UNIVERZA V LJUBLJANI
FAKULTETA ZA MATEMATIKO IN FIZIKO

Finančna matematika – 1. stopnja

Jaša Štefan

Kopule za končne porazdelitve

Delo diplomskega seminarja

Mentorica: prof. dr. Damjana Kokol Bukovšek

Somentor: asist. dr. Blaž Mojškerc

Ljubljana, 2018

Kazalo

Kopule za končne porazdelitve

POVZETEK

Pri obravnavanju odvisnosti med slučajnimi spremenljivkami pogosto uporabljamo funkcije s posebnimi lastnostmi, ki jim pravimo kopule. Za njih velja, da povezujejo porazdelitvene funkcije posameznih slučajnih spremenljivk z njihovo skupno porazdelitveno funkcijo. Kako so te povezane med sabo, nam pove Sklarov izrek, ki predstavlja teoretično podlago za uporabo kopul v praksi. V statistiki so pripravne, ker lahko s pomočjo kopule porazdelitev slučajnega vektorja ocenimo z ocenjenimi porazdelitvenimi funkcijami slučajnih spremenljivk.

Pri ugotavljanju odvisnosti med slučajnimi spremenljivkami si lahko pomagamo tudi z merami odvisnosti. Med njimi sta najpogosteje uporabljeni Kendallov τ in Spearmanov ρ, ki ustrezata Scarsinijevi definiciji in sta posledično meri skladnosti.

V primeru zveznih slučajnih spremenljivk je kopula za določeno skupno porazdelitveno funkcijo enolična. V diskretnem primeru pa temu ni več tako. Posledično naletimo na težave, saj se veliko lastnosti ne prenese v diskretni primer. Izkaže se, da teorija kopul v diskretnem primeru ni neuporabna, je pa potrebno biti previden pri njihovi uporabi.

Copulas on count data

ABSTRACT

When observing dependence between random variables we often use functions that have special characteristics, called copulas. They link distribution functions of one variable with their joint distribution function. Sklar's theorem tells us exactly how they are linked together, and this presents theoretical background for using copulas in practice. They are helpful in statistics: we can use them to estimate distribution of random vector by estimating distribution functions of random variables.

When measuring dependence between random variables, it can be helpful to use measures of dependence. Two of the most widely used are Kendall's τ and Spearman's ρ, and because they meet the criteria of Scarsini's definition, they are both measures of concordance.

In case of continuous random variables, copula linked with joint distribution functions is unique. However in discrete case that is no longer true. Therefore we run into troubles, as a lot of characteristics do not translate from continuous to discrete case. Nevertheless, copula theory can be useful in discrete case as well, but it has to be used with caution.

Math. Subj. Class. (2010): 62H05, 60E05

Ključne besede: kopula, podkopula, Sklarov izrek, Carleyine meje, skladnost

Keywords: copula, subcopula, Sklar's theorem, Carley's bounds, concordance

1. Uvod

Kopule so funkcije, ki povezujejo porazdelitvene funkcije slučajnega vektorja s porazdelitvenimi funkcijami posameznih slučajnih spremenljivk. Raziskovanje njihovih lastnosti se je začelo razmeroma pozno, intenzivneje šele pred približno petindvajsetimi leti, od takrat pa se je zanimanje zanje močno povečalo. Izkazalo se je namreč, da so zelo uporabne za opazovanje odvisnosti med spremenljivkami, ko so te zvezno porazdeljenene. Vedno več jih uporabljajo tudi v financah, denimo za določanje cen izvedenim finančnim instrumentom ali za modeliranje tveganja portfelja (angl. *portfolio risk*).

V članku, na katerem temelji delo diplomskega seminarja, skušata avtorja pokazati, kdaj lahko teorijo kopul iz zvezno porazdeljenih slučajnih spremenljivk uporabimo tudi pri diskretno porazdeljenih slučajnih spremenljivkah in na kaj moramo biti pri tem pazljivi. Avtorja sta se omejila na kopule dveh spremenljivk, a se da večino ugotovitev posplošiti na več spremenljivk.

