisu raspodelu.

Mnogi autori su proučavali kostrukciju raspodele dve slučajne promenljive sa zadatim marginalima, a mi to možemo posmatrati kao konstrukciju kopula.

Prvi istraživački rad u kojem se eksplicitno dovode u vezu kopule i zavisnost slučajnih promenljivih objavili su Švajcer (Schweizer) i Volf (Wolff) 1981. godine [18]. U tom radu Švajcer i Volf diskutovali i modifikovali su Renijev (Renyi [17]) kriterijum za meru zavisnosti dve slučajne promenljive i predstavili su osnovna invarijantna svojstva kopula pri strogo monotonoj transformaciji slučajnih promenljivih. Takođe, uveli su i meru zavisnosti koja danas nosi ime Švajcerovo i Volfovo σ.

Pre njih i mnogi drugi su dovodili u vezu kopule i zavisnost slučajnih promenljivih, pre svega Hefding (Hoeffding [7], [8]), koji je kopule koristio pri izučavanju osnovnih svojstava standardizovanih raspodela. Odnosno, on ih je koristio prilikom proučavanja mera saglasnosti kao što su Spermanovo ρ i njegov indeks zavisnosti Φ^2.

Deheuvels [3] je koristio empirijske funkcije zavisnosti kako bi izračunao kopule nad populacijama, kao i kako bi konstruisao različite neparametarske testove nezavisnosti. O kopulama i njihovoj povezanosti sa zavisnošću slučajnih promenljivih biće više reči u odeljku 3. Kako bismo uopšte došli do toga potrebno je pre svega da uvedemo pojam kopula, da damo osnovne definicije i teoreme, a to ćemo upravo i uraditi u sledećem odeljku. Dakle, tu će biti definisane prvo subkopule pa kopule, biće zadate njihove najvažnije osobine. Zatim doveščemo kopule u vezu sa slučajnim veličinama, naveščemo Sklarovu teoremu koja je jedna od najbitnijih kod kopula, uveščemo i objasniti pojam Freše-Hefdingovih granica, govorićemo o kopulama preživljavanja kao i o višedimenzionim kopulama kod kojih ćemo sve osobine dvodimenzionih kopula proširiti na veću dimenziju. U odeljku 3 će biti reči o zavisnosti slučajnih veličina pomoću kopula, gde ćemo najpre uvesti i razmatrati tzv. mere saglasnosti od kojih su svakako najpoznatije Kendalovo tau i Spirmanovo ro. Pored njih biće opisane još neke mere saglasnosti, a potom će biti reči i o konkretnim merama zavisnosti i njihovim osobinama.

2 Definicije i osnovna svojstva

U ovom odeljku predstavićemo kopule, navesti osnovnu definiciju i osobine kopula. Takođe, navešćemo Sklarovu teoremu koja predstavlja osnovu teorije kopula kao i njihove primene. Biće reči i o Freše-Hefdingovoj donjoj i gornjoj granici za kopule, kao i o pojmu „kopula preživljavanja“ koji će nam kasnije koristiti kada budemo govorili o zavisnosti slučajnih veličina koje imaju raspodele sa repom.

2.1 Dvodimenzione kopule

Pre nego što pređemo na naš konkretan zadatak u ovom radu, moramo uvesti neke oznake, kao i definicije, leme i teoreme koje ćemo kasnije koristiti prilikom definisanja kopula.

Definicija 2.1.1: Neka su data dva neprazna skupa $S_1, S_2 \subseteq \overline{\mathbf{R}}$, ($\overline{\mathbf{R}} = \mathbf{R} \cup \{-\infty, +\infty\}$), i neka je data funkcija $\mathrm{H}: S_1 \times S_2 \to \mathbf{R}$. H-mera skupa $\mathbf{B} = [x_1, x_2] \times [y_1, y_2]$ definisana je sa

$$V_{\mathrm{H}}(\mathbf{B}) = \mathrm{H}(x_2, y_2) - \mathrm{H}(x_2, y_1) - \mathrm{H}(x_1, y_2) + \mathrm{H}(x_1, y_1). \tag{2.1.1}$$

Definicija 2.1.2: Za funkciju H kažemo da je 2-*rastuća* ako je $V_{\mathrm{H}}(\mathbf{B}) \geq 0$ za svaki skup $\mathbf{B} \subset S_1 \times S_2$.

Definicija 2.1.3: Pretpostavimo da postoje $b_1 = \max S_1$ i $b_2 = \max S_2$. Tada marginalne funkcije F i G funkcije H date su na sledeći način:

$$\mathrm{F}: S_1 \to \mathbf{R}, \quad \mathrm{F}(x) = \mathrm{H}(x, b_2) \tag{2.1.2a}$$

$$\mathrm{G}: S_2 \to \mathbf{R}, \quad \mathrm{G}(y) = \mathrm{H}(b_1, y). \tag{2.1.2b}$$

Primetimo da b_1 i b_2 mogu da budu i $+\infty$ jer su to maksimumi skupova koji su podskupovi proširene realne prave.

Definicija 2.1.4: Pretpostavimo da postoje $a_1 = \min S_1$ i $a_2 = \min S_2$. Kažemo da je funkcija H *osnovana* ako

$$\mathrm{H}(a_1, y) = \mathrm{H}(x, a_2) = 0 \tag{2.1.3}$$

za svako $(x, y) \in S_1 \times S_2$.

Primetimo da a_1 i a_2 mogu da budu $-\infty$.

Lema 2.1.5: *Bilo koja osnovana, 2-rastuća funkcija* $\mathrm{H}: S_1 \times S_2 \to \mathbf{R}$ *je neopadajuća po svim argumentima (u ovom slučaju po oba argumenta) ako je za svako* $x_1, x_2 \in S_1$, $x_1 \le x_2$ *i* $y_1, y_2 \in S_2$, $y_1 \le y_2$

$$\mathrm{H}(\cdot, y_2) \ge \mathrm{H}(\cdot, y_1), \tag{2.1.4a}$$

$$\mathrm{H}(x_2, \cdot) \ge \mathrm{H}(x_1, \cdot). \tag{2.1.4b}$$

Dokaz: Kako je funkcija 2-rastuća, onda iz definicije 2.1.2 sledi da je $V_{\mathrm{H}}(\mathbf{B}) \ge 0$ za svaki skup $\mathbf{B} = [x_1, x_2] \times [y_1, y_2] \in S_1 \times S_2$, odnosno važi sledeće

$$\mathrm{H}(x_2, y_2) - \mathrm{H}(x_1, y_2) \ge \mathrm{H}(x_2, y_1) - \mathrm{H}(x_1, y_1) \tag{1}$$

i

$$\mathrm{H}(x_2, y_2) - \mathrm{H}(x_2, y_1) \ge \mathrm{H}(x_1, y_2) - \mathrm{H}(x_1, y_1). \tag{2}$$

Kako je funkcija i osnovana u jednačini (1) zamenimo $x_1 = a_1$ i u jednačini (2) $y_1 = a_2$, odnosno koristimo definiciju 2.1.3. Dobijamo sledeće

$$\mathrm{H}(x_2, y_2) \ge \mathrm{H}(x_2, y_1),\ \mathrm{H}(x_2, y_2) \ge \mathrm{H}(x_1, y_2),$$

a to je upravo ono što smo i želeli da pokažemo, da je funkcija neopadajuća po oba argumenta.

■

Lema 2.1.6: *Za bilo koju osnovanu, 2-rastuću funkciju* $\mathrm{H}: S_1 \times S_2 \to \mathbf{R}$ *i za svaki* $[x_1, x_2] \times [y_1, y_2] \in S_1 \times S_2$ *važi:*

$$|\mathrm{H}(x_2, y_2) - \mathrm{H}(x_1, y_1)| \le |\mathrm{F}(x_2) - \mathrm{F}(x_1)| + |\mathrm{G}(y_2) - \mathrm{G}(y_1)|. \tag{2.1.5}$$

Dokaz: Iz nejednakosti troglova, imamo

$$|\mathrm{H}(x_2, y_2) - \mathrm{H}(x_1, y_1)| \le |\mathrm{H}(x_2, y_2) - \mathrm{H}(x_1, y_2)| + |\mathrm{H}(x_1, y_2) - \mathrm{H}(x_1, y_1)|.$$

Pretpostavimo sada da je $x_1 \le x_2$. Kako je funkcija H osnovana, 2-rastuća i ima marginalne funkcije, prethodna lema i definicija 2.1.3 nam daju

$$0 \le \mathrm{H}(x_2, y_2) - \mathrm{H}(x_1, y_2) \le \mathrm{F}(x_2) - \mathrm{F}(x_1).$$

Analogno, važi nejednakost i kada je $x_2 \le x_1$, a odatle sledi da za svako $x_1, x_2 \in S_1$ važi

$$|\mathrm{H}(x_2, y_2) - \mathrm{H}(x_1, y_2)| \le |\mathrm{F}(x_2) - \mathrm{F}(x_1)|.$$

Na sličan način sledi da za svako $y_1, y_2 \in S_2$ važi

$$\left|\mathrm{H}(x_1,y_2)-\mathrm{H}(x_1,y_1)\right|\le\left|\mathrm{G}(y_2)-\mathrm{G}(y_1)\right|,$$

čime smo kompletirali dokaz ove leme.

■

Sada, kada smo uveli sve što nam je potrebno u stanju smo da definišemo funkcije *kopule*, koje su i glavni predmet izučavanja u ovom radu. Međutim, prvo ćemo definisati subkopule.

Definicija 2.1.7: *Dvodimenziona subkopula* (ili samo *subkopula*) je funkcija C' koja ima sledeća svojstva:

1. Domen funkcije C' je $S_1\times S_2$, gde su S_1 i S_2 podskupovi skupa **I**=[0,1].
2. C' je osnovana i 2-rastuća
3. Za svako $u\in S_1$ i za svako $v\in S_2$ važi

$$\mathrm{C}'(u,1)=u \; i \; \mathrm{C}'(1,v)=v. \tag{2.1.6}$$

Primetimo da za svako (u,v) iz domena funkcije C' važi $0\le\mathrm{C}'(u,v)\le1$, a to znači da je i kodomen funcije C' podskup skupa **I**=[0,1].

Definicija 2.1.8: *Dvodimenziona kopula* (ili jednostavno samo *kupul*a) je dvodimenziona subkopula C čiji je domen $\mathbf{I}^2$.

Ekvivalentno, kopule možemo da definišemo na sledeći način.

Definicija 2.1.9: Kopula je funkcija $\mathrm{C}:\mathbf{I}^2\to\mathbf{I}$, koja ima sledeća svojstva:

1. Za svako $u,v\in\mathbf{I}$

$$\mathrm{C}(u,0)=0=\mathrm{C}(0,v) \tag{2.1.7a}$$

$$\mathrm{C}(u,1)=u,\ \mathrm{C}(1,v)=v; \tag{2.1.7b}$$

2. Za sve $u_1,u_2,v_1,v_2\in\mathbf{I}$, takve da $u_1\le u_2$ i $v_1\le v_2$ važi

$$\mathrm{C}(u_2,v_2)-\mathrm{C}(u_2,v_1)-\mathrm{C}(u_1,v_2)+\mathrm{C}(u_1,v_1)\ge0. \tag{2.1.7c}$$

Razlike između kopula i subkopula su veoma male, ali te razlike će malo kasnije doći do izražaja kada budemo razmatrali Sklarovu teoremu. U svakom slučaju, sva bitnija svojstva kopula su ustvari svojstva subkopula jer je kopula i subkopula, kao što je navedeno gore u definiciji.

Teorema 2.1.10: *Neka je* C' *subkopula. Tada za svaki uređen par* (u,v) *iz domena funkcije* C' *važi*

$$\max(u+v-1,0)\le\mathrm{C}'(u,v)\le\min(u,v). \tag{2.1.8}$$

Dokaz: Neka je (u,v) proizvoljna tačka iz domena subkopule C'.

Tada je $C'(u,v) \le C'(u,1) = u$ i $C'(u,v) \le C'(1,v) = v$. Odatle sledi da je $C'(u,v) \le \min(u,v)$. Kako je subkopula rastuća funkcija važi da je $V_{C'}([u,1] \times [v,1]) \ge 0$. Odatle sledi,

$$C'(1,1) - C'(1,v) - C'(u,1) + C'(u,v) \ge 0,$$
$$1 - v - u + C'(u,v) \ge 0, \text{ tj.}$$
$$C'(u,v) \ge u + v - 1.$$

Iz ovog uslova i uslova $C'(u,v) \ge 0$ sledi da je $C'(u,v) \ge \max(u+v-1,0)$. Odnosno dobili smo da je zaista $\max(u+v-1,0) \le C'(u,v) \le \min(u,v)$.

■

Pošto je svaka kupula i subkopula, samim tim ova nejednakost će da važi i za kopule. Obično se ove granice obeležavaju na sledeći način

$$M(u,v) = \min(u,v)$$
$$W(u,v) = \max(u+v-1,0).$$

Dakle, za kopule važi nejednakost $W(u,v) \le C(u,v) \le M(u,v)$ (2.1.9)

Nejednakost (2.1.9) naziva se Freše-Hefdingova nejednakost, a kopule M i W zvaćemo gornja, odnosno donja Freše-Hefdingova granica. O tim granicama detaljnije ćemo se baviti u odeljku 2.4. Uvešćemo još jednu kopulu koja će nam biti bitna u nastavku izlaganja, a to je kopula proizvoda $\Pi(u,v) = uv$. Pokažimo u sledećem primeru da su zaista $M(u,v)$, $W(u,v)$ i $\Pi(u,v)$ kopule.

Primer 2.1: Dokazaćemo da su $M(u,v)$, $W(u,v)$ i $\Pi(u,v)$ $u,v \in \mathbf{I}$ kopule proveravajući da li one zadovoljavaju osobine navedene u definiciji kopula. Krenimo od gornje Freše-Hefdingove granice.

$M(u,v)$:

1. $M(u,0) = \min(u,0) = 0$, $M(0,v) = \min(0,v) = 0$;
   $M(u,1) = \min(u,1) = u$, $M(1,v) = \min(1,v) = v$.
2. Neka su $u_1 \le u_2$ i $v_1 \le v_2$, $u_1, u_2, v_1, v_2 \in \mathbf{I}$
   $M(u_2,v_2) - M(u_2,v_1) - M(u_1,v_2) + M(u_1,v_1) =$
   $= \min(u_2,v_2) - \min(u_2,v_1) - \min(u_1,v_2) + \min(u_1,v_1) \ge 0$ zbog uslova koji važe za u_1, u_2, v_1, v_2.

$W(u,v)$:

1. $W(u,0) = \max(u-1,0) = 0$, $W(0,v) = \max(v-1,0) = 0$;
   $W(u,1) = \max(u,0) = u$, $W(1,v) = \max(v,0) = v$.
2. Neka su $u_1 \le u_2$ i $v_1 \le v_2$, $u_1, u_2, v_1, v_2 \in \mathbf{I}$
   $W(u_2,v_2) - W(u_2,v_1) - W(u_1,v_2) + M(u_1,v_1) =$

$= \max(u_2+v_2-1,0)-\max(u_2+v_1-1,0)-\max(u_1+v_2-1,0)+\max(u_1+v_1-1,0)\geq 0$
zbog uslova koji važe za u_1, u_2, v_1, v_2.

$\Pi(u,v)$:

1. $\Pi(u,0)=u\cdot 0=0$, $\Pi(0,v)=0\cdot v=0$;
   $\Pi(u,1)=u\cdot 1=u$, $\Pi(1,v)=1\cdot v=v$.
2. Neka su $u_1 \leq u_2$ i $v_1 \leq v_2$, $u_1, u_2, v_1, v_2 \in \mathbf{I}$
   $\Pi(u_2,v_2)-\Pi(u_2,v_1)-\Pi(u_1,v_2)+\Pi(u_1,v_1)=u_2\cdot v_2-u_2\cdot v_1-u_1\cdot v_2+u_1\cdot v_1\geq 0$.

Sledeća teorema sledi direktno iz Leme 2.1.6 i obezbeđuje neprekidnost subkopula, a samim tim i na kopule pomoću Lipšicovog uslova na I^2.

Teorema 2.1.11: *Neka je* C' *subkopula. Tada za svaka dva uređena para* (u_1,u_2), (v_1,v_2) *iz domena subkopule važi*

$$|C'(u_2,v_2)-C'(u_1,v_1)|\leq |u_2-u_1|+|v_2-v_1|. \tag{2.1.10}$$

Dakle, C' je uniformno neprekidna na svom domenu.

■

Teorema 2.1.12: *Neka je data kopula* C. *Za bilo koje* $v\in\mathbf{I}$ *parcijalni izvod* $\partial C(u,v)/\partial u$ *postoji za skoro svako u i za takve v i u važi*

$$0\leq \frac{\partial}{\partial u}C(u,v)\leq 1. \tag{2.1.11a}$$

Slično, za bilo koje u iz **I**, *parcijalni izvod* $\partial C(u,v)/\partial v$ *postoji za skoro svako v i za tako zadate u i v važi*

$$0\leq \frac{\partial}{\partial v}C(u,v)\leq 1. \tag{2.1.11b}$$

Osim toga, funkcije $u\to\partial C(u,v)/\partial v$ *i* $v\to\partial C(u,v)/\partial u$ *su definisane i neopadajuće skoro svuda na* **I**.

Dokaz: Postojanje parcijalnih izvoda $\partial C(u,v)/\partial v$, $\partial C(u,v)/\partial u$ sledi direktno iz činjenice da su monotone funkcije diferencijabilne skoro svuda. Kada u nejednakosti (2.1.10) zamenimo $v_1=v_2$ i $u_1=u_2$ dobijamo da odatle direktno slede nejednakosti (2.1.11a) i (2.1.11b).
Pretpostavimo da je $v_1\leq v_2$ tada koristeći Lemu 2.1.5 imamo da je funkcija $u\to C(u,v_2)-C(u,v_1)$ neopadajuća. Kako je parcijalni izvod $\partial(C(u,v_2)-C(u,v_1))/\partial u$ definisan i neopadajuć skoro svuda na I, sledi da je funkcija $v\to\partial C(u,v)/\partial u$ definisana i neopadajuća skoro svuda na I. Slično, važi i za funkciju $u\to\partial C(u,v)/\partial v$.

■

Teorema 2.1.13: *Neka je data kopula* C. *Ako su* $\partial \mathrm{C}(u,v)/\partial v$ *i* $\partial^2 \mathrm{C}(u,v)/\partial u \partial v$ *neprekidni na* $\mathbf{I}^2$ *i* $\partial \mathrm{C}(u,v)/\partial u$ *postoji za svako* $u \in (0,1)$ *kada je v=0 tada* $\partial \mathrm{C}(u,v)/\partial u$ *i* $\partial^2 \mathrm{C}(u,v)/\partial v \partial u$ *postoje u* $(0,1)^2$ *i važi* $\partial^2 \mathrm{C}(u,v)/\partial u \partial v = \partial^2 \mathrm{C}(u,v)/\partial v \partial u$.

<u>Dokaz</u>*:* Pogledati [21].

■

Za sam kraj ovog odeljka navedimo neke od poznatijih familija kopula i pokazaćemo zašto one predstavljaju bas kopule ispitujući potrebne osobine.