V drugem poglavju dela diplomskega seminarja so opisane nekatere lastnosti kopul v primeru zveznih funkcij, nato pa še težave, ki nastopijo v diskretnem primeru. Glavni vir težav je, da kopul ni več mogoče enolično predstaviti. Posledično veliko lastnosti, ki veljajo za kopule v zveznem primeru, v diskretnem ne drži več. Če je množica kopul, ki ustrezajo enačbi v Sklarovem izreku, dovolj majhna, so te težave dokaj obvladljive. Težava je, da se izkaže, da je ta razred v splošnem razmeroma velik.

2. Kopule za zvezne slučajne spremenljivke

Poglavje povzeto po knjigi R. Nelsen: An Introduction to Copulas, [3].

2.1. Uvodne oznake, definicije in leme.

Tekom naloge uporabljamo naslednje oznake. Z $\mathbb{R}$ označimo množico realnih števil, tj. interval $(-\infty, \infty)$, z $\overline{\mathbb{R}}$ množico $[-\infty, \infty]$ in z $\overline{\mathbb{R}}^2$ množico $\overline{\mathbb{R}} \times \overline{\mathbb{R}}$. Naj $\mathbf{I}$ označuje interval $[0,1]$, $\mathbf{I}^2$ pa kvadrat $[0,1] \times [0,1]$. Z Z_F označimo zalogo vrednosti funkcije F in z D_F njeno definicijsko območje.

Denimo, da imamo slučajni spremenljivki X in Y s porazdelitvenima funkcijama $F(x) = P(X \leq x)$ in $G(y) = P(Y \leq y)$ ter skupno porazdelitveno funkcijo slučajnega vektorja (X, Y), $H(x, y) = P(X \leq x, Y \leq y)$. Potem lahko vsakemu paru števil (x, y) priredimo tri števila, $F(x)$, $G(y)$ in $H(x, y)$, pri čemer seveda vsa ležijo na intervalu $\mathbf{I}$. Kopula je funkcija, ki povezuje funkciji F in G s funkcijo H. Preden lahko zapišemo definicijo kopule, potrebujemo še naslednjo definicijo.

Naj bo f funkcija, ki je definirana na definicijskem območju D_f. Funkcija f je na D_f naraščajoča, če iz $x_2 > x_1$ sledi $f(x_2) \geq f(x_1)$ in strogo naraščajoča, če iz $x_2 > x_1$ sledi $f(x_2) > f(x_1)$. Podobno je funkcija f na D_f padajoča, če iz $x_2 > x_1$ sledi $f(x_2) \leq f(x_1)$ in strogo padajoča, če iz $x_2 > x_1$ sledi $f(x_2) < f(x_1)$.

Definicija 2.1. Naj bosta S_1 in S_2 neprazni podmnožici $\overline{\mathbb{R}}$ in naj bo H realna funkcija z definicijskim območjem $S_1 \times S_2$. Naj bo B pravokotnik $[x_1, x_2] \times [y_1, y_2]$, kjer je $x_1 < x_2$, $y_1 < y_2$, pri čemer vsa njegova oglišča ležijo v D_H. *H-ploščino* pravokotnika B označimo z

$$V_H(B) = H(x_2, y_2) - H(x_1, y_2) - H(x_2, y_1) + H(x_1, y_1).$$

Funkcija H je *2-naraščajoča*, če je $V_H(B) \geq 0$ za vse pravokotnike B z oglišči v D_H.

Če je funkcija 2-naraščajoča, ni nujno naraščajoča v vsaki spremenljivki posebej.

Zgled 2.2. Naj bo $H(x, y) = (2x - 1)(2y - 1)$. Najprej preverimo, da je funkcija H na območju $\mathbf{I}^2$ 2-naraščajoča. Izberimo si poljubna števila x_1, x_2, y_1 ter y_2, kjer je $x_1 < x_2$ in $y_1 < y_2$. Potem za pravokotnik $B = [x_1, x_2] \times [y_1, y_2]$ velja $V_H(B) \geq 0$, kar dobimo iz sledečega izračuna.