Primer 2.2: Frešeova familija kopula
To je dvoparametarska familija kopula koja sadrži M, W i Π i sledećeg je oblika.
Neka su $\alpha, \beta \in \mathbf{I}$ i $\alpha + \beta \leq 1$

$$\mathrm{C}_{\alpha,\beta}(u,v) = \alpha \mathrm{M}(u,v) + (1-\alpha-\beta)\Pi(u,v) + \beta \mathrm{W}(u,v).$$

1. $\mathrm{C}_{\alpha,\beta}(u,0) = \alpha \mathrm{M}(u,0) + (1-\alpha-\beta)\Pi(u,0) + \beta \mathrm{W}(u,0)$, u Primeru 2.1 pokazali smo da su $\mathrm{M}(u,0) = \mathrm{W}(u,0) = \Pi(u,0) = 0$, pa je tada $\mathrm{C}_{\alpha,\beta}(u,0) = 0$.Na isti način $\mathrm{C}_{\alpha,\beta}(0,v) = 0$.

$$\begin{aligned}\mathrm{C}_{\alpha,\beta}(u,1) &= \alpha \mathrm{M}(u,1) + (1-\alpha-\beta)\Pi(u,1) + \beta \mathrm{W}(u,1)\\ &= \alpha \cdot u + (1-\alpha-\beta)\cdot u + \beta \cdot u\\ &= \alpha \cdot u + u - \alpha \cdot u - \beta \cdot u + \beta \cdot u\\ &= u.\end{aligned}$$

Na isti način je i $\mathrm{C}_{\alpha,\beta}(1,v) = v$

2. $\mathrm{C}_{\alpha,\beta}(u_2,v_2) - \mathrm{C}_{\alpha,\beta}(u_2,v_1) - \mathrm{C}_{\alpha,\beta}(u_1,v_2) + \mathrm{C}_{\alpha,\beta}(u_1,v_1) \geq 0$, jer smo takođe u Primeru 2.1 pokazali da važe iste nejednakosti za M, W i Π, a imamo i uslov da $\alpha + \beta \leq 1$, a to znači da je $1-\alpha-\beta \geq 0$, odnosno ceo izraz će biti veći od nule što smo i hteli da pokažemo.

□

Primer 2.3: Mardia familija kopula.
To je jednoparametarska familija kopula koja takođe sadrži M, W i Π, gde ćemo parametar obeležavati sa θ i on će pripadati intervalu $[-1,1]$, a familija je sledećeg oblika

$$\mathrm{C}_\theta(u,v) = \frac{\theta^2(1+\theta)}{2}\mathrm{M}(u,v) + (1-\theta^2)\Pi(u,v) + \frac{\theta^2(1-\theta)}{2}\mathrm{W}(u,v).$$

1. $\mathrm{C}_\theta(u,0) = \frac{\theta^2(1+\theta)}{2}\mathrm{M}(u,0) + (1-\theta^2)\Pi(u,0) + \frac{\theta^2(1-\theta)}{2}\mathrm{W}(u,0) = 0$ iz istih razloga kao i u prethodnom primeru. Naravno, važi i da je $\mathrm{C}_\theta(0,v) = 0$.

$$\begin{aligned}C_\theta(u,1) &= \frac{\theta^2(1+\theta)}{2}\mathrm{M}(u,1)+(1-\theta^2)\Pi(u,1)+\frac{\theta^2(1-\theta)}{2}\mathrm{W}(u,1)\\ &= \frac{\theta^2(1+\theta)}{2}\cdot u+(1-\theta^2)\cdot u+\frac{\theta^2(1-\theta)}{2}\cdot u\\ &= u.\end{aligned}$$

Slično $C_\theta(1,v)=v$.

2. $C_\theta(u_2,v_2)-C_\theta(u_2,v_1)-C_\theta(u_1,v_2)+C_\theta(u_1,v_1)\geq 0$.

Objašnjenje je kao i u prethodnom primeru jedino treba da obratimo pažnju na vrednost parametra θ, ali kako on pripada intervalu $[-1,1]$ ova nejednakost je u potpunosti opravdana.

□

2.2 Sklarova teorema

Kao što je već rečeno, ova teorema predstavlja osnovu teorije kopula kao i temelj za mnoge, ako ne i sve primene kopula u statistici. Zapravo, Sklarova teorema objašnjava ulogu kopula u vezi između višedimenzionih funkcija raspodele i njihovih marginalnih raspodela. Počećemo ovaj odeljak sa definicijama funkcije raspodele i zajedničke funkcije raspodele, pri čemu nećemo uvoditi verovatnoću u te definicije jer nismo još uvek pominjali slučajne promenljive kao ni njihovu vezu sa kopulama. Bez obzira na sve to funkcije raspodele jedne, dve ili više slučajnih veličina zadovoljavaće uslove te dve definicije koje budemo naveli. Još uvek ćemo govoriti od dvodimenzionim kopulama, a kasnije, u nekom od sledećih odeljaka uvešćemo i višedimenzione kopule i videćemo da će za njih, takođe da važe sve definicije, osobine kao i Sklarova teorema.

Definicija 2.2.1: *Funkcija raspodele* je funkcija F sa domenom $\overline{\mathbf{R}}$ takva da važi

1. F je neopadajuća
2. $F(-\infty)=0$ i $F(+\infty)=1$.

Definicija 2.2.2: *Zajednička funkcija raspodele* je funkcija H sa domenom $\overline{\mathbf{R}}^2$ takva da

1. H je 2-rastuća
2. $H(x,-\infty)=H(-\infty,y)=0$ i $H(\infty,\infty)=1$.

Lema 2.2.3: *Neka je* H *zajednička funkcija raspodele sa marginalnim funkcijama* F *i* G. *Tada postoji jedinstvena subkopula* C' *takva da*

1. $\mathrm{DomC'}=\mathrm{RanF}\times\mathrm{RanG}$,
2. $H(x,y)=C'(F(x),G(y))$ *za svako* $x,y\in\overline{\mathbf{R}}$. (*)

Primedba: sa RanF i RanG označili smo kodomene funkcije F, odnosno G.

<u>Dokaz</u>: Da bi $\mathrm{C'}$ bila jedinstvena svaki uređeni par $(u,v)\in \mathrm{RanF}\times\mathrm{RanG}$ trebalo bi da ima jednu moguću sliku $\mathrm{C'}(u,v)$ konzistentnu sa (*). Pretpostavimo suprotno, da $\mathrm{C'}_1(u,v)\neq \mathrm{C'}_2(u,v)$ su obe konzistentne sa (*), odnosno postoje $(x_1,y_1),(x_2,y_2)\in \overline{\mathbf{R}}^2$ takvi da

$$\mathrm{C'}_1=\mathrm{C'}_1(\mathrm{F}(x_1),\mathrm{G}(y_1))=\mathrm{H}(x_1,y_1),$$
$$\mathrm{C'}_2=\mathrm{C'}_2(\mathrm{F}(x_2),\mathrm{G}(y_2))=\mathrm{H}(x_2,y_2).$$

Prema tome, mora da važi $u=\mathrm{F}(x_1)=\mathrm{F}(x_2)$ i $v=\mathrm{G}(y_1)=\mathrm{G}(y_2)$.
Kako je H zajednička funkcija raspodele, ona mora da zadovoljava uslove leme (2.1.5), odnosno vidimo da će biti

$$|\mathrm{H}(x_2,y_2)-\mathrm{H}(x_1,y_1)|\leq|\mathrm{F}(x_2)-\mathrm{F}(x_1)|+|\mathrm{G}(y_2)-\mathrm{G}(y_1)|=0.$$

Dakle, $\mathrm{C'}_1$ i $\mathrm{C'}_2$ poklapaju se u tački (u,v).
Definišimo sada $\mathrm{C'}$ kao funkciju koja povezuje uređeni par $(\mathrm{F}(x),\mathrm{G}(y))$ sa $\mathrm{H}(x,y)$ kada je $(x,y)\in\overline{\mathbf{R}}^2$. Potrebno je još da pokažemo da je $\mathrm{C'}$ subkopula, odnosno da zadovoljava uslove definicije subkopula.

- $\mathrm{C'}$ je osnovana:

$$\mathrm{C'}(0,\mathrm{G}(y))=\mathrm{C'}(\mathrm{F}(-\infty),\mathrm{G}(y))=\mathrm{H}(-\infty,y)=0$$
$$\mathrm{C'}(\mathrm{F}(x),0)=\mathrm{C'}(\mathrm{F}(x),\mathrm{G}(-\infty))=\mathrm{H}(x,-\infty)=0$$

-

$$\mathrm{C'}(1,\mathrm{G}(y))=\mathrm{C'}(\mathrm{F}(+\infty),\mathrm{G}(y))=\mathrm{H}(+\infty,y)=\mathrm{G}(y)$$
$$\mathrm{C'}(\mathrm{F}(x),1)=\mathrm{C'}(\mathrm{F}(x),\mathrm{G}(+\infty))=\mathrm{H}(x,+\infty)=\mathrm{F}(x)$$

- $\mathrm{C'}$ je 2-rastuća

Neka su $u_1\leq u_2$ iz kodomena funkcije F , a $v_1\leq v_2$ iz kodomena funkcije G. Kako su funkcije raspodela neopadajuće, postoje jedinstveni $x_1\leq x_2$ i $y_1\leq y_2$ takvi da $\mathrm{F}(x_1)=u_1$, $\mathrm{F}(x_2)=u_2$, $\mathrm{G}(y_1)=v_1$, $\mathrm{G}(y_2)=v_2$. Tada imamo

$$\begin{aligned}&\mathrm{C'}(u_2,v_2)-\mathrm{C'}(u_1,v_2)-\mathrm{C'}(u_2,v_1)+\mathrm{C'}(u_1,v_1)=\\&=\mathrm{C'}(\mathrm{F}(x_2),\mathrm{G}(y_2))-\mathrm{C'}(\mathrm{F}(x_1),\mathrm{G}(y_2))-\mathrm{C'}(\mathrm{F}(x_2),\mathrm{G}(y_1))+\mathrm{C'}(\mathrm{F}(x_1),\mathrm{G}(y_1))\\&=\mathrm{H}(x_2,y_2)-\mathrm{H}(x_1,y_2)-\mathrm{H}(x_2,y_1)+\mathrm{H}(x_1,y_1)\geq 0.\end{aligned}$$

■

Lema 2.2.4 (Nelsen 2007)**:** *Neka je* $\mathrm{C'}$ *subkopula. Tada postoji kopula* C *takva da* $\mathrm{C}(u,v)=\mathrm{C'}(u,v)$ *za svako* (u,v) *iz domena funkcije* $\mathrm{C'}$. *Drugim rečima, svaka subkopula može se proširiti do kopule. Ta proširenja nisu jedinstvena.*

Teorema 2.2.5 (Sklarova teorema): *Neka je* H *zajednička funkcija raspodele sa marginalnim funkcijama* F *i* G*. Tada postoji kopula* C *takva da za svako* $x, y \in \overline{\mathbf{R}}$ *važi*

$$\mathrm{H}(x, y) = \mathrm{C}(\mathrm{F}(x), \mathrm{G}(y)) \tag{2.2.1}$$

Ako su F *i* G *neprekidne, tada je* C *jedinstvena, inače,* C *je jedinstveno određena na* $\mathrm{RanF} \times \mathrm{RanG}$*. Obrnuto, ako je* C *kopula, a* F *i* G *funkcije raspodele, tada funkcija H definisana kao* (2.2.1) *je zajednička funkcija raspodele čije su marginalne funkcije raspodele* F *i* G*.*

<u>Dokaz</u>*:* Postojanje kopule C koja zadovoljava (2.2.1) sledi iz Leme 2.2.3 za svako $x, y \in \overline{\mathbf{R}}$. Ako su funkcije F i G neprekidne, to znači da su njihovi kodomeni jednaki jediničnom intervalu **I**. Ako nisu neprekidne tada nam Lema 2.2.4 govori da se subkopula može proširiti do kopule. Odnosno to znači da jedinstvena subkopula iz Leme 2.2.3 je kopula. Obrnuto se proverava direktnim računanjem.

■

Definicija 2.2.6: Neka je F funkcija raspodele. Tada *kvazi-inverzna* funkcija od F je svaka funkcija $\mathrm{F}^{(-1)}$ sa domenom **I** takva da

1. za svako *t* koje pripada kodomenu funkcije F važi $\mathrm{F}(\mathrm{F}^{(-1)}(t)) = t$;
2. ako *t* ne pripada kodomenu funkcije F tada za $x \in \overline{\mathbf{R}}$ važi
$$\mathrm{F}^{(-1)}(t) = \inf\{x \mid \mathrm{F}(x) \geq t\} = \sup\{x \mid \mathrm{F}(x) \leq t\}.$$

Posledica: *Neka su* H, F, G *i* C' *definisane kao u Lemi 2.2.3 i neka su* $\mathrm{F}^{(-1)}$ *i* $G^{(-1)}$ *kvazi-inverzne funkcije funkcija* F *i* G*. Tada za bilo koje* (*u*,*v*) *iz domena od* C' *važi*

$$\mathrm{C}'(u, v) = \mathrm{H}(\mathrm{F}^{(-1)}(x), \mathrm{G}^{(-1)}(y))..$$

Primer 2.4: Neka je data funkcija raspodele H na sledeći način

$$\mathrm{H}(x, y) = \begin{cases} \dfrac{(x+1)(e^y - 1)}{x + 2e^y - 1}, & (x,y) \in [-1,1] \times [0,\infty], \\ 1 - e^{-y}, & (x,y) \in (1,\infty] \times [0,\infty], \\ 0, & inace. \end{cases},$$

čije su marginalne funkcije raspodele date na sledeći način:

$$\mathrm{F}(x) = \begin{cases} 0, & x < -1, \\ (x+1)/2, & x \in [-1,1], \\ 1, & x > 1. \end{cases} \quad \text{i} \quad \mathrm{G}(y) = \begin{cases} 0, & y < 0, \\ 1 - e^{-y}, & y \geq 0. \end{cases}.$$

Može se pokazati da će za bilo koje vrednosti za *u* i *v* iz jediničnog intervala, kvazi inverzne funkcije funkcija F i G biti sledećeg oblika

$$\mathrm{F}^{(-1)}(u) = 2u - 1,$$

odnosno $\mathrm{G}^{(-1)}(v) = -\ln(1-v)$.

Kodomeni ove dve funkcije F i G su jednaki baš jediničnom intervalu pa nam odatle sledi da će kopula C biti oblika

$$\begin{aligned}
\mathrm{C}(u,v) &= \mathrm{H}(\mathrm{F}^{(-1)}(u), \mathrm{G}^{(-1)}(v)) \\
&= \frac{(2u-1+1)(e^{-\ln(1-v)}-1)}{2u-1+2e^{-\ln(1-v)}-1} \\
&= \frac{2u\left(\frac{1}{1-v}-1\right)}{2\left(u-1+\frac{1}{1-v}\right)} \\
&= \frac{\frac{uv}{1-v}}{\frac{u-uv+v}{1-v}}
\end{aligned}.$$

Odnosno

$$\mathrm{C}(u,v) = \frac{uv}{u+v-uv}.$$

□

Primer 2.5: Neka nam je data Gumbelova dvodimenzionalna eksponencijalna funkcija raspodele koja je sledećeg oblika

$$\mathrm{H}_\theta(x,y) = \begin{cases} 1-e^{-x}-e^{-y}+e^{-(x+y+\theta xy)}, & x \geq 0, y \geq 0, \\ 0, & \text{u ostalom.} \end{cases}.$$

Parametar θ pripada jediničnom intervalu. Primetimo da za tako zadat parametar marginalne funkcije raspodele će biti ustvari eksponencijalne raspodele, a njihove kvazi inverzne funkcije će biti oblika $\mathrm{F}^{(-1)}(u) = -\ln(1-u)$ i $\mathrm{G}^{(-1)}(v) = -\ln(1-v)$ za svako u i v iz **I**. Tada je odgovarajuća kopula

$$\begin{aligned}
\mathrm{C}_\theta(u,v) &= \mathrm{H}_\theta(\mathrm{F}^{(-1)}(u), \mathrm{G}^{(-1)}(v)) \\
&= 1-e^{\ln(1-u)}-e^{\ln(1-v)}+e^{-(-\ln(1-u)-\ln(1-v)+\theta\ln(1-u)\ln(1-v))} \\
&= 1-1+u-1+v+(1-u)(1-v)e^{-\theta\ln(1-u)\ln(1-v)} \\
&= u+v-1+(1-u)(1-v)e^{-\theta\ln(1-u)\ln(1-v)}.
\end{aligned}$$

□

Sa odgovarajućim proširenjem svog domena na $\overline{\mathbf{R}}^2$, svaka kopula predstavlja zajedničku funkciju raspodele sa marginalnim funkcijama raspodele koje su uniformne na jediničnom intervalu. Preciznije, neka je C kopula tada definišemo funkciju H_C kao

$$
\mathrm{H_C} = \begin{cases} 0, & x<0 \text{ ili } y<0, \\ \mathrm{C}(x,y), & (x,y)\in \mathbf{I}^2, \\ x, & y>1, x\in\mathbf{I}, \\ y, & x>1, y\in\mathbf{I}, \\ 1, & x>1, y>1. \end{cases}
$$

2.3 Kopule i slučajne promenljive

U ovom odeljku, konačno ćemo uvesti i slučajne promenljive i opisati povezanost sa kopulama. Po konvenciji slučajne veličine ćemo pisati velikim slovima npr. *X*, a funkciju raspodele te slučajne promenljive ćemo obeležavati $\mathrm{F}(x)=\mathrm{P}\{X\le x\}$, gde je $x\in\overline{\mathbf{R}}$ vrednost koju uzima ta promenljiva. Rekli smo da je Sklarova teorema osnova teorije kopula, pa kako smo uveli verovatnoće u celu priču, moći ćemo da Sklarovu teoremu navedemo u terminima slučajnih veličina.

Teorema 2.3.1: *Neka su X i Y slučajne veličine sa funkcijama raspodele redom* F *i* G *i zajedničkom funkcijom raspodele* H. *Tada postoji kopula* C *koja zadovoljava* (2.2.1). *Ako su* F *i* G *neprekidne funkcije, kopula* C *je jedinstvena. U suprotnom jedinstveno je definisana na* $\mathrm{RanF}\times\mathrm{RanG}$.

Kopulu C iz ove teoreme zvaćemo kopulom od *X* i *Y*, i obeležavaćemo je sa C_{XY}. Ranije u radu, tačnije u odeljku 2.1 pomenuli smo kopulu proizvoda $\Pi(u,v)=uv$ i rekli smo da će nam i ona biti važna. Nju ćemo koristiti u sledećoj teoremi, koja govori o nezavisnosti slučajnih veličina kada su njihove funkcije raspodele neprekidne.

Teorema 2.3.2: *Neka su X i Y neprekidne slučajne veličine. Tada su X i Y nezavisne ako i samo ako* $\mathrm{C}_{XY}=\Pi(u,v)$.

Dokaz ove teoreme sledi direktno iz Teoreme 2.3.1 i činjenice da su *X* i *Y* nezavisne ako i samo ako $\mathrm{H}(x,y)=\mathrm{F}(x)\mathrm{G}(y)$ za svako $(x,y)\in\overline{\mathbf{R}}^2$.

Kopule imaju jedno veoma lepo svojstvo, a to je da pri strogo monotonim transformacijama slučajnih veličina kopule su ili invarijantne ili se menjaju na prost, predvidiv način.

Podsetimo se da ako je funkcija raspodele slučajne veličine *X* neprekidna i ako je funkcija α strogo monotona čiji domen sadrži kodomen od *X*, tada je funkcija raspodele slučajne veličine $\alpha(X)$ takođe neprekidna.

Teorema 2.3.3: *Neka su X i Y neprkidne slučajne veličine sa kopulom* C_{XY}. *Ako su* α *i* β *strogo rasuće na kodomenima od X i Y redom, tada* $\mathrm{C}_{\alpha(X)\beta(Y)} = \mathrm{C}_{XY}$. *Prema tome,* C_{XY} *je invarijantna pri stogo rastućoj transformaciji X i Y.*

Dokaz: Obeležimo sa $\mathrm{F}_1, \mathrm{G}_1, \mathrm{F}_2$ i G_2 funkcije raspodele slučajnih veličina redom *X*, *Y*, $\alpha(X)$ i $\beta(Y)$. Kako su α i β strogo rastuće važi da je

$$\mathrm{F}_2(x) = \mathrm{P}\{\alpha(X) \le x\} = \mathrm{P}\{X \le \alpha^{-1}(x)\} = \mathrm{F}_1(\alpha^{-1}(x)),$$

i na isti način važi $\mathrm{G}_2(y) = \mathrm{G}_1(\beta^{-1}(y))$.
Odavde sledi da za svako $x, y \in \overline{\mathbf{R}}$

$$\begin{aligned} \mathrm{C}_{\alpha(X)\beta(Y)}(\mathrm{F}_2(x), \mathrm{G}_2(y)) &= \mathrm{P}\{\alpha(X) \le x, \beta(Y) \le y\} \\ &= \mathrm{P}\{X \le \alpha^{-1}(x), Y \le \beta^{-1}(y)\} \\ &= \mathrm{C}_{XY}(\mathrm{F}_1(\alpha^{-1}(x)), \mathrm{G}_1(\beta^{-1}(y))) \\ &= \mathrm{C}_{XY}(\mathrm{F}_2(x), \mathrm{G}_2(y)) \end{aligned}$$

X i *Y* su neprekidne i $\mathrm{Ran}\,\mathrm{F}_2 = \mathrm{Ran}\,\mathrm{G}_2 = \mathbf{I}$, odakle sledi da je $\mathrm{C}_{\alpha(X)\beta(Y)} = \mathrm{C}_{XY}$ na $\mathbf{I}^2$.

■

U slučaju da je najmanje jedna od funkcija α i β strogo opadajuća, dobijamo rezultat u kojem je kopula od $\alpha(X)$ i $\beta(Y)$ jednostavna transformacija kopule C_{XY}, odnosno imamo sledeću teoremu.

Teorema 2.3.4: *Neka su X i Y neprekidne slučajne veličine sa kopulom* C_{XY}. *Neka su* α *i* β *strogo monotone redom na* $\mathrm{Ran}\,X$ *i* $\mathrm{Ran}\,Y$.

1. *Ako je* α *strogo rastuća, a* β *strogo opadajuća, tada*
$$\mathrm{C}_{\alpha(X)\beta(Y)}(u,v) = u - \mathrm{C}_{XY}(u, 1-v).$$

2. *Ako je* α *strogo opadajuća, a* β *strogo rastuća, tada*
$$\mathrm{C}_{\alpha(X)\beta(Y)}(u,v) = v - \mathrm{C}_{XY}(1-u, v).$$

3. *Ako su obe* α *i* β *strogo opadajuće, tada*
$$\mathrm{C}_{\alpha(X)\beta(Y)}(u,v) = u + v - 1 - \mathrm{C}_{XY}(1-u, 1-v).$$

Dokaz: Obeležimo sa $\mathrm{F}_1, \mathrm{G}_1, \mathrm{F}_2$ i G_2 funkcije raspodele slučajnih veličina redom *X*, *Y*, $\alpha(X)$ i $\beta(Y)$.

1. α strogo rastuća, a β strogo opadajuća

$$
\begin{aligned}
C_{\alpha(X)\beta(Y)}(F_2(x), G_2(y)) &= P\{\alpha(X) \le x, \beta(Y) \le y\} \\
&= P\{\alpha(X) \le x, Y > \beta^{-1}(y)\} \\
&= P\{\alpha(X) \le x\} - P\{\alpha(X) \le x, Y \le \beta^{-1}(y)\} \\
&= F_2(x) - C_{\alpha(X),Y}(F_2(x), G_1(\beta^{-1}(y))) \\
&= F_2(x) - C_{\alpha(X),Y}(F_2(x), 1 - G_2(y))\ ^{*}
\end{aligned}
$$

Dalje koristeći prethodnu teoremu imamo da je

$$C_{\alpha(X),Y}(F_2(x), 1 - G_2(y)) = C_{XY}((F_2(x), 1 - G_2(y))).$$

I sada to vratimo u prethodnu jednakost i dobijamo

$$^{*} = F_2(x) - C_{X,Y}(F_2(x), 1 - G_2(y)).$$

Na isti način dokazuju se stavke 2. i 3.

■

Svaka zajednička funkcija raspodele H indukuje verovatnosnu meru na $\mathbf{R}^2$ preko $V_H((-\infty, x] \times (-\infty, y]) = H(x, y)$ i standardno proširenje na Borelove podskupove od $\mathbf{R}^2$ koristeći tehnike teorije mere. Kako su kopule zajedničke funkcije raspodele sa marginalnim funkcijama uniformnim na (0,1), to znači da one indukuju verovatnosnu meru na I^2 preko $V_C((0,u] \times (0,v]) = C(u,v)$, zvaćemo je C-*mera*. Intuitivno, smatraćemo da je C- mera nekog skupa koji je podskup od $\mathbf{I}^2$ verovatnoća da dve uniformno raspodeljene slučajne veličine na (0,1) sa zajedničkom funkcijom raspodele C uzimaju vrednosti iz tog podskupa. C-mere se često nazivaju dvostruke-stohastičke mere, jer za bilo koji merljiv podskup S od **I** važi $V_C(S\times\mathbf{I})=V_C(\mathbf{I}\times S)=\lambda(S)$, gde je λ Lebegova mera na **I**.

Navedimo sada još neke pojmove vezane za slučajne veličine i njihovu povezanost sa kopulama.Neka je za bilo koju kopulu C

$$C(u,v) = A_C(u,v) + S_C(u,v), \text{ gde je}$$

$$A_C(u,v) = \int_0^u \int_0^v \frac{\partial^2}{\partial s \partial t} C(s,t)\,dtds, \text{ a } S_C(u,v) = C(u,v) - A_C(u,v).$$

Ako je $C=A_C$ na $\mathbf{I}^2$ i kao takvu je posmatramo kao zajedničku funkciju raspodele, tada C ima gustinu datu sa $\frac{\partial^2 C(u,v)}{\partial u \partial v}$ i C je onda apsolutno neprekidna.

Primer 2.6: Kopula proizvoda $\Pi(u,v) = uv$ je apsolutno neprekidna jer je

$$A_\Pi(u,v) = \int_0^u \int_0^v \frac{\partial^2}{\partial s \partial t} \Pi(s,t)\,dtds = \int_0^u \int_0^v 1\,dtds = uv = \Pi(s,t).$$

□

Ako je C=S_C na $\mathbf{I}^2$, tada je $\dfrac{\partial^2 C(u,v)}{\partial u \partial v}=0$ skoro svuda na $\mathbf{I}^2$ i kažemo da je C singularna. U suprotnom, C ima apsolutno neprekidnu komponentu A_C i singularnu komponentu S_C. U tom slučaju ni A_C ni S_C nisu kopule, zato što nijedna od te dve komponente nema marginale uniformne na (0,1).

Primer 2.7: Neka su *X* i *Y* slučajne veličine takve da je njihova zajednička funkcija raspodele

$$\mathrm{H}(x)=(1+e^{-x}+e^{-y})^{-1} \text{ za svako } x,y\in\overline{\mathbf{R}}.$$

Marginalne funkcije raspodele će tada biti sledećeg oblika

$$\mathrm{F}(x)=(1+e^{-x})^{-1} \quad \text{i} \quad \mathrm{G}(x)=(1+e^{-y})^{-1}.$$

Možemo primetiti da slučajne veličine ustvari imaju logističke funkcije raspodele.
Nađimo sada kopulu od *X* i *Y*, a to ćemo uraditi kao i u pretodnim primerima inverznom metodom.
Kvazi inverzne funkcije raspodele slučajnih veličina *X* i *Y* su oblika

$$\mathrm{F}^{(-1)}(x)=-\ln\left(\frac{1}{x}-1\right) \quad \text{i} \quad \mathrm{G}^{(-1)}=-\ln\left(\frac{1}{y}-1\right).$$

Odnosno kopula od dve slučajne veličine sa logističkim raspodelama je

$$\begin{aligned}
\mathrm{C}_{X,Y}(u,v)&=\mathrm{H}\left(\mathrm{F}^{(-1)}(u),\mathrm{G}^{(-1)}(v)\right)\\
&=\left(1+e^{\ln\left(\frac{1}{u}-1\right)}+e^{\ln\left(\frac{1}{v}-1\right)}\right)^{-1}\\
&=\left(1+\frac{1}{u}-1+\frac{1}{v}-1\right)^{-1}\\
&=\frac{uv}{u+v-uv}.
\end{aligned}$$

□

2.4 Freše-Hefdingove granice za zajedničke funkcije raspodele

U jednom od prethodnih odeljaka pomenuli smo i Freše-Hefdingove granice koje ćemo ovde opisati. Zašto su nam one bitne uopšte? Pa pre svega ove dve granice su i same kopule. Sve druge kopule uzimaju vrednosti koje se nalaze između ovih granica u svakoj tački jediničnog kvadrata. Međutim, postoji još jedan razlog zbog kojeg će nam ove granice biti bitne, a to je zato što gornja Freše-Hefdingova granica odgovara pozitivnoj savršenoj zavisnosti, a donja granica odgovara negativnoj savršenoj zavisnosti, a o tome će biti više reči kada se budemo konkretno bavili zavisnošću slučajnih veličina. Podsetimo se kako je izgledala nejednakost gde smo prvi put uveli pomenute granice:

$$\mathrm{W}(u,v) = \max(u+v-1,0) \le \mathrm{C}(u,v) \le \mathrm{M}(u,v) = \min(u,v).$$

Kao posledicu Sklarove teoreme u slučaju kada su X i Y slučajne veličine sa zajedničkom funkcijom raspodele H sa marginalnim funkcijama F i G, tada za svako $x, y \in \overline{\mathbf{R}}$ imamo sledeću nejednakost:

$$\max(\mathrm{F}(x)+\mathrm{G}(y)-1,0) \le \mathrm{H}(x,y) \le \min(\mathrm{F}(x),\mathrm{G}(y)). \tag{2.4.1}$$

Kako su M i W kopule, granice u nejednakosti (2.4.1) su zajedničke funkcije raspodele i nazivaju se Freše-Hefdingove granice za zajedničku funkciju raspodele H sa marginalnim funkcijama F i G. U nastavku bavićemo se i pitanjem šta se dešava kad je funkcija H jednaka nekoj od ovih granica. Ali, pre toga uvedimo pojmove neopadajućih i nerastučih skupova na $\overline{\mathbf{R}}^2$.

Definicija 2.4.1: Neka je $S = S_1 \times S_2 \subseteq \overline{\mathbf{R}}^2$. Tada ako za sve $(x,y),(u,v) \in S$

a) važi da ako je $x < u \Rightarrow y \le v$, skup S je neopadajući
b) važi da ako je $x < u \Rightarrow y \ge v$, skup S je nerastući.

Navešćemo sada definiciju nosača funkcije, jer ćemo taj pojam koristiti u jednoj od sledećih teorema gde ćemo dokazati da je zajednička funkcija raspodele H para slučajnih promenljivih (X,Y) upravo gornja Freše-Hefdingova granica.

Definicija 2.4.2: Nosač funkcije raspodele H je komplement unije svih otvorenih podskupova od $\mathbf{R}^2$ koji su H-mere nula.

Lema 2.4.3: *Neka je dat skup* $S \subseteq \overline{\mathbf{R}}^2$. *Tada, S je neopadajući ako i samo ako za svaki par* $(x,y) \in \overline{\mathbf{R}}^2$

a) za svaki par $(u,v) \in \overline{\mathbf{R}}^2$, $u \le x \Rightarrow v \le y$; ili (2.4.3a)
b) za svaki par $(u,v) \in \overline{\mathbf{R}}^2$, $v \le y \Rightarrow u \le x$. (2.4.3b)

<u>*Dokaz*</u>*:* Pretpostavimo da je S neopadajući i da nijedna od relacija (2.4.3a), (2.4.3b) ne važi. Tada postoje tačke (a,b) i (c,d) u S takve da $a \le x, b > y, d \le y, c > x$. Odatle, $a < c$ i $b > d$, a to je kontradikcija.
U obrnutom smeru pretpostavimo da S nije neopadajući. Tada postoje tačke (a,b) i (c,d) u S takve da $a < c$ i $b > d$. Za $(x,y) = (\frac{a+c}{2}, \frac{b+d}{2})$ nijedna od relacija (2.4.3a), (2.4.3b) ne važi. Odnosno pokazali smo da je S neopadajući.

■

Lema 2.4.4: *Neka su X i Y slučajne veličine sa zajedničkom funkcijom raspodele* H. *Tada* H *je jednaka Freše-Hefdingovoj gornjoj granici ako i samo ako za svaki par* $(x, y) \in \overline{\mathbf{R}}^2$ *važi* $P\{X > x, Y \le y\} = 0$ *ili* $P\{X \le x, Y > y\} = 0$.

Dokaz: Neka F i G predstavljaju marginale funkcije H. Tada

$$\begin{aligned} \mathrm{F}(x) = \mathrm{P}\{X \le x\} &= \mathrm{P}\{X \le x, Y \le y\} + \mathrm{P}\{X \le x, Y > y\} \\ &= \mathrm{H}(x, y) + \mathrm{P}\{X \le x, Y > y\}, \text{ i} \end{aligned}$$

$$\begin{aligned} \mathrm{G}(y) = \mathrm{P}\{Y \le y\} &= \mathrm{P}\{X \le x, Y \le y\} + \mathrm{P}\{X > x, Y \le y\} \\ &= \mathrm{H}(x, y) + \mathrm{P}\{X > x, Y \le y\}. \end{aligned}$$

$$\mathrm{H}(x, y) = \mathrm{M}(\mathrm{F}(x), \mathrm{G}(y)) \Leftrightarrow \min(\mathrm{P}\{X \le x, Y > y\}, \mathrm{P}\{X > x, Y \le y\}) = 0$$

■

Teorema 2.4.5: *Neka su X i Y slučajne veličine sa zajedničkom funkcijom raspodele* H. *Tada* H *je identički jednaka svojoj gornjoj Freše-Hefdingovoj granici ako i samo ako je nosač funkcije* H *neopadajući podskup od* $\overline{\mathbf{R}}^2$.

Dokaz: Označimo sa *S* nosač funkcije *H* i neka je (*x*,*y*) proizvoljna tačka iz $\overline{\mathbf{R}}^2$. Tada (2.4.3a) važi ako i samo ako je $\{(u,v) \mid u \le x, v > y\} \cap S = \varnothing$, odnosno ekvivalentno, ako i samo ako je $\mathrm{P}\{X \le x, Y > y\} = 0$.
Na sličan način, pokazuje sa da važi (2.4.3b) ako i samo ako je $\mathrm{P}\{X > x, Y \le y\} = 0$.
Primenama Leme 2.4.3 i Leme 2.4.4 kompletirali smo dokaz ove teoreme.

■

Naravno, analagni rezultat postoji i za donju Freše-Hefdingovu granicu (postoje analogoni Leme 2.4.3 i Leme 2.4.4 za nerastući skup odnosno donju granicu), a sledeća teorema koju navodimo, opisuje upravo to.

Teorema 2.4.6: *Neka su X i Y slučajne veličine sa zajedničkom funkcijom raspodele* H. *Tada* H *je identički jednaka svojoj donjoj Freše-Hefdingovoj granici ako i samo ako je nosač funkcije* H *nerastući podskup od* $\overline{\mathbf{R}}^2$.

■

2.5 Kopula preživljavanja

Kod izučavanja slučajnih veličina od posebnog interesa predstvalja nam i životni vek slučajne veličine ili nekih objekata iz zadate populacije. Verovatnoću da neka slučajna veličina premaši određenu vrednost data je sa $\overline{\mathrm{F}}(x) = \mathrm{P}\{X > x\} = 1 - \mathrm{F}(x)$. Tu verovatnoću nazivamo *funkcija preživljavanja*, a funkcija F je funkcija raspodele slučajne veličine *X*. Funkciju preživljavanja dve slučajne veličine sa zajedničkom funkcijom raspodele definišemo na sledeći način.

Definicija 2.5.1: Za par slučajnih veličina (*X*,*Y*) sa zajedničkom funkcijom raspodele H, *zajednička funkcija preživljavanja* data je sa $\overline{\mathrm{H}}(x, y) = \mathrm{P}\{X > x, Y > y\}$. Marginalne funkcije funkcije $\overline{\mathrm{H}}$ su $\overline{\mathrm{H}}(x, -\infty) = \overline{\mathrm{F}}(x)$ i $\overline{\mathrm{H}}(-\infty, y) = \overline{\mathrm{G}}(x)$.

Neka nam je data kopula C od *X* i *Y*. Tada imamo

$$\begin{aligned}\overline{\mathrm{H}}(x, y) &= 1 - \mathrm{F}(x) - \mathrm{G}(y) + \mathrm{H}(x, y) \\ &= \overline{\mathrm{F}}(x) + \overline{\mathrm{G}}(y) - 1 + \mathrm{C}(\mathrm{F}(x), \mathrm{G}(y)) \\ &= \overline{\mathrm{F}}(x) + \overline{\mathrm{G}}(y) - 1 + \mathrm{C}(1 - \overline{\mathrm{F}}(x), 1 - \overline{\mathrm{G}}(y)).\end{aligned}$$

Ako definišemo sada funkciju

$$\hat{\mathrm{C}}(u, v) = u + v - 1 + \mathrm{C}(1 - u, 1 - v) \ ,$$

koja slika $\mathbf{I}^2$ u **I**, imaćemo sledeće

$$\overline{\mathrm{H}}(x, y) = \hat{\mathrm{C}}(\overline{\mathrm{F}}(x), \overline{\mathrm{G}}(y)).$$

Lako se pokazuje da je $\hat{C}$ kopula, potrebno je samo proveriti one osobine koju su navedene u definiciji kopula, što ćemo sada i uraditi.

- $\hat{\mathrm{C}}(o, v) = o + v - 1 + \mathrm{C}(1, 1 - v) = v - 1 + 1 - v = 0$
- $\hat{\mathrm{C}}(1, v) = 1 + v - 1 + \mathrm{C}(0, 1 - v) = v + 0 = v$
- $\hat{\mathrm{C}}(u_2, v_2) - \hat{\mathrm{C}}(u_2, v_1) - \hat{\mathrm{C}}(u_1, v_2) + \hat{\mathrm{C}}(u_1, v_1) =$

$$\begin{aligned}&= u_2 + v_2 - 1 + \mathrm{C}(1 - u_2, 1 - v_2) \\ &- u_2 - v_1 + 1 - \mathrm{C}(1 - u_2, 1 - v_1) \\ &- u_1 - v_2 + 1 - \mathrm{C}(1 - u_1, 1 - v_2) \\ &+ u_1 + v_1 - 1 + \mathrm{C}(1 - u_1, 1 - v_1) \\ &= \mathrm{C}(1 - u_2, 1 - v_2) - \mathrm{C}(1 - u_2, 1 - v_1) - \mathrm{C}(1 - u_1, 1 - v_2) + \mathrm{C}(1 - u_1, 1 - v_1) \geq 0 ,\end{aligned}$$

Definicija 2.5.2: Kopula $\hat{C}$ povezana sa kopulom C je

$$\hat{C}(u,v) = u + v - 1 + C(1-u, 1-v), \tag{2.5.1}$$
i zvaćemo je *kopulom preživljavanja* slučajnih veličina X i Y.