$$\begin{aligned}
V_H(B) &= V_H([x_1, x_2] \times [y_1, y_2]) = \\
&= H(x_2, y_2) - H(x_2, y_1) - H(x_1, y_2) + H(x_1, y_1) = \\
&= (2x_2 - 1)(2y_2 - 1) - (2x_2 - 1)(2y_1 - 1) - (2x_1 - 1)(2y_2 - 1) + (2x_1 - 1)(2y_1 - 1) = \\
&= (2x_2 - 1)(2y_2 - 2y_1) - (2x_1 - 1)(2y_2 - 2y_1) = \\
&= 2(y_2 - y_1)(2x_2 - 2x_1) = \\
&= 4(y_2 - y_1)(x_2 - x_1) \geq 0.
\end{aligned}$$

Sedaj preverimo, da ni naraščajoča v vsaki spremenljivki posebej. Naj bo funkcija $h_x(x, y) = 2(2y - 1)$ odvod funkcije H po x in $h_y(x, y) = 2(2x - 1)$ odvod funkcije H po y. Vidimo, da je $h_y < 0$ za $x < 1/2$ in $h_x < 0$ za $y < 1/2$. Od tod sledi, da H ni naraščajoča v x za $y < 1/2$ in ne v y za $x < 1/2$. $\diamondsuit$

Velja tudi obratno, če je H naraščajoča v vsaki spremenljivki posebej, ni nujno 2-naraščajoča.

Zgled 2.3. Naj bo $H(x, y) = \max\{x, y\}$. Očitno je H naraščajoča v vsaki spremenljivki posebej. Izberimo $x_1 = 0$, $x_2 = 1$, $y_1 = 0$ in $y_2 = 1$. Potem je

$$\begin{aligned}
H(x_2, y_2) - H(x_1, y_2) - H(x_2, y_1) + H(x_1, y_1) &= \\
= H(1, 1) - H(1, 0) - H(0, 1) + H(0, 0) &= \\
= 1 - 1 - 1 + 0 = -1&,
\end{aligned}$$

od koder sledi, da ni 2-naraščajoča. $\diamondsuit$

Dokažimo nekaj preprostih lem, ki nam bodo pomagale definirati kopule.

Lema 2.4. *Naj bosta S_1 in S_2 neprazni podmnožici $\overline{\mathbb{R}}$ in naj bo H 2-naraščajoča funkcija z definicijskim območjem $S_1 \times S_2$. Naj bosta x_1, x_2 v S_1 in y_1, y_2 v S_2, pri čemer je $x_1 < x_2$ in $y_1 < y_2$. Potem je funkcija $f : S_1 \mapsto \mathbb{R}$, definirana s predpisom $f(t) = H(t, y_2) - H(t, y_1)$, naraščajoča na S_1 in funkcija $g : S_2 \mapsto \mathbb{R}$, definirana s predpisom $g(t) = H(x_2, t) - H(x_1, t)$, naraščajoča na S_2.*

Dokaz. Dokazujemo s protislovjem. Naj bo funkcija H na območju $S_1 \times S_2$ 2-naraščajoča, torej je $V_H(B) \geq 0$ za vse pravokotnike B z oglišči v $S_1 \times S_2$. Denimo, da obstajata $y_1, y_2 \in S_2$, $y_1 < y_2$, za katera funkcija $f(t) = H(t, y_2) - H(t, y_1)$ ni naraščajoča na S_1. Torej obstajata $t_1, t_2 \in S_1$, $t_1 < t_2$, da je $f(t_1) > f(t_2)$. Naj bo A pravokotnik $[t_1, t_2] \times [y_1, y_2]$.

$$0 > f(t_2) - f(t_1) = H(t_2, y_2) - H(t_2, y_1) - H(t_1, y_2) - H(t_1, y_1) = V_H(A)$$

Torej $V_H(B)$ ni nenegativno za vse pravokotnike z oglišči v $S_1 \times S_2$, kar je protislovje.

Podobno postopamo v primeru, ko funkcija g ni naraščajoča. $\square$

Definicija 2.5. Naj bosta S_1 in S_2 neprazni podmnožici $\overline{\mathbb{R}}$, ki vsebujeta svoj najmanjši element a_1 oziroma a_2. Pravimo, da je $H : S_1 \times S_2 \mapsto \mathbb{R}$ *prizemljena*, če velja $H(x, a_2) = 0 = H(a_1, y)$ za vsak $(x, y) \in S_1 \times S_2$.