Potrebno je da izbegnemo mešanje kopule preživljavanja $\hat{C}$ sa funkcijom preživljavanja $\overline{C}$ dve uniformno (0,1) raspodeljene slučajne veličine čija je zajednička funkcija raspodele kopula C, jer:

$$\overline{C}(u,v) = P\{U > u, V > v\} = 1 - u - v + C(u,v) = \hat{C}(1-u, 1-v). \tag{2.5.2}$$

Primer 2.8: Neka su X i Y slučajne veličine takve da obe imaju Paretovu raspodelu. Njihova zajednička funkcija preživljavanja je sledeća

$$\overline{H}_\theta(x,y) = \begin{cases} (1+x+y)^{-\theta}, & x \geq 0, y \geq 0, \\ (1+x)^{-\theta}, & x \geq 0, y < 0, \\ (1+y)^{-\theta}, & x < 0, y \geq 0, \\ 1, & x < 0, y < 0. \end{cases}, \quad \theta > 0.$$

To znači da će njihove marginalne funkcije preživljavanja biti oblika redom

$$\overline{F} = \begin{cases} (1+x)^{-\theta}, & x \geq 0 \\ 1, & x < 0. \end{cases} \quad \text{i} \quad \overline{G} = \begin{cases} (1+y)^{-\theta}, & y \geq 0 \\ 1, & y < 0. \end{cases}$$

Sada ćemo invertovati marginalne funkcije preživljavanja kako bismo našli kopulu preživljavanja. Kada to uradimo dobijamo sledeće rezultate

$$\overline{F}^{(-1)}(u) = u^{-1/\theta} - 1 \quad \text{i} \quad \overline{G}^{(-1)}(v) = v^{-1/\theta} - 1.$$

Sada možemo da nađemo kopulu preživljavanja za slučajne veličine X i Y koje imaju Paretovu raspodelu i ona izgleda ovako:

$$\begin{aligned} \hat{C}(u,v) &= \overline{H}\left(\overline{F}^{(-1)}, \overline{G}^{(-1)}\right) \\ &= \left(1 + u^{-1/\theta} - 1 + v^{-1/\theta} - 1\right)^{-\theta} \\ &= \left(u^{-1/\theta} + v^{-1/\theta} - 1\right)^{-\theta}. \end{aligned}$$

□

Primer 2.9: Neka su X i Y slučajne veličine sa standardnom eksponencijalnom raspodelom, tj. funkcije raspodele su sledećeg oblika

$$F(x) = 1 - e^{-x} \text{ i } G(x) = 1 - e^{-y}, \quad x, y \geq 0.$$

Njihova zajednička funkcija preživljavanja je

$$\overline{H}(x,y) = (e^x + e^y - 1)^{-1}, \quad x, y \geq 0,$$

dok će marginalne funkcije preživljavanja će biti $\overline{\mathrm{F}}(x)=e^{-x}$ i $\overline{\mathrm{G}}(x)=e^{-y}$.

Nađimo kopulu preživljavanja ove dve slučajne veličine. Kvazi inverzne funkcije su sledeće $\overline{\mathrm{F}}^{(-1)}(u)=\ln\left(\frac{1}{u}\right)$ i $\overline{\mathrm{G}}^{(-1)}(v)=\ln\left(\frac{1}{v}\right)$.

$$\begin{aligned}\hat{\mathrm{C}}(u,v)&=\overline{\mathrm{H}}\left(\overline{\mathrm{F}}^{(-1)}(u),\overline{\mathrm{G}}^{(-1)}(v)\right)\\&=\left(e^{\ln(1/u)}+e^{\ln(1/v)}-1\right)^{-1}\\&=\left(\frac{1}{u}+\frac{1}{v}-1\right)^{-1}\\&=\frac{uv}{u+v-uv}.\end{aligned}$$

□

2.6 Višedimenzione kopule

U ovom odeljku generalizovaćemo prethodne rezultate na više dimenzije tj. $n\geq 2$. Iako je većina definicija i teorema analogna i u višedimezionom slučaju, imaćemo par izuzetaka na koje treba posebno da obratimo pažnju. Pokazaćemo na primer da donja Freše-Hefdingova granica nije kopula u slučaju kada je $n\geq 3$. Kako ne bi bilo zabune prilikom navođenja teorema mi ćemo predefinisati sve pojmove u višedimenzionom slučaju. Uvedimo neke oznake koje ćemo pominjati često u našoj priči o višedimenionim kopulama, kako bismo izbegli glomaznost u pisanju.

Neka je $\overline{\mathbf{R}}^{\mathbf{n}}=\overline{\mathbf{R}}\times\overline{\mathbf{R}}\times\mathbf{...}\overline{\mathbf{R}}$ prošireni n-dimenzioni vektorski prostor, a vektori iz ovog prostora biće $\mathbf{a}=(a_1,a_2,...,a_n)$.

Za $\mathbf{a}\leq\mathbf{b}$, $\mathbf{B}=[\mathbf{a},\mathbf{b}]=[a_1,b_1]\times[a_2,b_2]\times...[a_n,b_n]$.

Ivice n-dimenzionog segmenta **B** su vektori $\mathbf{c}=(c_1,c_2,...,c_n)$, gde je svako od c_k jednako ili a_k ili b_k. Jedinčni n-dimezioni segment je $\mathbf{I}^{\mathbf{n}}=\mathbf{I}\times\mathbf{I}\times\mathbf{...}\times\mathbf{I}$.

Definicija 2.6.1: Neka su S_1, S_2,..., S_n neprazni podskupovi od $\overline{\mathbf{R}}$ i neka je $\mathrm{H}:S_1\times S_2\times\cdots\times S_n\to\mathbf{R}$ funkcija n promenljivih. Neka je **B** n-dimenzioni segment čije se ivice nalaze unutar domena funkcije H. Tada H-meru od B definišemo kao

$$\mathrm{V_H}(\mathbf{B}) = \sum_{\mathbf{c}\in\mathbf{B}} \mathrm{sgn}(\mathbf{c})\,\mathrm{H}(\mathbf{c}), \tag{2.6.1}$$

gde je

$$\mathrm{sgn}(\mathbf{c}) = \begin{cases} 1, & c_k = a_k \ \text{za parno k}, \\ -1, & c_k = a_k \ \text{za neparno k}. \end{cases}$$

Definicija 2.6.2: Funkcija n promenljivih H je n-*rastuća* ako $\mathrm{V_H}(\mathbf{B}) \geq 0$ za svako **B** čije ivice leže u domenu funkcije H.

Definicija 2.6.3: Neka je $\mathrm{H}: S_1 \times S_2 \times \cdots \times S_n \to \mathbf{R}$, gde svaki od skupova S_i ima minimum a_i. Kažemo da je funkcija H *osnovana* ako za svako **u** iz domena H takvo da za bar jedan indeks k je $u_k=a_k$, važi

$$\mathrm{H}(\mathbf{u}) = \mathrm{H}(u_1,...,u_{k-1},a_k,u_{k+1},...,u_n) = 0.$$

Ako je svaki skup S_k neprazan i ima najveći elemnt b_k tada kažemo da funkcija H ima marginalne funkcije. Jednodimenzionalnu marginalnu funkciju obeležavaćemo sa H_k čiji će domen biti S_k i za svako $x \in S_k$ važi

$$\mathrm{H}_k(x) = \mathrm{H}(b_1,...,b_{k-1},x,b_{k+1},...,b_n) \tag{2.6.2}$$

Lema 2.6.4: *Neka su* $S_1, S_2, ..., S_n$ *neprazni podskupovi od* $\overline{\mathbf{R}}$ *i neka je* $\mathrm{H}: S_1 \times S_2 \times \cdots \times S_n \to \mathbf{R}$ *osnovana, n-rastuća funkcija. Tada je funkcija* H *neopadajuća po svim argumentima, ako* $(t_1,...,t_{k-1},x,t_{k+1},...,t_n)$ *i* $(t_1,...,t_{k-1},y,t_{k+1},...,t_n)$ *pripadaju domenu funkcije* H *tj. ako je* $x < y$ *tada je*

$$\mathrm{H}(t_1,...,t_{k-1},x,t_{k+1},...,t_n) \leq \mathrm{H}(t_1,...,t_{k-1},y,t_{k+1},...,t_n).$$

Lema 2.6.5: *Neka su* $S_1, S_2, ..., S_n$ *neprazni podskupovi od* $\overline{\mathbf{R}}$ *i neka je* $\mathrm{H}: S_1 \times S_2 \times \cdots \times S_n \to \mathbf{R}$ *osnovana, n-rastuća funkcija sa marginalnim funkcijama. Neka su* $\mathbf{x} = (x_1, x_2, ..., x_n)$ *i* $\mathbf{y} = (y_1, y_2, ..., y_n)$ *proizvoljne tačke iz* $S_1 \times S_2 \times \cdots \times S_n$. *Tada*

$$\left|\mathrm{H}(\mathbf{x}) - \mathrm{H}(\mathbf{y})\right| \leq \sum_{k=1}^{n} \left|\mathrm{H}_k(x_k) - \mathrm{H}_k(y_k)\right|.$$

Primetimo da sve što smo do sada naveli predstavlja proširenja do dimenzije *n* definicija koje smo navodili u odeljku 2.1 pre nego što smo definisali dvodimenzione kopule. Sada kad smo to uradili u stanju smo da definišemo višedimenzione kopule. Na primer iz prethodne leme koju smo naveli sledi da su *n*-dimenzione kopule uniformno neprekidne, što korisimo u dokazu Sklarove teoreme za višedimenzione kopule. Sada prelazimo na definisanje kopula.

Definicija 2.6.6: *n-dimenziona subkopula* je funkcija $\mathrm{C'}$ koja ima sledeća svojstva

1. $\mathrm{DomC'} = S_1 \times S_2 \times \cdots \times S_n$ gde je svaki S_k podskup od $\mathbf{I} = [0,1]$;
2. $\mathrm{C'}$ je osnovana i n-rastuća;
3. $\mathrm{C'}$ ima jednodimenzione marginalne funkcije C'_k, k=1,2,...,n koje zadovoljavaju

$$\mathrm{C}'_k(u) = u \text{ za svako } u \text{ iz } S_k. \tag{2.6.3}$$

Primetimo da za svaki vektor **u** iz domena funkcije $\mathrm{C'}$ važi $0 \le \mathrm{C'}(\mathbf{u}) \le 1$, tako da je i kodomen funkcije $\mathrm{C'}$ podskup od **I**.

Definicija 2.6.7: *n-dimenziona kopula* je n-dimenziona subkopula čiji je domen $\mathbf{I^n}$**.**

Ekvivalentno, kopula se može definisati i na sledeći način.
Definicija 2.6.8: n-dimenziona kopula je funkcija $\mathrm{C} : \mathbf{I}^n \to \mathbf{I}$, koja ima sledeća svojstva

1. Za svako **u** iz $\mathbf{I}^n$,
   ako je bar jedna koordinata od **u** jednaka nuli važi $\mathrm{C}(\mathbf{u}) = 0$ (2.6.4a)
   i ako su sve koordinate od **u** jednake 1 sem u_k, tada je $\mathrm{C}(\mathbf{u}) = u_k$; (2.6.4b)
2. Za svako **a** i **b** iz $\mathbf{I}^n$ takve da $\mathbf{a} \le \mathbf{b}$ važi $V_{\mathrm{C}}([\mathbf{a},\mathbf{b}]) \ge 0$. (2.6.4c)

Kako je kopula zajednička funkcija raspodele, ona uvodi verovatnosnu meru preko

$$V_{\mathrm{C}}([0,u_1] \times [0,u_2] \times \cdots \times [0,u_n]) = \mathrm{C}(u_1, u_2, ..., u_n).$$

Sledeća teorema je posledica Leme 2.6.5, sličnu stvar smo imali i kada smo opisivali dvodimenzione kopule.

Teorema 2.6.9: *Neka je* $\mathrm{C'}$ *n-dimenziona subkopula. Tada za svaki* **u** *i* **v** *iz domena funkcije* $\mathrm{C'}$ važi

$$\left|\mathrm{C'}(\mathbf{v}) - \mathrm{C'}(\mathbf{u})\right| \le \sum_{k=1}^{n} |v_k - u_k| \tag{2.6.5}$$

Dakle, $\mathrm{C'}$ *je uniformno neprekidna na svom domenu.*

Teorema 2.6.10 (**Sklarova teorema):** *Neka je* H *n-dimenziona funkcija raspodele sa marginlnim funkcijama* $\mathrm{F_1, F_2, ..., F}_n$**.** *Tada postoji n-dimenziona kopula* C *takva da za svako* $\mathbf{x} \in \overline{\mathbf{R}}^n$ *važi*

$$\mathrm{H}(x_1, x_2, ..., x_n) = \mathrm{C}(\mathrm{F}_1(x_1), \mathrm{F}_2(x_2), ..., \mathrm{F}_n(x_n)) \tag{2.6.6}$$

Ako su $\mathrm{F_1, F_2, ..., F}_n$ *neprekidne, tada* C *je jednistvena. U suprotnom C je jedinstveno definisana na* $\mathrm{RanF_1 \times RanF_1 \times \cdots \times RanF_1}$. *Obrnuto, ako je* C *n-dimenziona kopula i* $\mathrm{F_1, F_2, ..., F}_n$ *funkcije raspodele, tada funkcija* H *koja zadovoljava* (2.6.6) *je n-dimenziona funkcija raspodele sa marginama* $\mathrm{F_1, F_2, ..., F}_n$.

Naravno, potrebno je znati šta je to n-dimenziona funkcija raspodele, ali to je jednostavno shvatiti. Samo je potrebno proširiti definiciju koju smo naveli za dvodimenzionu funkciju raspodele u odeljku gde smo obimnije opisivali Sklarovu teoremu. Teoremu 2.6.10

nećemo dokazivati. Potrebno je samo pomenuti da će se dokaz bazirati na proširenjima lema 2.2.3 i 2.2.4 do dimenzije *n*. Dokaz Leme 2.2.3 u *n*-dimenzionom slučaju ne bi trebalo da predstavlja veliki problem, matematičkom indukcijom lako se dolazi do konačnog rezultata. Međutim, Lemu 2.2.4 komplikovano je dokazati i kada imamo slučaj dimenzije *n*=2, ali smo rekli da se dokaz može naći u knjizi Rodžera Nelsena. Dokaz u *n*-dimenzionom slučaju se može pronaći u nekom od sledećih radova Spruill 1975; Deheuvels 1978; Sklar 1996. Kada ubacimo i slučajne promenljive u celu priču o *n*-dimenzionim kopulama, Sklarova teorema izgleda ovako.

Teorema 2.6.11: *Neka su* $X_1, X_2, ..., X_n$ *slučajne promenljive čije su funkcije raspodele redom* $\mathrm{F}_1, \mathrm{F}_2, ..., \mathrm{F}_\mathrm{n}$ *i zajednička funkcija raspodele* H. *Tada postoji n-dimenziona kopula* C *takva da zadovoljava* (2.6.6). *Ako su* $\mathrm{F}_1, \mathrm{F}_2, ..., \mathrm{F}_\mathrm{n}$ *neprekidne* C *je jedinstvena. U suprotnom,* C *je jedinstveno definisana na* $\mathrm{RanF}_1 \times \mathrm{RanF}_1 \times \cdots \times \mathrm{RanF}_1$.

Proširenja kopula M, W i Π do dimenzije *n* su:

$$\mathrm{M}^n(\mathbf{u}) = \min(u_1, u_2, ..., u_n);$$

$$\Pi^n(\mathbf{u}) = u_1 u_2 \cdots u_n,$$

$$\mathrm{W}^n(\mathbf{u}) = \max(u_1 + u_2 + ... + u_n - n + 1, 0).$$

Kopule M^n i Π^n su kopule za $n \geq 2$, ali kopula W^n nije kopula ni za jedno $n > 2$. Bez obzira na tu činjenicu važiće *n*-dimenziona verzija Freše-Hefdingove nejednakosti.

Teorema 2.6.12: *Ako je* C *bilo koja n-dimenziona kopula, tada za svako* **u** *iz* $[0,1]^n$ *važi*

$$\mathrm{W}^n(\mathbf{u}) \leq \mathrm{C}(\mathbf{u}) \leq \mathrm{M}^n(\mathbf{u}) \tag{2.6.7}$$

Dokaz sledi direktno iz Leme 2.6.4 i Leme 2.6.5.

Iako donja Freše-Hefdingova granica nikada nije kopula za $n > 2$ leva strana nejednakosti (2.6.7) je najbolja moguća u slučaju koji ćemo navesti u sledećoj teoremi.

Teorema 2.6.13: *Za bilo koje* $n \geq 3$ i **u** *iz* $[0,1]^n$ *postoji kopula* C, *koja zavisi od* **u**, *takva da*

$$\mathrm{C}(\mathbf{u}) = \mathrm{W}^n(\mathbf{u}).$$

<u>*Dokaz*</u>*:* Neka je $\mathbf{u} = (u_1, u_2, ..., u_n)$ fiksirana tačka iz $\mathbf{I}^n$ koja je različita od $0 = (0,0,...,0)$ ili $1 = (1,1,...,1)$. Razmatramo dva slučaja:

1. Pretpostavimo da $0 < u_1 + u_2 + \cdots + u_n \leq n - 1$. Posmatrajmo vektor $\mathbf{v} = (v_1, v_2, ..., v_n)$, gde je svako v_k jednako ili 0 ili 1 ili $t_k = \min\{(n-1)u_k / (u_1 + \cdots + u_n), 1\}$. Definišemo funkciju C' na ovim tačkama kao $\mathrm{C}'(\mathbf{v}) = \mathrm{W}^n(\mathbf{v})$. Lako se može proveriti da ovako definisana funkcija ima svojstva jedne *n*-dimenzione subkopule. Zatim, ona se može proširiti do n-dimenzione kopule, to nam kaže Lema 2.2.4 proširena do dimenzije n. Tada za svako $\mathbf{x}$ iz $[0,\mathbf{t}]$ (i $\mathbf{u}$ se nalazi tu) , gde je $\mathbf{t}=(t_1,t_2,...,t_n)$ važi da je $\mathrm{C}(\mathbf{x})=\mathrm{W}^n(\mathbf{x})=0$.

2. Pretpostavimo sada da je $n-1<u_1+u_2+\cdots+u_n\le n$ i posmatrajmo vektor $\mathbf{v}=(v_1,v_2,...,v_n)$, gde je svaki od v_k jednak ili 0 ili 1 ili $s_k=(1-k)/[n-(u_1+u_2+\cdots+u_n)]$. Definišemo funkciju C' na isti način kao i u prethodnom slučaju, a zatim je takođe proširujemo da n-dimenzione kopule C. Tada za svako $\mathbf{x}$ iz $[\mathbf{s},1]$ (i $\mathbf{u}$ se nalazi tu) , gde je $\mathbf{s}=(s_1,s_2,...,s_n)$ važi da je $\mathrm{C}(\mathbf{x})=\mathrm{W}^n(\mathbf{x})=x_1+x_2+\cdots+x_n-n+1$.

■

Odeljak o n-dimenzionim kopulama završićemo pričom o kopulama i njihovom povezanošću sa slučajnim veličinama u slučaju kada imamo dimenziju $n\ge 3$. Standardna reprezentacija n-dimenzione kopule slučajnog vektora $(X_1,X_2,...,X_n)^{\mathrm{T}}$ je

$$\mathrm{H}(x_1,x_2,...,x_n)=\mathrm{P}\{X_1\le x_1,X_2\le x_2,...,X_n\le x_n\}=\mathrm{C}(\mathrm{F}_1(x_1),\mathrm{F}_2(x_2),...,\mathrm{F}_n(x_n)) .$$

Naravno, ovde $\mathrm{F}_1,\mathrm{F}_2,...,\mathrm{F}_n$ su funkcije raspodele slučajnih promenljivih redom $X_1,X_2,...,X_n$ i za tako definisanu kopulu kažemo da je to kopula od $X_1,X_2,...,X_n$. Navedimo sada analognu teoremu Teoremi 2.3.2.

Teorema 2.6.14: *Neka je* $(X_1,X_2,...,X_n)^{\mathrm{T}}$ *vektor neprekidnih slučajnih promenljivih sa kopulom* C*, tada su* $X_1,X_2,...,X_n$ *nezavisne ako i samo ako je* $\mathrm{C}=\Pi^n$.

I u slučaju kada imamo posla sa n-dimenzionim kopulama važiće ono lepo svojstvo da za striktno monotone transformacije slučajnih promenljivih kopule će biti invarijantne ili će se neznatno menjati. Pa navedimo sada, za sam kraj ovog odeljka i teoreme koje govore o tome, koje takođe nećemo dokazivati a njihovi dokazi mogu se pronaći u [4].

Teorema 2.6.15: *Neka je* $(X_1,X_2,...,X_n)^{\mathrm{T}}$ *vektor neprekidnih slučajnih veličina sa kopulom* C. *Ako su* $\alpha_1,\alpha_2,...,\alpha_n$ *strogo rastuće funkcije na* $\mathrm{Ran}\,X_1,\mathrm{Ran}\,X_2,\cdots,\mathrm{Ran}\,X_n$, *tada i vektor* $(\alpha_1(X_1),\alpha_2(X_2),...,\alpha_n(X_n))^{\mathrm{T}}$ *ima kopulu* C.

Teorema 2.6.16: *Neka je* $(X_1,X_2,...,X_n)^{\mathrm{T}}$ *vektor neprekidnih slučajnih veličina sa kopulom* C. *Neka su* $\alpha_1,\alpha_2,...,\alpha_n$ *strogo rastuće funkcije na* $\mathrm{Ran}\,X_1,\mathrm{Ran}\,X_2,\cdots,\mathrm{Ran}\,X_n$ *i neka* $(\alpha_1(X_1),\alpha_2(X_2),...,\alpha_n(X_n))^{\mathrm{T}}$ *ima kopulu* $\mathrm{C}_{\alpha_1(X_1),...,\alpha_n(X_n)}$. *Neka* α_k *strogo opada za neko k. Bez gubitka opštosti uzmimo da je k=*1*, tada*

$$\mathrm{C}_{\alpha_1(X_1),...,\alpha_n(X_n)}(u_1,u_2,...,u_n)=\mathrm{C}_{\alpha_2(X_2),...,\alpha_n(X_n)}(u_1,u_2,...,u_n)-\mathrm{C}_{X_1,\alpha_2(X_2),...,\alpha_n(X_n)}(1-u_1,u_2,...,u_n) .$$

3 Zavisnost

Kao što i sam naslov kaže u ovom odeljku, bavićemo se zavisnošću između slučajnih veličina pomoću kopula, odnosno istraživaćemo kakvu ulogu igraju kopule u proučavanju zavisnosti. Struktura zavisnosti između slučajnih veličina je kompletno opisana njihovom zajedničkom funkcijom raspodele. Zavisnost je jedan od najviše proučavanih pojmova u verovatnoći i statistici. Postoje različiti načini na koje možemo da diskutujemo i merimo zavisnost među slučajnim veličinama. Jedna od svakako najpoznatijih mera zavisnosti je linearna korelacija ili linearni koeficijent korelacije, koji je do sedamdesetih godina bio praktično jedina mera zavisnosti. Međutim, linearni koeficijent korelacije je funkcija marginalnih raspodela i zbog toga će svaka promena oblika marginalnih raspodela uticati na promene vrednosti ovog koeficijenta. Uvođenjem kopula želimo da se oslobodimo uticaja marginalnih raspodela, jer nam Teorema 2.3.3 upravo i govori da su kopule invarijantne pri strogo rastućim transformacijama slučajnih veličina. Osobine zavisnosti i mere saglasnosti su veoma povezane pa ćemo u nastavku izlaganja posvetiti pažnju i toj vezi.

3.1 Mere saglasnosti, osnovne definicije

Pre nego što uvedemo svakako najpoznatije mere Kendalovo (Kendall) tau i Spirmanovo (Spearman) ro, koje mere oblik zavisnosti koji je poznat kao saglasnost, objasnimo upravo pojam saglasnosti i mere saglasnosti. Neformalno, par slučajnih veličina je saglasan ako se velike vrednosti jedne slučajne veličine javljaju sa velikim vrednostima druge slučajne veličine. Slično par slučajnih veličina je saglasan ako se male vrednosti jedne javljaju sa malim vrednostima druge slučajne veličine. Dakle, za par slučajnih veličina reći ćemo da je nesaglasan ako se veće vrednosti jedne slučajne veličine javljaju sa manjim vrednostima druge slučajne veličine.

Definicija 3.1.1:

a) Dve realizacije (x_i, y_i), (x_j, y_j) vektora neprekidnih slučajnih veličina (X,Y) su *saglasne* ako $x_i < x_j$ i $y_i < y_j$ ili ako $x_i > x_j$ i $y_i > y_j$. Ekvivalentna karakterizacija saglasnosti je da važi $(x_i - x_j)(y_i - y_j) > 0$. Kažemo da su realizacije (x_i, y_i), (x_j, y_j) *nesaglasne* ako je $(x_i - x_j)(y_i - y_j) < 0$.

b) Ako su C_1 i C_2 kopule, kažemo da je kopula C_1 *manje saglasna* od C_2 (ili C_2 *više saglasna* od C_1) i pišemo $\mathrm{C}_1 \prec \mathrm{C}_2$ ($\mathrm{C}_2 \succ \mathrm{C}_1$) ako $\mathrm{C}_1(u,v) \le \mathrm{C}_2(u,v)$ ($\mathrm{C}_1(u,v) \ge \mathrm{C}_2(u,v)$) za svako $(u,v) \in \mathbf{I}^2$.

Kako bismo pokazali kakvu ulogu kopule imaju u priči o saglasnosti, definisaćemo funkciju saglasnosti Q koja ustvari predstavlja razliku između verovatnoća saglasnosti i nesaglasnosti dva slučajna vektora (X_1, Y_1) i (X_2, Y_2) neprekidnih slučajnih veličina sa različitim funkcijama raspodele H_1 i H_2, ali poznatim marginalnim funkcijama raspodele F i G. Sada ćemo navesti i dokazati teoremu koja nam govori da ova funkcija saglasnosti zavisi od vektora (X_1, Y_1) i (X_2, Y_2) samo preko njihovih kopula.

Teorema 3.1.2: *Neka su* (X_1, Y_1) i (X_2, Y_2) *dva nezavisna vektora neprekidnih slučajnih veličina sa zajedničkim funkcijama raspodele redom* H_1 i H_2 *i poznatim marginalnim funkcijama* F *(od* X_1 *i* X_2*) i* G *(od* Y_1 *i* Y_2*). Neka su* C_1 i C_2 *kopule od* (X_1, Y_1) i (X_2, Y_2) *redom, odnosno* $\mathrm{H}_1(x,y) = \mathrm{C}_1(\mathrm{F}(x), \mathrm{G}(y))$ *i* $\mathrm{H}_2(x,y) = \mathrm{C}_2(\mathrm{F}(x), \mathrm{G}(y))$. *Neka* Q *označava razliku između verovatnoća saglasnosti i nesaglasnosti od* (X_1, Y_1) i (X_2, Y_2), *odnosno*

$$Q = \mathrm{P}\{(X_1 - X_2)(Y_1 - Y_2) > 0\} - \mathrm{P}\{(X_1 - X_2)(Y_1 - Y_2) < 0\}, \tag{3.1.1}$$

tada

$$Q = Q(\mathrm{C}_1, \mathrm{C}_2) = 4\iint_{\mathbf{I}^2} \mathrm{C}_2(u,v)\, d\,\mathrm{C}_1(u,v) - 1 \tag{3.1.2}$$

<u>Dokaz</u>*:* Iz osobina neprekidnih slučajnih veličina znamo da važi sledeće $\mathrm{P}\{(X_1 - X_2)(Y_1 - Y_2) < 0\} = 1 - \mathrm{P}\{(X_1 - X_2)(Y_1 - Y_2) > 0\}$. Kada to vratimo u (3.1.1) imamo da je $Q = 2\mathrm{P}\{(X_1 - X_2)(Y_1 - Y_2) > 0\} - 1$. (**)

Verovatnoću $\mathrm{P}\{(X_1 - X_2)(Y_1 - Y_2) > 0\}$ možemo da napišemo koristeći definiciju saglasnosti na sledeći način

$\mathrm{P}\{(X_1 - X_2)(Y_1 - Y_2) > 0\} = \mathrm{P}\{X_1 > X_2, Y_1 > Y_2\} + \mathrm{P}\{X_1 < X_2, Y_1 < Y_2\}$.

Verovatnoće $\mathrm{P}\{X_1 > X_2, Y_1 > Y_2\}$ i $\mathrm{P}\{X_1 < X_2, Y_1 < Y_2\}$ računamo tako što integralimo po raspodeli jednog od vektora (X_1, Y_1) ili (X_2, Y_2). Na primer uzmimo vektor (X_1, Y_1), tada imamo

$$\begin{aligned}
\mathrm{P}\{X_1 > X_2, Y_1 > Y_2\} &= \mathrm{P}\{X_2 < X_1, Y_2 < Y_1\}, \\
&= \iint_{\mathbf{R}^2} \mathrm{P}\{X_2 \le X_1, Y_2 \le Y_1\}\, d\,\mathrm{C}_1(\mathrm{F}(x), \mathrm{G}(y)), \\
&= \iint_{\mathbf{R}^2} \mathrm{C}_2(\mathrm{F}(x), \mathrm{G}(y))\, d\,\mathrm{C}_1(\mathrm{F}(x), \mathrm{G}(y)),
\end{aligned}$$

uvedemo smenu $u = \mathrm{F}(x), v = \mathrm{G}(y)$ i dobijamo

$$= \iint_{\mathbf{I}^2} \mathrm{C}_2(u,v)\, d\,\mathrm{C}_1(u,v).$$

Na sličan način

$$\begin{aligned}
\mathrm{P}\{X_1 < X_2, Y_1 < Y_2\} &= \mathrm{P}\{X_2 > X_1, Y_2 > Y_1\}, \\
&= \iint_{\mathbf{R}^2} \mathrm{P}\{X_2 > X_1, Y_2 > Y_1\}\, d\,\mathrm{C}_1(\mathrm{F}(x), \mathrm{G}(y)), \\
&= \iint_{\mathbf{R}^2} (1 - \mathrm{F}(x) - \mathrm{G}(y) + \mathrm{C}_2(\mathrm{F}(x), \mathrm{G}(y)))\, d\,\mathrm{C}_1(\mathrm{F}(x), \mathrm{G}(y)),
\end{aligned}$$

Kada uvedemo istu smenu kao gore dobijamo

$$= \iint_{\mathbf{I}^2} (1 - u - v + \mathrm{C}_2(u,v))\, d\,\mathrm{C}_1(u,v).$$

Ali kako je C_1 zajednička funkcija raspodele para slučajnih veličina (U,V) koje su uniformno raspodeljene na (0,1) važiće $\mathrm{E}(U)=\mathrm{E}(V)=\frac{1}{2}$.

$$=\iint_{\mathbf{I}^2} d\,\mathrm{C}_1(u,v)-\iint_{\mathbf{I}^2} ud\,\mathrm{C}_1(u,v)-\iint_{\mathbf{I}^2} vd\,\mathrm{C}_1(u,v)+\iint_{\mathbf{I}^2}\mathrm{C}_2(u,v)d\,\mathrm{C}_1(u,v)$$

$$=1-\frac{1}{2}-\frac{1}{2}+\iint_{\mathbf{I}^2}\mathrm{C}_2(u,v)d\,\mathrm{C}_1(u,v)$$

$$=\iint_{\mathbf{I}^2}\mathrm{C}_2(u,v)d\,\mathrm{C}_1(u,v).$$

Dakle, pokazali samo da je $\mathrm{P}\{(X_1-X_2)(Y_1-Y_2)>0\}=2\iint_{\mathbf{I}^2}\mathrm{C}_2(u,v)d\,\mathrm{C}_1(u,v)$, tj. kada sada taj dobijeni rezultat vratimo u (**) vidimo da je zaista $Q=4\iint_{\mathbf{I}^2}\mathrm{C}_2(u,v)d\,\mathrm{C}_1(u,v)-1$.

■

Posledica : Neka su $\mathrm{C}_1,\mathrm{C}_2$ i Q date kao u Teoremi 3.1.2, tada:

1. Q je simetrična, odnosno $Q(\mathrm{C}_1,\mathrm{C}_2)=Q(\mathrm{C}_2,\mathrm{C}_1)$.
2. Q je neopadajuća po svakom argumentu: ako $\mathrm{C}_1 \prec \mathrm{C}_3$, tada $Q(\mathrm{C}_1,\mathrm{C}_2)\le Q(\mathrm{C}_3,\mathrm{C}_2)$.
3. Kopule se u Q mogu zameniti kopulama preživljavanja, tj. $Q(\mathrm{C}_1,\mathrm{C}_2)=Q(\hat{\mathrm{C}}_1,\hat{\mathrm{C}}_2)$.

Napomena: Neka je C kopula. Kako je Q razlika verovatnoća to znači da ona uzima vrednosti iz intervala $[-1,1]$.

Primer 3.1: Funkcija saglasnosti para kopula Π se može lako izračunati jer su margine kopule proizvoda uniformne na jediničnom intervalu. Odnosno funkcija preživljavanja u tom slučaju data je u sledećem obliku

$$Q(\Pi,\Pi)=4\iint_{\mathbf{I}^2} uv\,d\Pi(u,v)-1$$

$$=4\int_0^1\int_0^1 uv\;dudv-1$$

$$=0.$$

Na sličan način mogu se izračunati funkcije saglasnosti za parove Freše-Hefdingovih granica kao i za njihove kombinacije, ali nećemo o tome detaljnije.

□

Definicija 3.1.3: Kažemo da je mera $\kappa_{\mathrm{C}}=\kappa_{X,Y}$ *mera saglasnosti* ako zadovoljava sledeće uslove:

1. $\kappa_{X,Y}$ je definisana za svaki par slučajnih veličina *X,Y*;
2. $-1\le\kappa_{X,Y}\le 1$, $\kappa_{X,X}=1$, $\kappa_{-X,X}=-1$;
3. $\kappa_{X,Y}=\kappa_{Y,X}$;

4. ako su X i Y nezavisne tada je $\kappa_{X,Y} = \kappa_{\Pi} = 0$;

5. $\kappa_{-X,Y} = \kappa_{X,-Y} = -\kappa_{X,Y}$;

6. ako su C_1 i C_2 kopule takve da $\mathrm{C}_1 \prec \mathrm{C}_2$ tada $\kappa_{\mathrm{C}_1} \le \kappa_{\mathrm{C}_2}$;

7. ako je $\{(X_n, Y_n)\}$ niz neprekidnih slučajnih veličina sa kopulama C_n i ako $\{\mathrm{C}_n\}$ konvergira ka C, tada $\lim_{n\to\infty} \kappa_{X_n,Y_n} = \kappa_{\mathrm{C}}$.

Navedimo sada jedu važnu posledicu ove definicije. Ako na primer primenimo uslov 6 dva puta dobijamo da je $\mathrm{C}_1 = \mathrm{C}_2$ što nam dalje govori da je $\kappa_{\mathrm{C}_1} = \kappa_{\mathrm{C}_2}$. Ako su X i Y slučajne veličine sa kopulom C i ako su transformacije tih slučajnih veličina α i β strogo rastuće mi znamo da važi $\mathrm{C}_{X,Y} = \mathrm{C}_{\alpha(X),\beta(Y)}$. Dakle biće $\kappa_{X,Y} = \kappa_{\alpha(X),\beta(Y)}$. Iz uslova 5 može se videti da će važiti sličan rezultat i kada su transformacije strogo opadajuće, odnosno, kažemo da mere saglasnosti su invarijantne pri strogo monotonim transformacijama slučajnih veličina.
Ako je $\alpha(X)$ strogo rastuća (opadajuća) transformacija od X, tada $\mathrm{C}_{X,\alpha(X)} = \mathrm{C}_{X,X}$. Koristeći uslov 2 imamo da je $\kappa_{X,X} = \kappa_{\mathrm{M}} = 1 \; (= \kappa_{\mathrm{W}} = -1)$. Odnosno drugim rečima mera saglasnosti ima svoju maksimalnu (minimalnu) vrednost ako nosač zajedničke funkcije raspodele od X i $\alpha(X)$ sadrži samo saglasne (nesaglasne) parove. Sve ovo može se iskazati i u obliku teoreme.

Teorema 3.1.4: *Neka je κ mera saglasnosti za nezavisne slučajne veličine X i Y:*

1. ako je Y skoro sigurno rastuća funkcija od X, tada $\kappa_{X,Y} = \kappa_{\mathrm{M}} = 1$,

2. ako je Y skoro sigurno opadajuća funkcija od X, tada $\kappa_{X,Y} = \kappa_{\mathrm{W}} = -1$,

3. ako su α i β skoro sigurno strogo monotone funkcije na kodomenima od X i Y redom, tada $\kappa_{\alpha(X),\beta(Y)} = \kappa_{X,Y}$.

Istu stvar govori i sledeća lema:
Lema 3.1.5:
a) Mera saglasnosti je invarijantna pri strogo monotonim transformacijama slučajnih veličina.
b) Mera saglasnosti dostiže svoju maksimalnu (minimalnu) vrednost ako nosač zajedničke funkcije raspodele od X i $\alpha(X)$ sadrži samo saglasne (nesaglasne) parove.

U nastavku, bavićemo se dvema najpoznatijim merama saglasnosti Kendalovim tau i Spirmanovim ro.

3.2 Kendalovo tau

Kendalovo tau ili Kendalov koeficijent korelacije ranga je neparameteraka mera veze bazirana na broju saglasnih i nesaglasnih parova slučajnih promenljivih iz uzorka. Drugim rečima, neka je $\{(x_1,y_1),(x_2,y_2),...,(x_n,y_n)\}$ slučajan uzorak od *n* realizacija vektora (X,Y) neprekidnih slučajnih veličina. Postoji $\binom{n}{2}$ različitih parova (x_i,y_i) i (x_j,y_j) iz uzorka i svaki od tih parova su ili saglasni ili nesaglasni. Označimo sa *c* broj saglasnih parova, a sa *d* broj nesaglasih parova. Tada je Kendalovo tau definisano kao

$$t=\frac{c-d}{c+d}=(c-d)\Big/\binom{n}{2}. \tag{3.2.1}$$

t predstavlja razliku između verovatnoće saglasnosti i verovatnoće nesaglasnosti za par realizacija (x_i,y_i) i (x_j,y_j) uzetih iz uzorka na slučajan način.