Lema 2.6. *Naj bosta S_1 in S_2 neprazni podmnožici $\overline{\mathbb{R}}$, ki vsebujeta svoj najmanjši element a_1 oziroma a_2. Naj bo H prizemljena 2-naraščajoča funkcija z definicijskim območjem $S_1 \times S_2$. Potem je H naraščajoča v vsaki spremenljivki.*

Dokaz. Naj bo a_1 najmanjši element v S_1 in a_2 najmanjši element v S_2. Izberimo $x_1 = a_1$ in $y_1 = a_2$. Ker je $H(t, a_2) = 0$, sledi $f(t) = H(t, y_2) - H(t, a_2) = H(t, y_2)$ in $g(t) = H(x_2, t) - H(a_1, t)$. Od tod veljavnost sledi iz prejšnje leme. □

Definicija 2.7. Denimo, da sta S_1 in S_2 neprazni podmnožici $\overline{\mathbb{R}}$ ter b_1 in b_2 največja elementa v S_1 oziroma S_2. Naj bo H funkcija dveh spremenljivk, definirana na $S_1 \times S_2$. Potem lahko za funkcijo H definiramo *robni funkciji* F in G kot $F(x) = H(x, b_2)$ in $G(y) = H(b_1, y)$.

Lema 2.8. *Naj bosta S_1 in S_2 neprazni podmnožici $\overline{\mathbb{R}}$, ki vsebujeta svoj največji in najmanjši element. Naj bo H prizemljena 2-naraščajoča funkcija z definicijskim območjem $S_1 \times S_2$ in z robnima funkcijama F in G. Potem za poljubni točki (x_1, y_1) in (x_2, y_2) iz $S_1 \times S_2$ velja*

$$|H(x_2, y_2) - H(x_1, y_1)| \leq |F(x_2) - F(x_1)| + |G(y_2) - G(y_1)|.$$

Dokaz. Iz trikotniške enakosti sledi

$$|H(x_2, y_2) - H(x_1, y_1)| \leq |H(x_2, y_2) - H(x_1, y_2)| - |H(x_1, y_2) - H(x_1, y_1)|.$$

Zaradi simetrije lahko predpostavimo, da je $x_1 \leq x_2$. Funkcija H je prizemljena in 2-naraščajoča, zato si lahko pomagamo z lemama 2.4 in 2.6 . Dobimo

$$0 \leq H(x_2, y_2) - H(x_1, y_2) \leq F(x_2) - F(x_1).$$

Za $x_1 > x_2$ velja

$$0 \leq -\big(H(x_2, y_2) - H(x_1, y_2)\big) \leq -\big(F(x_2) - F(x_1)\big),$$

torej za vse $x_1, x_2 \in S_1$ velja

$$|H(x_2, y_2) - H(x_1, y_2)| \leq |F(x_2) - F(x_1)|.$$

Podoben razmislek nam pove, da tudi za vse $y_1, y_2 \in S_2$ velja $|H(x_1, y_2) - H(x_1, y_1)| \leq |G(y_2) - G(y_1)|$, torej je $|H(x_2, y_2) - H(x_1, y_1)| \leq |F(x_2) - F(x_1)| + |G(y_2) - G(y_1)|$. □

2.2. **Kopule.**

Sedaj lahko definiramo, kaj so podkopule in kopule. Na prvi pogled se bo morda razlika zdela skorajda zanemarljiva, a je, kot bomo videli v nadaljevanju, izjemno pomembna.

Definicija 2.9. *2-dimenzionalna podkopula* je funkcija dveh spremenljivk C', ki ima naslednje lastnosti:

- $D_{C'} = S_1 \times S_2$, kjer sta S_1 in S_2 podmnožici intervala $\mathbf{I}$, ki vsebujeta točki 0 in 1.
- C' je prizemljena in 2-naraščajoča.
- Za vsak $u \in S_1$ in za vsak $v \in S_2$, $C'(u, 1) = u$ in $C'(1, v) = v$.

Dokažimo, da je $0 \leq C'(u, v) \leq 1$.
Izberimo poljuben $v \in S_2$. Množica S_1 vsebuje $x_1 = 0$ in ker je funkcija C' prizemljena, je $C'(0, v) = 0$. Podobno je $C'(u, 0) = 0$ za vsak $u \in S_1$.
Ker je $C'(u, v)$ 2-naraščajoča funkcija, iz leme 2.6 sledi, da je naraščajoča v vsaki spremenljivki posebej. Iz zadnje alineje sledi, da je $C'(1, v) = v$, torej za vsak $v \in S_1$ velja, da je $0 \leq C'(u, v) \leq v$. Podobno je tudi $C'(u, v) \leq u$.

Ker za vsak $(u, v) \in D_{C'}$ velja $0 \leq C'(u, v) \leq 1$, opazimo, da je tudi zaloga vrednosti podkopule $Z_{C'}$ podmnožica intervala $\mathbf{I}$.

Definicija 2.10. *2-dimenzionalna kopula* je 2-dimenzionalna podkopula, ki ima za definicijsko območje celoten $\mathbf{I}^2$.

Zgled 2.11. Funkcija $M(u, v) = \min\{u, v\}$ je kopula. ♢

Dokaz. Očitno je funkcija $\min\{u, v\}$ definirana na celotnem območju $\mathbf{I}^2$. Za vsak $(u, v) \in \mathbf{I}^2$ velja $\min\{u, 1\} = u$ in $\min\{1, v\} = v$. Prav tako velja $\min\{u, 0\} = 0$ in $\min\{0, v\} = 0$, torej je funkcija M prizemljena.

Dokažimo, da je še 2-naraščajoča. Naj bo $x_2 > x_1$ in $y_2 > y_1$. Ločimo različne možnosti.
Če velja $y_2 > y_1 \geq x_2 > x_1$, je $M(x_2, y_2) - M(x_1, y_2) - M(x_2, y_1) + M(x_1, y_1) =$
$= x_2 - x_2 - x_1 + x_1 = 0$.
Če velja $y_2 \geq x_2 > y_1 \geq x_1$, dobimo $x_2 - x_1 - y_1 + x_1 = x_2 - y_1 \geq 0$.
Če velja $y_2 \geq x_2 > x_1 > y_2$, dobimo $x_2 - y_1 - x_1 + x_1 = x_2 - y_1 \geq 0$.
Če velja $x_2 > y_2 > y_1 \geq x_1$, dobimo $y_2 - y_1 - x_1 + x_1 = y_2 - y_1 \geq 0$.
Če velja $x_2 > y_2 \geq x_1 > y_1$, dobimo $y_2 - y_1 - x_1 + y_1 = y_2 - x_1 \geq 0$.
Če velja $x_2 > x_1 > y_2 > y_1$, dobimo $y_2 - y_1 - y_2 + y_1 = 0$.

S tem smo zajeli vse primere in tako pokazali, da je funkcija M tudi 2-naraščajoča. □

Zgled 2.12. Funkcija $W(u, v) = \max\{u + v - 1, 0\}$ je kopula. ♢

Dokaz. Najprej preverimo, da je $W(u, 1) = u$ in $W(1, v) = v$.
$W(u, 1) = \max\{u+1-1, 0\} = \max\{u, 0\} = u$. Podobno je $W(1, v) = \max\{1+v-1, 0\} = v$.

Preverimo prizemljenost. Za $u = 0$ je $W(0, v) = \max\{v - 1, 0\}$. Ker je $u \in [0, 1]$, je posledično $W(0, v) = 0$. Podobno dobimo v primeru, ko je $v = 0$.

Preverimo še, da je funkcija W 2-naraščajoča, torej računamo
$V_W([x_1, x_2] \times [y_1, y_2]) = W(x_2, y_2) - W(x_1, y_2) - W(x_2, y_1) + W(x_1, y_1)$.
Naj velja $x_2 > x_1$ in $y_2 > y_1$.