Međutim, mi možemo da pričamo o Kendalovom tau i na nivou populacije. Pa tako pretpostavimo da imamo dva slučajna vektora (X_1,Y_1) i (X_2,Y_2), takva da su nezavisni, jednako raspodeljeni i svaki od njih ima zajedničku funkciju raspodele H. U tom slučaju Kendalovo tau slučajnog vektora (X,Y) sa funkcijom raspodele H je

$$\tau=\tau_{X,Y}=\mathrm{P}\{(X_1-X_2)(Y_1-Y_2)>0\}-\mathrm{P}\{(X_1-X_2)(Y_1-Y_2)<0\}. \tag{3.2.2}$$

Primetimo da je (3.2.2) identična (3.1.1), ali smo kod funkcije saglasnosti imali da slučajni vektori (X_1,Y_1) i (X_2,Y_2) nemaju istu zajedničku funkciju raspodele. U tome je jedina razlika, pa ćemo zbog toga navesti teoremu koristeći (3.2.2), (3.1.1) i (3.1.2), koja važi za Kendalovo tau.

Teorema 3.2.1: *Neka su X i Y neprekidne slučajne veličine sa kopulom* C. *Tada Kendalovo tau za X i Y dato je sa*

$$\tau_{X,Y}=\tau_C=Q(C,C)=4\iint_{\mathbf{I}^2}\mathrm{C}(u,v)\,d\,\mathrm{C}(u,v)-1\ . \tag{3.2.3}$$

Ovu teoremu nećemo dokazivati jer je dokaz identičan onome iz Teoreme 3.1.2, samo sa zamenjenim oznakama i slučajnim veličinama.

Može se primetiti da integral u formuli (3.2.3) ustvari predstavlja matematičko očekivanje funkcije C(*U*,*V*), gde su *U* i *V* slučajne veličine sa raspodelom U(0,1), a C je njihova zajednička funkcija raspodele. Odnosno Kendalovo tau je ustvari

$$\tau_C=4\mathrm{E}(\mathrm{C}(U,V))-1. \tag{3.2.4}$$

Primer 3.2: Neka nam je data kopula C_θ koja pripada Farli-Gumbel-Morgenstern familiji kopula, tj.

$$\mathrm{C}_\theta(u,v)=uv[1+\theta(1-u)(1-v)],\ \theta\in[-1,1].$$

Želimo da izračunamo vrednost Kendalovog tau za ovako zadatu kopulu.

$$\tau_{C_\theta} = Q(\mathrm{C}_\theta, \mathrm{C}_\theta) = 4\iint_{\mathbf{I}^2} \mathrm{C}_\theta(u,v) d\,\mathrm{C}_\theta(u,v) - 1,$$

$$= 4\iint_{\mathbf{I}^2} uv[1+\theta(1-u)(1-v)] d\,\mathrm{C}_\theta(u,v) - 1,$$

gde je $d\,\mathrm{C}_\theta(u,v) = \dfrac{\partial^2 \mathrm{C}_\theta(u,v)}{\partial u \partial v} = (1+\theta(1-2u)(1-2v))dudv.$

Kada se to zameni u prethodnoj jednakosti i reši jednostavan ali malo robustan dvostruki integral dobijamo da je

$$\tau_{C_\theta} = Q(\mathrm{C}_\theta, \mathrm{C}_\theta) = \frac{2\theta}{9}.$$

Kako parametar $\theta \in [-1,1]$ to znači da je vrednost Kendalovog tau za ovakvu familiju kopula ograničena na intervalu, tj. $\tau_\theta \in [-2/9, 2/9]$. Sama činjenica da je zavisnost ograničena govori nam da je korist i primena ovakve familije kopula veoma mala.

□

Teorema 3.2.2: *Neka su* C_1 *i* C_2 *kopule. Tada važi*

$$\iint_{\mathbf{I}^2} \mathrm{C}_1(u,v) d\,\mathrm{C}_2(u,v) = \frac{1}{2} - \iint_{\mathbf{I}^2} \frac{\partial}{\partial u}\mathrm{C}_1(u,v) \frac{\partial}{\partial v}\mathrm{C}_2(u,v) dudv. \qquad (3.2.5)$$

<u>Dokaz</u>*:* Videti [13].

Kao posledicu ove teoreme imamo jednakost izraza (3.2.3) i sledećeg izraza

$$\tau_C = 1 - 4\iint_{\mathbf{I}^2} \frac{\partial}{\partial u}\mathrm{C}(u,v) \frac{\partial}{\partial v}\mathrm{C}(u,v) dudv. \qquad (3.2.6)$$

Ova jednakost nam omogućava lakse izračunavanje, jer u jednakosti (3.2.3) može doći do određenih problema ukoliko su kopule singularne ili ako imaju i singularni i apsolutno neprekidni deo, pa je za većinu kopula pogodnije koristiti (3.2.6).

Kao jednu od primena formule (3.2.6) navedimo sledeću lemu.

Lema 3.2.3: $\tau_C = \tau_{\overline{\mathrm{C}}}$, *gde je* $\overline{\mathrm{C}}$ *definisana kao u jednakosti (2.5.2).*

<u>Dokaz</u>*:* $\tau_{\overline{\mathrm{C}}} = 1 - 4\iint_{\mathbf{I}^2} \dfrac{\partial \overline{\mathrm{C}}}{\partial u} \dfrac{\partial \overline{\mathrm{C}}}{\partial v} dudv,$

$$= 1 - 4\iint_{\mathbf{I}^2} [1 - \frac{\partial \mathrm{C}}{\partial u}][1 - \frac{\partial \mathrm{C}}{\partial v}] dudv,$$

$$= \tau_{\mathrm{C}} - 4\iint_{\mathbf{I}^2} [1 - \frac{\partial \mathrm{C}}{\partial u} - \frac{\partial \mathrm{C}}{\partial v}] dudv,$$

$$= \tau_{\mathrm{C}} - 4\left[\iint_{\mathbf{I}^2} 1 dudv - \iint_{\mathbf{I}^2} \frac{\partial \mathrm{C}}{\partial u} dudv - \iint_{\mathbf{I}^2} \frac{\partial \mathrm{C}}{\partial v} dudv\right] = (*); \qquad (1)$$

Integral $\iint_{\mathrm{I}^2} \frac{\partial \mathrm{C}}{\partial u} dudv$ iz poslednje nejednakosti možemo da izračunamo i on će biti jednak:

$$\iint_{\mathrm{I}^2} \frac{\partial \mathrm{C}}{\partial u} dudv = \int_0^1 [\mathrm{C}(1,v) - \mathrm{C}(0,v)] dv = \int_0^1 \mathrm{C}(1,v) dv = \int_0^1 v dv = \frac{1}{2}.$$

Na sličan način je $\iint_{\mathrm{I}^2} \frac{\partial \mathrm{C}}{\partial v} dudv = \frac{1}{2}$. Sada vratimo dobijene rezultate u (1), odakle imamo da je :

$$(*) = \tau_{\mathrm{C}} - 4(\frac{5}{4} - \frac{1}{2} - \frac{1}{2}) = \tau_{\mathrm{C}}.$$

Odnosno dokazali smo da zaista važi lema.

■

3.3 Spirmanovo ro

Neka su data tri nezavisna slučajna vektora (X_1, Y_1), (X_2, Y_2) i (X_3, Y_3) sa istom zajedničkom funkcijom raspodele H i kopulom C. Mera saglasnosti poznata kao Spirmanovo ro proporcionalna je razlici verovatnoće saglasnosti i verovatnoće nesaglasnosti dva vektora (X_1, Y_1) i (X_2, Y_3) (može se posmatrati i vektor (X_3, Y_2)). Ova dva para vektora imaju iste marginalne funkcije F i G funkcije H, ali H je funkcija raspodele samo jednog od tih vektora, dok komponente drugog vektora moraju biti nezavisne. Za tako definisane vektore Spirmanovo ro je sledećeg oblika

$$\rho_{X,Y} = 3[\mathrm{P}\{(X_1 - X_2)(Y_1 - Y_3) > 0\} - \mathrm{P}\{(X_1 - X_2)(Y_1 - Y_3) < 0\}]. \qquad (3.3.1)$$

Kako je funkcija $\mathrm{H}(x,y)$ zajednička funkcija raspodele vektora (X_1, Y_1), tada će funkcija raspodele vektora (X_2, Y_3) biti $\mathrm{F}(x)\mathrm{G}(y)$ jer su komponente tog vektora nezavisne kao što smo već i rekli u prethodnoj rečenici. Odnosno vektor (X_2, Y_3) će imati kopulu Π.

Teorema 3.3.1: *Neka su X i Y neprekidne slučajne veličine sa kopulom* C. *Tada Spirmanovo ro od X i Y dato je kao*

$$\rho_{X,Y} = \rho_{\mathrm{C}} = 3Q(\mathrm{C}, \Pi), \qquad (3.3.2a)$$

$$= 12 \iint_{\mathrm{I}^2} uv\, d\mathrm{C}(u,v) - 3. \qquad (3.3.2b)$$

$$= 12 \iint_{\mathrm{I}^2} \mathrm{C}(u,v) dudv - 3. \qquad (3.3.2c)$$

I dokaz ove teoreme sledi iz dokaza Teoreme 3.1.2. Potrebno je napomenuti i da jednakost (3.3.2c) važi iz osobine da je funkcija Q simetrična.

Primer 3.3: Neka nam je data kopula $\mathrm{C}_{\alpha,\beta}, \alpha > 0, \beta < 1$ koja pripada Maršal-Olkin familiji kopula. Tj.

$$\mathrm{C}_{\alpha,\beta}(u,v) = \begin{cases} u^{1-\alpha}v, & u^{\alpha} \geq v^{\beta} \\ uv^{1-\beta}, & u^{\alpha} \leq v^{\beta}. \end{cases}$$

Pa Spirmanovo ro za ovako definisanu kopulu računamo na sledeći način

$$\rho_{\mathrm{C}_{\alpha,\beta}} = 12\iint_{\mathrm{I}^2} \mathrm{C}_{\alpha,\beta}(u,v)\,dudv - 3.$$

Kako postoje dva slučaja u zavisnosti od kojih rešavamo ovaj integral će biti

$$\iint_{\mathrm{I}^2} \mathrm{C}_{\alpha,\beta}(u,v)\,dudv = \frac{1}{2}\left(\frac{\alpha+\beta}{2\alpha-\alpha\beta+2\beta}\right), \text{ odnosno}$$

$$\rho_{\mathrm{C}_{\alpha,\beta}} = 12\cdot\frac{1}{2}\left(\frac{\alpha+\beta}{2\alpha-\alpha\beta+2\beta}\right) - 3$$

$$= \frac{3\alpha\beta}{2\alpha-\alpha\beta+2\beta}.$$

□

Spirmanovo ro možemo računati i kao

$$\rho_C = 12\,\mathrm{E}(U,V) - 3,$$

ili drugačije zapisano

$$\rho_C = \frac{\mathrm{E}(U,V) - \frac{1}{4}}{\frac{1}{12}}, \tag{3.3.3}$$

Odakle se može primeti da Spirmanovo ro ne predstavlja ništa drugo do Pirsonov koeficijent korelacije dve slučajne veličine sa U(0,1) raspodelom. Odnosno,

$$\rho_C = \frac{\mathrm{E}(U,V) - \frac{1}{4}}{\frac{1}{12}} = \frac{\mathrm{E}(UV) - \mathrm{E}(U)\,\mathrm{E}(V)}{\sqrt{\mathrm{D}(U)\,\mathrm{D}(V)}}.$$

Pored ove interpretacije Spirmanovo ro možemo predstaviti na još jedan način. Posmatrajmo jednakost (3.3.2c). Integral u toj jednakosti predstavlja zapreminu grafika kopule na jediničnom kvadratu i njega možemo da zamenimo sledećim integralom

$$\rho_{\mathrm{C}} = 12\iint_{\mathrm{I}^2} (\mathrm{C}(u,v) - uv)\,dudv, \tag{3.3.4}$$

gde dobijamo da će mera ρ_{C} biti proporcijonalna zapremini između grafika kopule C i kopule proizvoda Π. Drugim rečima, kažemo da je ρ_{C} mera srednjeg rastojanja između raspodele od *X* i *Y* i nezavisnosti.

Sada ćemo navesti teoremu u kojoj ćemo videti da su i Kendalovo tau i Spirmanovo ro mere saglasnosti.

Teorema 3.3.2: *Ako su X i Y neprekidne slučajne veličine sa kopulom* C, *tada Kendalovo tau i Spirmanovo ro računati na celoj populaciji zadovoljavaju uslove iz Definicije 3.1.3 Teoreme 3.1.4 za meru saglasnosti.*

<u>Dokaz</u>*:* Prvih šest osobina iz Definicije 3.1.3 zadovoljavaće i Kendalovo tau i Spirmanovo ro, a to sve jasno sledi zbog osobina funkcije *Q* (pogledati Teoremu 3.1.2, Posledicu i Napomenu koje smo naveli u odeljku o merama saglasnosti). Što se tiče sedme osobine važiće i ona, potrebno je samo da se prisetimo Teoreme 2.1.10, odnosno Lipšicovog kriterijuma za kopule, odakle sledi da niz $\{\mathrm{C}_n\}$ konvergira uniformno ka C.

Pored osobina iz Definicije 3.1.3 koje mere saglasnosti moraju da ispunjavaju, postoje još neka dodatna svojstva koja su jako korisna. Međutim, sledeće osobine koje ćemo navesti ne važe generalno za sve mere saglasnosti.

Teorema 3.3.3: *Neka su X i Y neprekidne slučajne veličine sa kopulom* C, *neka su* τ_{C} *i* ρ_{C} *Kendalovo tau i Spirmanovo ro. Tada važi*

1. $\mathrm{C}=\mathrm{M} \Leftrightarrow \tau_{\mathrm{C}}=\rho_{\mathrm{C}}=1$;
2. $\mathrm{C}=\mathrm{W} \Leftrightarrow \tau_{\mathrm{C}}=\rho_{\mathrm{C}}=-1$,

gde su M *i* W *donja, odnosno gornja Freše-Hefdingova granica.*

<u>Dokaz</u>*:* Videti [5].

3.4 Veza između Kendalovog tau i Spirmanovog ro

Vrednosti ova dva koeficijenta se često veoma razlikuju i pored činjenice da oba predstavljaju razlike verovatnoća saglasnosti, odnosno nesaglasnosti između slučajnih veličina sa zadatom kopulom. U teoremama koje ćemo u nastavku navesti pokazaćemo u kolikoj meri mogu ove dve mere saglasnosti da se razlikuju.

Teorema 3.4.1: *Neka su X i Y neprekidne slučajne veličine i neka su* τ *i* ρ *Kendalovo tau i Spirmanovo ro definisani kao* (3.2.2)*, odnosno* (3.3.1)*. Tada*

$$-1 \leq 3\tau - 2\rho \leq 1. \tag{3.4.1}$$

<u>*Dokaz*</u>*:* Neka su (X_1, Y_1), (X_2, Y_2) i (X_3, Y_3) tri nezavisna slučajna vektora sa poznatom zajedničkom raspodelom. Zbog neprekidnosti, definicije Kendalovog tau i Spirmanovog ro možemo zameniti sledećim zapisom

$$\tau = 2\mathrm{P}\{(X_1 - X_2)(Y_1 - Y_2) > 0\} - 1 \tag{3.4.2}$$

i

$$\rho = 6\mathrm{P}\{(X_1 - X_2)(Y_1 - Y_3) > 0\} - 3.$$

Indeksi slučajnih veličina *X* i *Y* mogu se ciklično menjati kako bismo dobili sledeće simetrične forme za τ i ρ:

$$\tau = \frac{2}{3}[\mathrm{P}\{(X_1 - X_2)(Y_1 - Y_2) > 0\} + \mathrm{P}\{(X_2 - X_3)(Y_2 - Y_3) > 0\} \\ + \mathrm{P}\{(X_3 - X_1)(Y_3 - Y_1) > 0\}]\text{-}1;$$

$$\rho = [\mathrm{P}\{(X_1 - X_2)(Y_1 - Y_3) > 0\} + \mathrm{P}\{(X_1 - X_3)(Y_1 - Y_2) > 0\} \\ + \mathrm{P}\{(X_2 - X_1)(Y_2 - Y_3) > 0\} \\ + \mathrm{P}\{(X_3 - X_2)(Y_3 - Y_1) > 0\} \\ + \mathrm{P}\{(X_2 - X_3)(Y_2 - Y_1) > 0\} \\ + \mathrm{P}\{(X_3 - X_1)(Y_3 - Y_2) > 0\}\text{-}3.$$

Pošto su izrazi za τ i ρ invarijantni pri bilo kojim permutacijama indeksa slučajnih veličina možemo pretpostaviti da je $X_1 < X_2 < X_3$. U tom slučaju biće

$$\tau = \frac{2}{3}[\mathrm{P}\{Y_1 < Y_2\} + \mathrm{P}\{Y_2 < Y_3\} + \mathrm{P}\{Y_1 < Y_3\}] - 1;$$

i

$$\rho = [\mathrm{P}\{Y_1 < Y_3\} + \mathrm{P}\{Y_1 < Y_2\} + \mathrm{P}\{Y_2 > Y_3\} + \mathrm{P}\{Y_3 > Y_1\} + \mathrm{P}\{Y_2 < Y_1\} + \mathrm{P}\{Y_3 > Y_2\}] - 3$$

$= 2\,\mathrm{P}\{Y_1 < Y_3\} - 1.$

Definišimo sada verovatnoću $p_{ijk} = \mathrm{P}\{Y_i < Y_j < Y_k \mid X_1 < X_2 < X_3\}$. Sa ovako definisanom verovatnoćom imamo da je

$$\tau = \frac{2}{3}[(p_{123} + p_{132} + p_{312}) + (p_{123} + p_{213} + p_{231}) + (p_{123} + p_{132} + p_{213})] - 1$$
$$= p_{123} + \frac{1}{3}(p_{132} + p_{213}) - \frac{1}{3}(p_{231} + p_{312}) - p_{321};$$

i

$$\rho = 2[(p_{123} + p_{132} + p_{213})] - 1$$
$$= p_{123} + p_{132} + p_{213} - p_{231} - p_{312} - p_{321}.$$

Sa ovakvim τ i ρ važi da je

$$3\tau - 2\rho = p_{123} - p_{132} - p_{213} + p_{231} + p_{312} - p_{321}$$
$$= (p_{123} + p_{231} + p_{312}) - (p_{132} + p_{213} + p_{321}),$$

odakle vidimo da će važiti $-1 \le 3\tau - 2\rho \le 1$, jer je $3\tau - 2\rho$ razlika verovatnoća.