- Naj bo $x_1 + y_1 - 1 \geq 0$, dobimo
$$(x_2 + y_2 - 1) - (x_1 + y_2 - 1) - (x_2 + y_1 - 1) + (x_1 + y_1 - 1) = 0.$$
- Naj bo $x_2 + y_2 - 1 \geq 0$, $x_1 + y_2 - 1 \geq 0$, $x_2 + y_1 - 1 \geq 0$ in $x_1 + y_1 - 1 < 0$. Potem dobimo
$$\begin{aligned} V_W([x_1, x_2] \times [y_1, y_2]) &= (x_2 + y_2 - 1) - (x_1 + y_2 - 1) - (x_2 + y_1 - 1) + 0 = \\ &= -x_1 - y_1 + 1 = -(x_1 + y_1 - 1) > 0. \end{aligned}$$
- Naj bo $x_2 + y_2 - 1 \geq 0$, $x_1 + y_2 - 1 \geq 0$, $x_2 + y_1 - 1 < 0$ in $x_1 + y_1 - 1 < 0$. Potem dobimo
$$\begin{aligned} V_W([x_1, x_2] \times [y_1, y_2]) &= (x_2 + y_2 - 1) - (x_1 + y_2 - 1) - 0 + 0 = \\ &= x_2 - x_1 > 0. \end{aligned}$$
- Naj bo $x_2 + y_2 - 1 \geq 0$, $x_1 + y_2 - 1 < 0$, $x_2 + y_1 - 1 \geq 0$ in $x_1 + y_1 - 1 < 0$. Potem dobimo
$$\begin{aligned} V_W([x_1, x_2] \times [y_1, y_2]) &= (x_2 + y_2 - 1) - 0 - (x_2 + y_1 - 1) + 0 = \\ &= y_2 - y_1 > 0. \end{aligned}$$
- Naj bo $x_2 + y_2 - 1 \geq 0$, $x_1 + y_2 - 1 < 0$, $x_2 + y_1 - 1 < 0$ in $x_1 + y_1 - 1 < 0$. Potem dobimo
$$\begin{aligned} V_W([x_1, x_2] \times [y_1, y_2]) &= (x_2 + y_2 - 1) - 0 - 0 + 0 = \\ &= x_2 + y_2 - 1 \geq 0. \end{aligned}$$
- Naj bo $x_2 + y_2 - 1 < 0$, $x_1 + y_2 - 1 < 0$, $x_2 + y_1 - 1 < 0$ in $x_1 + y_1 - 1 < 0$. Potem dobimo
$$V_W([x_1, x_2] \times [y_1, y_2]) = 0 - 0 - 0 + 0 = 0$$

S tem smo za vse možne primere pokazali, da je $V_W([x_1, x_2] \times [y_1, y_2]) \geq 0$, torej je 2-naraščajoča. □

Zgled 2.13. Funkcija $\Pi(u, v) = uv$ je kopula. ♢

Dokaz. Očitno je v primeru $u = 0$ ali $v = 0$ funkcija $\Pi(u, v) = 0$, torej je prizemljena. Velja tudi $\Pi(u, 1) = u \cdot 1 = u$ in $\Pi(1, v) = 1 \cdot v = v$.

Pokažimo še 2-naraščanje. Naj velja $x_2 > x_1$ in $y_2 > y_1$.

$$\begin{aligned} V_\Pi([x_1, x_2] \times [y_1, y_2]) &= x_2 y_2 - x_2 y_1 - x_1 y_2 + x_1 y_1 = \\ &= x_2 (y_2 - y_1) - x_1 (y_2 - y_1) = \\ &= (y_2 - y_1)(x_2 - x_1) > 0. \end{aligned}$$

□

Izrek 2.14 (Fréchet - Hoeffdingovi meji)**.** *Naj bo C' podkopula. Potem za vsak $(u, v) \in D_{C'}$ velja*

$$W(u, v) \leq C'(u, v) \leq M(u, v).$$

Ker je vsaka kopula tudi podkopula, neenakosti veljata tudi v primeru kopul.