■

Teorema 3.4.2: *Neka su X, Y,* τ *i* ρ *definisani kao u Teoremi* 3.3.2. *Tada*

$$\frac{1+\rho}{2} \ge \left(\frac{1+\tau}{2}\right)^2 \tag{3.4.5a}$$

i

$$\frac{1-\rho}{2} \ge \left(\frac{1-\tau}{2}\right)^2 \tag{3.4.5a}$$

<u>*Dokaz*</u>*:* Neka su (X_1, Y_1), (X_2, Y_2) i (X_3, Y_3) tri nezavisna slučajna vektora sa istom zajedničkom funkcijom raspodele H. Definišimo verovatnoću *p* kao

$$p = \mathrm{P}\{(X_2, Y_2)\ \mathrm{i}\ (X_3, Y_3)\ \text{su saglasni sa}\ (X_1, Y_1)\};$$
$$p = \iint_{\mathbf{R}^2} \mathrm{P}\{(X_2, Y_2)\ \mathrm{i}\ (X_3, Y_3)\ \text{su saglasni sa}\ (x, y)\}d\,\mathrm{H}(x, y),$$
$$= \iint_{\mathbf{R}^2} \mathrm{P}\{(X_2 - x)(Y_2 - y) > 0\}\mathrm{P}\{(X_3 - x)(Y_3 - y) > 0\}d\,\mathrm{H}(x, y),$$
$$= \iint_{\mathbf{R}^2} \left(\mathrm{P}\{(X_2 - x)(Y_2 - y) > 0\}\right)^2 d\,\mathrm{H}(x, y),$$

Iskoristimo sada činjenicu, koju ćemo primeniti u prethodnoj jednakosti, da za uslovnu verovatnoću oblika $Z = \mathrm{P}\{(X_1 - X_2)(Y_1 - Y_2) > 0 \mid (X_1, Y_1)\}$ važi nejednakost $\mathrm{E}(Z^2) \ge (\mathrm{E}(Z))^2$. Odnosno u našem slučaju sledi

$$\ge \left(\iint_{\mathbf{R}^2} \mathrm{P}\{(X_2 - x)(Y_2 - y) > 0\}d\,\mathrm{H}(x, y)\right)^2,$$

$$= \left(\mathrm{P}\{(X_2 - x)(Y_2 - y) > 0\}\right)^2$$
$$= \left(\frac{1+\tau}{2}\right)^2$$

Permutacijama indeksa dobijamo da je

$$p = \frac{1}{3}[\mathrm{P}\{(X_2, Y_2) \text{ i } (X_3, Y_3) \text{ su saglasni sa } (X_1, Y_1)\}$$
$$+ \mathrm{P}\{(X_3, Y_3) \text{ i } (X_1, Y_1) \text{ su saglasni sa } (X_2, Y_2)\}$$
$$+ \mathrm{P}\{(X_1, Y_1) \text{ i } (X_2, Y_2) \text{ su saglasni sa } (X_3, Y_3)\}].$$

Kao i u dokazu prethodne teoreme pretpostavimo da je $X_1 < X_2 < X_3$, kao i da je $p_{ijk} = \mathrm{P}\{Y_i < Y_j < Y_k \mid X_1 < X_2 < X_3\}$, tada je

$$p = \frac{1}{3}[(p_{123} + p_{132}) + p_{123} + (p_{123} + p_{213})],$$
$$= p_{123} + \frac{1}{3} p_{132} + \frac{1}{3} p_{213}.$$

Prisetimo se da smo takođe u dokazu prethodne teoreme pokazali da važi $\rho = 2[(p_{123} + p_{132} + p_{213})] - 1$, odnosno drugačije napisano $\frac{1+\rho}{2} = p_{123} + p_{132} + p_{213}$. Pa možemo da primetimo da će biti $p \leq \frac{1+\rho}{2}$.

Odnosno $\frac{1+\rho}{2} \geq p \geq \left(\frac{1+\tau}{2}\right)^2$, tj. dokazali smo teoremu.

Da bismo dokazali drugu nejednakost o kojoj govori teorema potrebno je da u definiciji verovatnoće p stavimo reč nesaglasni umesto saglasni i dokaz će se prirodno odvijati slično kao i u dokazivanju ove nejednakosti.

■

I za kraj ovog odeljka navedimo još jednu teoremu koju nećemo dokazivati, a koja nam govori u kakvom sve još odnosu mogu da budu Kendalovo tau i Spirmanovo ro.

Teorema 3.4.3: *Neka su X, Y,* τ *i* ρ *definisani kao u Teoremi* 3.3.2. *Tada*

$$\frac{3\tau - 1}{2} \leq \rho \leq \frac{1 + 2\tau - \tau^2}{2}, \; za \;\; \tau \geq 0$$

i

$$\frac{\tau^2 + 2\tau - 1}{2} \leq \rho \leq \frac{1 + 3\tau}{2}, \text{ za } \; \tau \leq 0$$

Primer 3.4: Neka nam je data kopula iz familije kopula Farli-Gumbel-Morgenstern za koju smo već računali Kendalovo tau $\tau_{C_\theta} = \frac{2\theta}{9}$. Sada ćemo izračunati i Spirmanovo ro za ovu familiji proveriti da li važi neka od ovih jednakosti koje su bile date teoremama.

$$\begin{aligned}\rho_{C_\theta} &= 12\iint_{I^2} C_\theta(u,v)\,dudv - 3\\ &= 12\left(\frac{1}{4}+\frac{\theta}{36}\right)-3\\ &= \frac{\theta}{3}.\end{aligned}$$

Proverimo na primer da li važi nejednakost iz Teoreme 3.4.1:

$$3\tau - 2\rho = \frac{6\theta}{9} - \frac{2\theta}{3} = 0.$$

Zaista, ova razlika pripada intervalu $[-1,1]$.

Sada ćemo proveriti jednu od nejednakosti iz Teoreme 3.4.2, a druga se dokazuje na isti način:

$$\begin{aligned}\frac{1+\rho}{2} - \left(\frac{1+\tau}{2}\right)^2 &= \frac{1+\frac{\theta}{3}}{2} - \left(\frac{1+\frac{2\theta}{9}}{2}\right)^2\\ &= \frac{3+\theta}{6} - \left(\frac{9+2\theta}{18}\right)^2\\ &= \frac{162+54\theta-81-36\theta-4\theta^2}{324}\\ &= \left(\frac{9-2\theta}{18}\right)^2 \geq 0.\end{aligned}$$

Tj. $\frac{1+\rho}{2} \geq \left(\frac{1+\tau}{2}\right)^2$.

□

3.5 Ostale mere saglasnosti

Način na koji smo naveli Teoremu 3.3.1 govori nam da Spirmanovo ro nije ništa drugo do mera saglasnosti između kopule C slučajnih veličina *X* i *Y* i nezavisnosti koja je data kopulom proizvoda Π. Drugačije još kažemo da Spirmanovo ro meri rastojanje između ove dve kopule. U smislu tog rastojanja, definisaćemo koeficijent po imenu Đinijevo (Gini) gama koje će meriti saglasnost između kopule C i monotone zavisnosti, odnosno kopula M i W. On je definisao koeficijent γ kao

$$\gamma = 2\iint_{\mathbf{I}^2}\left(|u+v-1|-|u-v|\right)d\,\mathrm{C}(u,v), \tag{3.5.1}$$

čije se detaljno izvođenje može naći u knjizi Nelsen (2006.). Formula (3.5.1) je ekvivalentna sa

$$\gamma = 2\mathrm{E}(|U+V-1|-|U-V|), \tag{3.5.2}$$

što možemo da protumačimo kao očekivanu razliku između (U,V) i dijagonale $[0,1]\times[0,1]$.

Teorema 3.5.1: *Neka su X i Y neprekidne slučajne veličine sa kopulom* C. *Tada je mera saglasnosti, poznatija kao Đinijevo gama, koju ćemo obeležavati sa* $\gamma_{X,Y}$ *ili* γ_{C} *data sa*

$$\gamma_{X,Y} = \gamma_{\mathrm{C}} = Q(\mathrm{C},\mathrm{M}) + Q(\mathrm{C},\mathrm{W}). \tag{3.5.3}$$

<u>*Dokaz:*</u> Znamo da je $\mathrm{M} = \min(u,v)$, međutim, primetimo da to možemo da napišemo i na sledeći način $\mathrm{M} = \frac{1}{2}(u+v-|u-v|)$. Izračunajmo prvo koliko nam je $Q(\mathrm{C},\mathrm{M})$. Koristeći definiciju funkcije saglasnosti imamo da je

$$\begin{aligned} Q(\mathrm{C},\mathrm{M}) &= 4\iint_{\mathbf{I}^2}\mathrm{M}(u,v)d\,\mathrm{C}(u,v)-1, \\ &= 2\iint_{\mathbf{I}^2}(u+v-|u-v|)dC(u,v)-1. \end{aligned}$$

Znamo da je svaka kopula zajednička funkcija raspodele sa uniformnim (0,1) marginama pa će važiti

$$\iint_{\mathbf{I}^2} u\,dC(u,v) = \iint_{\mathbf{I}^2} v\,dC(u,v) = \frac{1}{2},$$

pa odatle zaključujemo da je

$$Q(\mathrm{C},\mathrm{M}) = 1 - 2\iint_{\mathbf{I}^2}|u-v|\,dC(u,v).$$

Na sličan način pošto je $\mathrm{W}(u,v)=\max(u+v-1,0)$, što možemo napisati na drugačiji način kao $\mathrm{W}(u,v)=\frac{1}{2}[u+v-1+|u+v-1|]$, imamo

$$Q(\mathrm{C},\mathrm{W})=4\iint_{\mathbf{I}^2}\mathrm{W}(u,v)d\,\mathrm{C}(u,v)-1.$$
$$=2\iint_{\mathbf{I}^2}(u+v-1-|u+v-1|)dC(u,v)-1,$$
$$=2\iint_{\mathbf{I}^2}|u+v-1|\,dC(u,v)-1.$$

Sada ćemo izračunati zbir $Q(\mathrm{C},\mathrm{M})+Q(\mathrm{C},\mathrm{W})=$
$=1-2\iint_{\mathbf{I}^2}|u-v|\,dC(u,v)+2\iint_{\mathbf{I}^2}|u+v-1|\,dC(u,v)-1$,
$=2\iint_{\mathbf{I}^2}(|u+v-1|-|u-v|)d\,\mathrm{C}(u,v)=\gamma$.

■

Posledica: Pod uslovima prethodne teoreme važi sledeće

$$\gamma_{\mathrm{C}}=4\left[\int_0^1\mathrm{C}(u,1-u)du-\int_0^1[u-\mathrm{C}(u.u)]du\right]. \tag{3.5.4}$$

<u>*Dokaz:*</u> $\gamma_{\mathrm{C}}=Q(\mathrm{C},\mathrm{M})+Q(\mathrm{C},\mathrm{W})$, gde su

$$Q(\mathrm{C},\mathrm{M})=4\iint_{\mathbf{I}^2}\mathrm{C}(u,v)d\,\mathrm{M}(u,v)-1=4\int_0^1\mathrm{C}(u,u)du-1$$

i

$$Q(\mathrm{C},\mathrm{W})=4\iint_{\mathbf{I}^2}\mathrm{C}(u,v)d\,\mathrm{W}(u,v)-1=4\int_0^1\mathrm{C}(u,1-u)du-1.$$

Kada vratimo to u jednačinu za izračinavanje Đinijevog gama dobijamo upravo rezultat o kome govori ova posledica.

■

Za kraj ovog odeljka o merama saglasnosti, navedimo još jedan koeficijent koji možemo svrstati u ovu grupu, a to je Blomkvistovo (Blomqvist [2]) beta. Pretpostavimo da u izrazu funkcije Q koristimo slučajan vektor (X,Y) i fiksiranu tačku $(x_0,y_0)\in\mathbf{R}^2$. Tada funkcija Q će izgledati ovako

$$Q=\mathrm{P}\{(X-x_0)(Y-y_0)>0\}-\mathrm{P}\{(X-x_0)(Y-y_0)<0\}.$$

Blomkvist je proučavao meru saglasnosti koristeći medijane populacije. Često se Blomkvistovo beta naziva i medijalan koeficijent korelacije. Neka su $\tilde{x}$ i $\tilde{y}$ medijane od X i Y, tada Blomkvistovo beta je izraženo kao

$$\beta=\beta_{X,Y}=\mathrm{P}\{(X-\tilde{x})(Y-\tilde{y})>0\}-\mathrm{P}\{(X-\tilde{x})(Y-\tilde{y})<0\}. \tag{3.5.5}$$

U slučaju da su X i Y neprekidne sa zajedničkom funkcijom raspodele H i marginama redom F i G, čija je kopula C, važiće da je $\mathrm{F}(\tilde{x})=\mathrm{G}(\tilde{y})=\frac{1}{2}$, pa će Blomkvistovo beta biti

$$\beta=2\mathrm{P}\{(X-\tilde{x})(Y-\tilde{y})>0\}-1,$$

$$= 2[\mathrm{P}\{X < \widetilde{x}, Y < \widetilde{y}\} + \mathrm{P}\{X > \widetilde{x}, Y > \widetilde{y}\}] - 1,$$
$$= 2[\mathrm{H}(\widetilde{x}, \widetilde{y}) + (1 - \mathrm{F}(\widetilde{x}) - \mathrm{G}(\widetilde{y}) + \mathrm{H}(\widetilde{x}, \widetilde{y}))] - 1,$$
$$= 4\mathrm{H}(\widetilde{x}, \widetilde{y}) - 1.$$

Kako je $\mathrm{F}(\widetilde{x}) = \mathrm{G}(\widetilde{y}) = \frac{1}{2}$, tada je $\mathrm{H}(\widetilde{x}, \widetilde{y}) = \mathrm{C}(1/2, 1/2)$, odnosno

$$\beta = \beta_{\mathrm{C}} = 4\mathrm{C}\left(\frac{1}{2}, \frac{1}{2}\right) - 1. \tag{3.5.6}$$

Priču o merama saglasnosti završavamo teoremom koja je ekvivalentna Teoremi 3.3.2 samo što ovde imamo druge koeficijente.

Teorema 3.5.2: *Ako su X i Y neprekidne slučajne veličine sa kopulom* C*, tada Đinijevo gama i Blomkvistovo beta računati na celoj populaciji zadovoljavaju uslove izDefinicije 3.1.3 i Teoreme 3.1.4 za mere saglasnosti.*

3.7 Zavisnost, svojstva i mere zavisnosti

Jedno od svojstava zavisnosti sa kojima se najčešće možemo sresti je tzv. nedostatak zavisnosti, poznatije kao svojstvo nezavisnosti. Na primer ako su X i Y neprekidne slučajne veličine sa zajedničkom funkcijom raspodele H, kažemo da su te dve slučajne veličine nezavisne, ako se njihova funkcija raspodele može napisati kao proizvod svojih marginalnih funkcija. Prisetimo se Teoreme 2.3.2 u kojoj smo uveli pojam nezavisnosti preko kopule proizvoda Π. Takođe, pomenuli smo u odeljku 2.4 da Freše-Hefdingove granice možemo povezati sa pojmom zavisnosti. Tačnije rekli smo da će donja (gornja) Freše-Hefdingova granica opisivati negativnu (pozitivnu) savršenu zavisnost. Ali, sta je to ustvari savršena zavisnost? Osobina pozitivne zavisnosti je da se velike vrednosti neke slučajne veličine javljaju zajedno, a osobina negativne zavisnosti je da se velike vrednosti jedne slučajne veličine javljaju sa malim vrednostima neke druge slučajne veličine. Sledeću definiciju koja govori o pozitivnoj i negativnoj zavisnosti uveo je Leman (Lehmann [12]).

Definicija 3.6.1: Neka su X i Y slučajne veličine. Kažemo da su X i Y pozitivno zavisne ako za svako $(x, y) \in \mathbf{R}^2$ važi

$$\mathrm{P}\{X \le x, Y \le y\} \ge \mathrm{P}\{X \le x\}\mathrm{P}\{Y \le y\}.$$

Ekvivalentno X i Y su negativno zavisne ako za svako $(x, y) \in \mathbf{R}^2$ važi

$$\mathrm{P}\{X > x, Y > y\} \ge \mathrm{P}\{X > x\}\mathrm{P}\{Y > y\}.$$

Na osnovu ove definicije i Freše-Hefdingove nejednakosti $W(u,v) \le C(u,v) \le M(u,v)$, zaključili smo da donja i gornja Freše-Hefdingova granica opisuju negativnu, odnosno pozitivnu zavisnost.

U prethodnim odeljcima govorili smo o merama saglasnosti, koje predstavljaju na neki način i sami mere zavisnosti. Videćemo u definiciji mera zavisnosti da su osobine koje one moraju da ispunjavaju ustvari sve osobine mera saglasnosti, jedina razlika biće u četvrtoj osobini Definicije 3.1.3. Ta četvrta osobina kaže da ako su slučajne veličine nezavisne ona je mera saglasnosti jednaka nuli, međutim obrnuto ne važi. Odnosno, može da se desi da mera saglasnosti bude jednaka nuli, a da slučajne veličine budu zavisne. Mere zavisnosti u određenom smislu nam govore u kakvoj su vezi X i Y sa ekstremima nezavisnosti i zavisnosti, dok su nam kod mera saglasnosti ekstremi bili donja i gornja Freše-Hefdingova granica. Definišimo konačno mere zavisnosti.

Definicija 3.6.2: Neka su X i Y dve neprekidne slučajne veličine sa kopulom C. Kažemo da je mera δ (često pišemo $\delta_{X,Y}$ ili δ_{C}) mera zavisnosti ako ispunjava sledeće osobine:

1. $\delta_{X,Y}$ je definisana za svaki par neprekidnih slučajnih veličina X,Y;
2. $0 \le \delta_{X,Y} \le 1$;
3. $\delta_{X,Y} = \delta_{Y,X}$;
4. $\delta_{X,Y} = 0$ ako i samo ako su X i Y nezavisne;
5. $\delta_{X,Y} = 1$ ako i samo ako svaki od X i Y je skoro sigurno strogo monotona funkcija od drugog;
6. ako su α i β skoro sigurno strogo monotone funkcije na kodomenima slučajnih veličina X i Y redom, tada $\delta_{\alpha(X)\beta(Y)} = \delta_{X,Y}$;
7. ako je $\{(X_n, Y_n)\}$ niz neprekidnih slučajnih veličina sa kopulama C_n i ako $\{C_n\}$ konvergira ka C, tada $\lim_{n\to\infty} \delta_{C_n} = \delta_{C}$.

3.7 Švajcer-Volfovo sigma

Podsetimo se Spirmanovog ro i načina na koji on može biti zapisan, tačnije, jednakosti (3.3.4). Prvu meru zavisnosti koju uvodimo usko je povezana sa ovim zapisom Spirmanovog ro-a. Odnosno, neka imamo

$$\rho_{X,Y} = \rho_{\mathrm{C}} = 12\iint_{\mathrm{I}^2}(\mathrm{C}(u,v)-uv)dudv .$$

Kada u ovoj jednakosti umesto običnih zagrada stavimo apsolutnu vrednost razlike, tj. $|\mathrm{C}(u,v)-uv|$, tada imamo meru koja predstavlja rastojanje između grafika kopula C i Π. Ta mera je poznatija pod imenom Švajcer-Volfovo sigma i oblika je

$$\sigma_{X,Y} = \sigma_{\mathrm{C}} = 12\iint_{\mathrm{I}^2}|\mathrm{C}(u,v)-uv|\, dudv . \tag{3.7.1}$$

Pokažimo da je to zaista jedna mera zavisnosti, odnosno pokažimo da važi sledeća teorema.

Teorema 3.7.1: *Neka su X i Y dve neprekidne slučajne veličine sa kopulom* C. *Tada mera* σ_{C} *definisana u (3.7.1) je mera zavisnosti, odnosno zadovoljava uslove Definicije 3.6.2.*

Dokaz: Prva i treća osobina slede intuitivno iz definicije Švajcer-Volfovog sigma. Drugu osobinu je lako dokazati, potrebno je prvo pokazati da za bilo koju kopulu C važi $\iint_{\mathrm{I}^2}|\mathrm{C}(u,v)-uv|\, dudv \le \frac{1}{12}$. Što se tiče četvrte osobine ona sledi direktno iz Teoreme 2.3.2, dok peta osbina sledi iz činjenice da će važiti jednakost (3.7.1) ako i samo ako je C jednako M ili W. Ako je α strogo rastuća funkcija, a β strogo opadajuća tada možemo da iskoristimo Teoremu 2.3.4 i činjenicu da je $\mathrm{C}_{\alpha(X),\beta(Y)}(u,v)-\Pi(u,v)=\Pi(1-u,v)-\mathrm{C}_{X,Y}(1-u,v)$. U ostalim slučajevima za funkcije α i β dobijamo slične rezultate. Sedmo svojstvo sledi iz Lipšicovog kriterijuma, pa niz $\{\mathrm{C}_n\}$ konvergira uniformno ka C.