Dokaz. Izberimo poljubno točko (u, v) iz $D_{C'}$. Vemo, da je $C'(u, v) \leq C'(u, 1) = u$ in $C'(u, v) \leq C'(1, v) = v$. To nam da $C'(u, v) \leq \min\{u, v\}$.

Ker je funkcija C' 2-naraščajoča, velja $V_{C'}([u, 1] \times [v, 1]) \geq 0$, od koder sledi $C'(u, v) \geq u + v - 1$. Za podkopule velja tudi $C'(u, v) \geq 0$ in tako sledi $C'(u, v) \geq \max\{u + v - 1, 0\}$. □

Ta izrek utemeljuje naslednjo definicijo.

Definicija 2.15. Kopulo $M(u, v) = \min\{u, v\}$ imenujemo zgornja, $W(u, v) = \max\{u + v - 1, 0\}$ pa spodnja *Fréchet-Hoeffdingova meja.* Poleg teh dveh bomo pogosto uporabljali tudi *produktno kopulo*, ki smo jo videli že zgoraj in je definirana kot $\Pi(u, v) = uv$.

Definicija 2.16. Naj bo C kopula in $a \in \mathbf{I}$. *Vodoravni prerez* funkcije C pri a je funkcija iz $\mathbf{I}$ v $\mathbf{I}$, podana s predpisom $t \mapsto C(t, a)$, *navpični prerez* funkcije C pri a je funkcija $t \mapsto C(a, t)$, ki slika iz $\mathbf{I}$ v $\mathbf{I}$, *diagonalni prerez* funkcije C pa je funkcija δ_C, ki slika iz $\mathbf{I}$ v $\mathbf{I}$, definirana z $\delta_C(t) = C(t, t)$.

Trditev 2.17. *Vodoravni, navpični in diagonalni prerezi kopule C so nepadajoče in enakomerno zvezne funkcije.*

Dokaz. Funkcija C je kopula, zato je 2-naraščajoča. Posledično je naraščajoča v vsaki spremenljivki posebej, od koder sledi, da so prerezi nepadajoče funkcije. Podkopule (in tudi kopule) imajo lastnost, da so enakomerno zvezne na svojem definicijskem območju (dokaz te lastnosti je v [3] na strani 11). Sledi, da so prerezi enakomerno zvezne funkcije. □

2.3. **Sklarov izrek.**

Sklarov izrek je ključni izrek v teoriji kopul in je temelj za večino uporab te teorije v statistiki. Z njim lahko razložimo povezavo med skupno porazdelitveno funkcijo slučajnega vektorja in porazdelitvenimi funkcijami posameznih slučajnih spremenljivk.

Iz teorije verjetnosti se spomnimo naslednje trditve.

Trditev 2.18. *Naj bo X slučajna spremenljivka s porazdelitveno funkcijo F, ki je definirana na $\overline{\mathbb{R}}$. Funkcija F je naraščajoča in zanjo velja $F(-\infty) = 0, F(\infty) = 1$. Naj bo (X, Y) slučajni vektor s skupno porazdelitveno funkcijo H, ki je definirana na $\overline{\mathbb{R}}^2$. Funkcija H je 2-naraščajoča in velja $H(x, -\infty) = 0, H(-\infty, y) = 0, H(\infty, \infty) = 1$.*

Izrek 2.19 (Sklarov izrek)**.** *Naj bo H skupna porazdelitvena funkcija z robnima porazdelitvenima funkcijama F in G. Potem obstaja taka kopula C, da za vsak $x, y \in \overline{\mathbb{R}}$ velja*

$$H(x, y) = C(F(x), G(y)).$$

Če sta funkciji F in G zvezni, potem je kopula C enolično določena s to formulo, sicer pa je C enolično določena samo na množici $Z_F \times Z_G$.

Velja tudi obratno, če je C kopula in sta F in G porazdelitveni funkciji slučajnih spremenljivk X in Y, potem je H skupna porazdelitvena funkcija slučajnega vektorja (X, Y) z robnima porazdelitvenima funkcijama F in G.

Sklarov izrek bomo dokazali s pomočjo naslednjih dveh lem, katerih dokaza bomo izpustili. Bralec ju lahko najde v [3] na straneh 18-21.