■

Primetimo da što se tiče povezanosti Spirmanovog ro-a i Švajcer-Volfovog sigma možemo zaključiti da ako su *X* i *Y* pozitivno zavisne slučajne veličine tada je $\sigma_{\mathrm{C}} = \rho_{\mathrm{C}}$, a ako su negativno zavisne onda je $\sigma_{\mathrm{C}} = -\rho_{\mathrm{C}}$. U slučaju kada *X* i *Y* nisu ni pozitivno ni negativno zavisne, često je Švajcer-Volfovog sigma bolja mera od ρ.

3.8 Hefdingov indeks zavisnosti

Švajcer i Volf su primetili da bilo koja normalizovana mera rastojanja između kopule C i Π, treba da rezultuje simetričnom neparametarskom merom zavisnosti. Uvedimo meru rastojanja koju ćemo obeležavati sa L_p. Za bilo koje p takvo da $1 \leq p < \infty$, mera rastojanja L_p između kopula C i Π data je sa

$$L_p = \left(k_p \iint_{\mathbf{I}^2} |\mathrm{C}(u,v) - uv|^p \, dudv\right)^{1/p}, \tag{3.8.1}$$

gde je ovo k_p konstanta koju biramo tako da jednakost (3.8.1) bude jednaka 1 ako je C=M ili W. Može se pokazati proveravajući osobine Definicije 3.6.2 da će mera data u obliku jednakosti (3.8.1) biti mera zavisnosti. Pogledati [16].

Mi ćemo razmatrati slučaj kada je p=2. Tada dobijamo sledeću meru

$$\Phi_{X,Y} = \Phi_{\mathrm{C}} = \left(90 \iint_{\mathbf{I}^2} |\mathrm{C}(u,v) - uv|^2\right)^{1/2}. \tag{3.8.2}$$

Kada kvadriramo ovu meru zavisnosti dobijemo indeks koji je poznat kao Hefdingov indeks zavisnosti i njega obeležavamo sa $\Phi_{X,Y}^2$.

3.9 Repna zavisnost

Mnoge mere zavisnosti o kojima smo do sada pričali opisuju kako se velike (male) vrednosti jedne slučajne veličine javljaju sa velikim (malim) vrednostima neke druge slučajne veličine. Sada ćemo govoriti o tzv. repnoj zavisnosti koja meri zavisnost između slučajnih veličina koje se nalaze u gornjem desnom i donjem levom kvadarantu u $\mathbf{I}^2$. Prvo ćemo definisati koeficijente repne zavisnosti.

Definicija 3.9.1: Neka su X i Y dve neprekidne slučajne veličine sa funkcijama raspodele redom F i G i neka je data njihova kopula C.

1.Neka je $\lambda_{\mathrm{L}} = \lim_{u \downarrow 0} \mathrm{P}\{\mathrm{F}(X) < u \mid \mathrm{G}(Y) < u\}$. (3.9.1)

Kažemo da je λ_{L} *donji repni koeficijent zavisnosti* ako i samo ako je $\lambda_{\mathrm{L}} \neq 0$.

2. Neka je $\lambda_{\mathrm{U}} = \lim_{u \uparrow 1} \mathrm{P}\{\mathrm{F}(X) > u \mid \mathrm{G}(Y) > u\}$. (3.9.2)

Kažemo da je λ_{U} *gornji repni koeficijent zavisnosti* ako i samo ako je $\lambda_{\mathrm{U}} \neq 0$.

(L je kao lower odnosno donji, a U kao upper tj. gornji).

Ukoliko je donji (gornji) koeficijent baš jednak nuli tada kažemo da kopula C nema donju (gornju) repnu zavisnost.

Teorema 3.9.2: *Neka su X, Y,* F, G, λ_{U} i λ_{L} *dati kao u Definiciji* 3.9.1 *i neka je* C *kopula od X i Y. Tada važi*

$$\lambda_{\mathrm{L}} = \lim_{u\downarrow 0} \mathrm{P}\{\mathrm{F}(X) < u \mid \mathrm{G}(Y) < u\} = \lim_{u\to 0^+} \frac{\mathrm{C}(u,u)}{u}$$

i

$$\lambda_{\mathrm{U}} = \lim_{u\uparrow 1} \mathrm{P}\{\mathrm{F}(X) > u \mid \mathrm{G}(Y) > u\} = 2 - \lim_{u\to 1^-} \frac{1-\mathrm{C}(u,u)}{u}.$$

Dokaz: Može se naći u [16].

Primer 3.5: Posmatrajmo Gumbelovu familiju kopula i neka nam je iz te familije data kopula koja je sledećeg oblika

$$\mathrm{C}_\theta(u,v) = \exp\{-((-\ln u)^\theta + (-\ln v)^\theta)^{1/\theta}\}, \theta \geq 1.$$

Izračunaćemo gornji repni koeficijent.

$$\frac{1-2u+\mathrm{C}(u,u)}{1-u} = \frac{1-2u+\exp\{2^{1/\theta}\ln u\}}{1-u} = \frac{1-2u+u^{2^{1/\theta}}}{1-u},$$

$$\lambda_{\mathrm{U}} = \lim_{u\uparrow 1} \frac{1-2u+\mathrm{C}(u,u)}{1-u} = \lim_{u\uparrow 1} 2^{1/\theta} u^{2^{1/\theta}-1} = 2-2^{1/\theta}.$$

Za $\theta > 1$ kopula ima gornju repnu zavisnost.

□

Lema 3.9.3: *Označimo sa* $\overline{\lambda}_{\mathrm{L}}$ *i* $\overline{\lambda}_{\mathrm{U}}$ *repne koeficijente zavisnosti kopule preživljavanja. Tada će važiti:*

$$\lambda_{\mathrm{L}} = \overline{\lambda}_{\mathrm{U}};$$
$$\lambda_{\mathrm{U}} = \overline{\lambda}_{\mathrm{L}}.$$

Dokaz:

$$\lambda_{\mathrm{L}} = \lim_{u\downarrow 0} \frac{\mathrm{C}(u,u)}{u} = \lim_{v\uparrow 1} \frac{\mathrm{C}(1-v,1-v)}{1-v} = \lim_{v\uparrow 1} \frac{1-2v+\hat{\mathrm{C}}(v,v)}{1-v} = \overline{\lambda}_{\mathrm{U}}.$$

$$\lambda_{\mathrm{U}} = \lim_{u\uparrow 1} \frac{1-2u+\mathrm{C}(u,u)}{1-u} = \lim_{v\downarrow 0} \frac{2v-1+\mathrm{C}(1-v,1-v)}{v} = \lim_{v\downarrow 0} \frac{\hat{\mathrm{C}}(v,v)}{v} = \overline{\lambda}_{\mathrm{L}}.$$

■

3.10 Empirijske kopule

U dosadašnjoj priči smo uglavnom koristili mere saglasnosti i mere zavisnosti koje su bazirane na čitavoj populaciji. Sada ćemo izraziti kopule u uzoračkoj verziji u terminima poznatijim kao empirijska kopula kao i empirijska funkcija frekvencije.

Definicija 3.10.1: Neka je dat uzorak obima n $\{(x_1,y_1),(x_2,y_2),...,(x_n,y_n)\}$ neke dvodimenzionalne neprekidne slučajne veličine. *Empirijska kopula* je funkcija C_n data sa

$$\mathrm{C}_n\left(\frac{i}{n},\frac{j}{n}\right)=\frac{\text{broj parova (x, y) takvih da } x\le x_{(i)},\, y\le y_{(j)}}{n},$$

gde su $x_{(i)}$ i $y_{(j)}$ statistike poretka iz uzorka. *Empirijska kopula frekvencije* je c_n data sa

$$\mathrm{c}_n\left(\frac{i}{n},\frac{j}{n}\right)=\begin{cases}\frac{1}{n}, & \text{ako } (x_{(i)},y_{(j)}) \text{ je element iz uzorka,}\\ 0, & \text{ostalo.}\end{cases}$$

Možemo da primetimo da su empirijska kopula i empirijska kopula frekvencije c_n povezane na sledeći način:

$$\mathrm{C}_n\left(\frac{i}{n},\frac{j}{n}\right)=\sum_{p=1}^{i}\sum_{q=}^{j}\mathrm{c}_n\left(\frac{p}{n},\frac{q}{n}\right).$$

U uvodu smo pomenuli da je prvi empirijske kopule proučavao Deheuvels i on ih je nazivao *empirijske funkcije zavisnosti*. Navešćemo sada jednu teoremu koja nam govori o povezanosti empirijskih kopula i mera kao što su Kendalovo tau, Spirmanovo ro i Đinijevo gama. Dokaz te teoreme nećemo navoditi, a on se može naći u [16].

Teorema 3.10.2: *Neka su date empirijska kopula* C_n *i empirijska kopula frekvencije* c_n. *Ako sa* r, t *i* g *obeležimo uzoračke vrednosti Spirmanovog ro-a, Kendalovom tau i Đinijevog gama redom, tada*

$$\mathrm{r}=\frac{12}{n^2-1}\sum_{i=1}^{n}\sum_{j=1}^{n}\mathrm{C}_n\left(\frac{i}{n},\frac{j}{n}\right)-\frac{i}{n}\cdot\frac{j}{n},$$

$$\mathrm{t}=\frac{2n}{n-1}\sum_{i=2}^{n}\sum_{j=2}^{n}\sum_{p=1}^{i-1}\sum_{q=1}^{j-1}\left[\mathrm{c}_n\left(\frac{i}{n},\frac{j}{n}\right)\mathrm{c}_n\left(\frac{p}{n},\frac{q}{n}\right)-\mathrm{c}_n\left(\frac{i}{n},\frac{q}{n}\right)\mathrm{c}_n\left(\frac{p}{n},\frac{j}{n}\right)\right],$$

$$\mathrm{g}=\frac{2n}{n^2/2}\left\{\sum_{i=1}^{n-1}\mathrm{C}_n\left(\frac{i}{n},1-\frac{i}{n}\right)-\sum_{i=1}^{n}\left[\frac{i}{n}-\mathrm{C}_n\left(\frac{i}{n},\frac{i}{n}\right)\right]\right\}.$$

3.11 Zavisnost više promenljivih

Skoro sve mere saglasnosti i zavisnosti, koje smo do sada opisivali imaju svoja proširenja do većih dimenzija. Naravno u tim višedimenzionim slučajevima priča se prilično komplikuje, teže je izračunati mere koje su nam potrebne i slično, ali bez obzira na to sve osobine zavisnosti koje smo obrađivali do sada mogu se preneti i u slučaju kada imamo veći broj promenljivih. Definišimo sada generalizovane vrednosti najosnovnijih mera kao što su Kendalovo tau, Spirmanovo ro i Švajcer-Volfovo sigma.

Neka su $\mathbf{X}=(X_1, X_2,..., X_m)$ i $\mathbf{Y}=(Y_1, Y_2,..., Y_m)$ slučajni vektori sa kopulom C i definišimo slučajnu promenljivu $D_j = X_j - Y_j$. Obeležimo sa $\mathrm{B}_{k,m-k}$ skup *m*-torki iz $\mathbf{R}^m$ koji sadrži *k* pozitivnih i *m-k* negativnih komponenti. Tada generalizovana verzija Kendalovog tau data je sa

$$\tau_{\mathrm{C}} = \sum_{k=\frac{m+1}{2}}^{m} w_k \,\mathrm{P}\{(D_1,..., D_m) \in \mathrm{B}_{k,m-k}\},$$

gde su w_k takve da za $\frac{m+1}{2} \leq k \leq m$ važi

(1) $\tau_{\mathrm{C}} = 1$, ako je C=W,

(2) $\tau_{\mathrm{C}} = 0$, ako je C=M,

(3) $\tau_{\mathrm{C}_1} < \tau_{\mathrm{C}_2}$ kad god je $\mathrm{C}_1 \prec \mathrm{C}_2$.

Implikacije prve i druge osobine za w_k slede prirodno , tj.

(1) $w_m = 1$,

(2) $\sum_{k=0}^{m} w_k \binom{m}{k} = 0$, $w_k = w_{m-k}$ za $k < \frac{m+1}{2}$,

dok je treća osobina komplikovanija za objašnjavanje, ali može se detaljnije naći u Joe 1996.

Generalizacija Spirmanovog ro-a za funkciju raspodele *m* promenljivih koje imaju kopulu C data je sa

$$\rho_{\mathrm{C}} = \left(\int\limits_{\mathbf{I}^m} \!\!...\!\! \int \mathrm{C}(u) du_1 ... du_m - 2^m \right) \Big/ \left((m+1)^{-1} + 2^{-m} \right).$$

Što se tiče mera zavisnosti pomenućemo jedino generalizaciju Švajcer-Volfovog sigma, koje je oblika

$$\sigma_{\mathrm{C}} = \frac{2^m (m+1)}{2^m - (m+1)} \int \dots \int_{\mathbf{I}^m} | \mathrm{C}(u_1, \dots, u_m) - u_1 \cdots u_m | du_1 \dots du_m ,$$

gde je C *m*-dimenziona kopula.

4 Zaključak

Priča o kopulama bi mogla da se nastavi u nedogled, jer je to jedno od polja statistike, matematike, ekonomije koje se i dalje istražuje i razvija. Svakim daljim istraživanjem pojma kopula dolazi se do vrlo interesantnih saznanja o njihovim osobinama i primenama. U ovom radu prikazana su osnovna svojstva kopula, povezanost kopula sa slučajnim veličinama i zavisnosti slučajnih veličina pomoću kopula. Naravno, kao što sam rekla teorija kopula je veoma širok pojam i moglo je ovde da bude reči o još mnogo toga, ali ja sam se generalno koncentrisala na ovo što je i objašnjeno u radu. Osobama koje budu čitale ovaj rad toplo preporučujem ukoliko žele da saznaju nešto više o kopulama od ovoga što je prikazano u mom radu da pogledaju i takozvane Arhimedove kopule, zatim Gausovu kopulu, t-kopulu kao i mnoge druge familije kopula koje su par puta i pomenute u nekim primerima koje sam navodila. Sve te kopule imaju rasprostranjene primene u finansijama, aktuarstvu, hidrologiji, medicinskim naukama, informacionim tehnologijama, menadžmentu i tako dalje. Čitav spisak dosadašnjih primena kopula može se naći u [1]. Pored svih kvaliteta koje kopule poseduju, naravno, pojavile su se i određene kritike (pogledati [14]). Uglavnom su se te kritike svodile na sledeće: ne vidi se smisao u tome da dajemo prednost njima umesto da uzmemo bilo koju višedimenzionu raspodelu koja je pogodna za problem koji se rešava i onda da primenimo klasične statističke metode. Često se kopule koriste jer su pogodne matematički, a da se pritom ne obrati pažnja na to da li primena te kopule ima smisla. Jedna bitna stvar kod kopula je ta što nije pogodna za modeliranje zavisnosti kroz vreme, tj. nisu pogodne kada imamo stohastičke procese i vremenske serije. Sam pristup kod kopula svodi se na to da neki veoma komplikovan problem rešimo tako što nađemo dvodimenzionalnu ili višedimenzionalnu raspodelu i računamo marginalne raspodele i kopulu odvojeno. To je ustvari i najbitnija stvar kod kopula, jer smo odvojili zavisnost od marginalnih raspodela. Ponekad je jako teško naći odgovarajuću kopulu shodno problemu koji rešavamo pa se često koriste parametarske familije kopula gde se podešavaju paramatri, a podešavanjem parametara opet može da dođe do određenih nedostataka čime se stvaraju određena odstupanja. Međutim, pored svega toga kopule igraju veoma važnu ulogu danas, široko su rasprostranjene o čemu govori i činjenica da se i dalje intenzivno istražuju.

Reference

[1] Balakrishnan N., Chin Diew Lai, *Continuous Bivariate Distributions*, Springer (2009.).

[2] Blomqvist N.: *On a measure of dependence between two random variables*. Ann Math Statist 21:593-600 (1950.).

[3] Deheuvels P.: *La fonction de dependence empirique et ses propietes. On test non parametrique d'independence*, Acad Roy Belg Bul Cl Sci (5) 65:274-292 (1979.).

[4] Embrechts P., Lindskog F., McNeil A.: *Modelling Dependence with Copulas and Applications to Risk Management*, In: Rachev's (ed) Handbook of Heavy Tailed Distributions in Finance. Elsevier, New York, pp. 329-384.

[5] Embrechts, P., McNeil A., and Straumann D.: *Correlation and Dependence in Risk Management: Properties and Pitfalls*, To appear in *Risk Management: Value at Risk and Beyond*, ed. byM.Dempster and H.K.Moffatt, Cambridge University Press (2001).

[6] Fisher: *Copulas*, Encyclopedia of Statistical Sciences 159-163 (1997.).

[7] Hoeffding W (1940.): *Masstabinvariante Korrelationstheorie*: Schriften des Matematischen Instituts und des Instituts fuer Angewandte Mathematik der Universitaet Berlin 5 Heft 3:179-233 (Reprinted as Scale-invariant correlation theory. In: Fisher NI, Sen PK (eds) The Collected Works of Wassily Hoeffding. Springer, New York pp 57-107).

[8] Hoeffding W (1941) *Masstabinvariante Korrelationsmasse fuer diskontinuierliche Verteilungen.* Arkiv fuer matematischen Wirtschaften und Sozialforschung 7:49-70 [Reprinted as Scale-invariant correlation measures for discontinuous distributions. In: Fisher NI, Sen PK (eds) The Collected Works of Wassily Hoeffding. Springer, New York pp 109-133].

[9] Joe H.: *Multivariate Models and Dependence Concepts*, Chapman & Hall, London (1997.).

[10] Joe H.: *Multivariate Concordance,* Journal of Multivariate Analysis (1990.).

[11] Jan de Kart, 2007., *Modelling Tail Dependence using Copula.*

[12] Lehmann EL: *Some concepts of dependence.* Ann Math Statist 37:1137-1153 (1966).

[13] Li X, Mikusinski P, Sherwood H, Taylor MD: *Some integration-by-parts formulas involving 2-copulas*, In: Cuadras CM, Fortiana J, Rodrguez Lallena JA (eds) Distributions with Given Marginals and Statistical Modelling, Kluwer, Dordrecht, pp 153-159 (2002).

[14] Mikosch, T.: *Copulas: Tales and facts*, Extremes **9**, 3–20 (2006).

[15] Marshall A.W.: *Copulas, Marginals and joint Distribution*, In: Ruechendorf L, Schweizer B, Taylor MD (eds) Distributions with Fixed Marginals and Related Topics. Institute of Mathematical Statistics, Hayward, CA, pp 213-222 (1996.).

[16] Nelsen Roger B.: *An Introduction to Copulas*, Springer (2007.).

[17] Renyi A: *On measures of dependence*. Acta Math Acad Sci Hungar 10:441-451 (1959.).

[18] Schweizer B., Wollf E.: *On nonparametric measures of dependence for random variables,* Annals of Statistics 9, 879-885 (1981.).

[19] Schweizer B., Sklar A.: *Probabilistics Metric Spaces*, North-Holland, New York (1983.).

[20] Sklar A.: *Fonctions de repartition dimensions et leurs marges*, Publ Inst Statist Univ Paris 8:229-231 (1959.).

[21] Seeley: *Fubini implies Leibniz implies* $F_{x,y} = F_{y,x}$, Am math Monthly 68: 56-57 (1961.).

[22] www.wikipedia.com