Lema 2.20. *Naj bo H skupna porazdelitvena funkcija z robnima porazdelitvenima funkcijama F in G. Potem obstaja enolična podkopula C', za katero velja $D_{C'} = Z_F \times Z_G$ in za vsak $x, y \in D_{C'}$ velja $H(x, y) = C'(F(x), G(y))$.*

Lema 2.21. *Naj bo C' podkopula. Potem obstaja kopula C, za katero velja $C(u, v) = C'(u, v)$ za vsak $(u, v) \in D_{C'}$. Vsako podkopulo se torej da razširiti v kopulo, a razširitev v splošnem ni enolična.*

Dokaz Sklarovovega izreka. Iz lem 2.20 in 2.21 sledi, da obstaja kopula C, ki zadošča enačbi $H(x, y) = C(F(x), G(y))$ za vse $x, y \in \overline{\mathbb{R}}$. Če sta porazdelitveni funkciji F in G zvezni, je $Z_F \times Z_G = \mathbf{I}^2$, zato je C enolična. V nasprotnem primeru pa je C' enolično določena samo na $Z_F \times Z_G$.

Da dokažemo izrek v nasprotno stran, moramo preveriti lastnosti skupne porazdelitvene funkcije. Ker je C kopula, je 2-naraščajoča in posledično naraščajoča v vsaki spremenljvki posebej. Sledi, da je funkcija H naraščajoča v vsaki spremenljivki posebej. F in G sta porazdelitveni funkciji, zato velja, da v limiti, ko gresta x in y proti $-\infty$, velja $\lim_{x \to -\infty} F(x) = 0$ in $\lim_{y \to -\infty} G(y) = 0$. C je kopula, zato velja $C(0, 0) = 0$, torej je $\lim_{x,y \to -\infty} H(x, y) = 0$. Iz podobnega razloga velja, da ko gresta x in y proti ∞, velja $\lim_{x \to \infty} F(x) = 1$ in $\lim_{y \to \infty} G(y) = 1$. C je kopula, zato velja $C(1, 1) = 1$, torej je $\lim_{x,y \to \infty} H(x, y) = 1$. Poleg tega za kopule velja, da je njihova zaloga vrednosti enaka $\mathbf{I}$. Ker so podkopule enakokomerno zvezne na svojem definicijskem območju ter sta funkciji F in G zvezni z desne, je tudi funkcija H zvezna z desne. $\square$

Formula iz Sklarovega izreka nam da skupno porazdelitveno funkcijo iz robnih porazdelitvenih funkcij. Včasih pa bi želeli izraziti kopulo s pomočjo skupne porazdelitvene funkcije in inverzov posameznih porazdelitvenih funkcij. Ker funkciji F in G nista nujno strogo naraščajoči, ni nujno, da inverz obstaja, zato definiramo kvazi inverz porazdelitvene funkcije.

Definicija 2.22. Naj bo F porazdelitvena funkcija. Potem je *kvazi inverz* funkcije F funkcija $F^{(-1)}$, definirana na $\mathbf{I}$, za katero velja $F^{(-1)}(t) = \inf\{x \in \mathbb{R} | F(x) \geq t\}$.

Če je F strogo naraščajoča, je kvazi inverz kar običajni inverz in ga označevali s F^{-1}.

Posledica 2.23. *Naj bodo H, F, G in C' definirane tako kot v lemi 2.20 in naj bosta $F^{(-1)}$ in $G^{(-1)}$ kvazi inverza za F in G. Potem za poljuben $(u, v) \in D_{C'}$ velja*

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

Ko sta F in G zvezni, je podkopula C' kopula. To nam daje orodje za konstrukcijo kopul iz skupne porazdelitvene funkcije.

Zgled 2.24. Naj bo H porazdelitvena funkcija slučajnega vektorja (X, Y), definirana z

$$H(x, y) = \begin{cases} \frac{x(e^y - 1)}{x + e^y - 1}, & \text{če } (x, y) \in [0, 1] \times [0, \infty], \\ 1 - e^{-y}, & \text{če } (x, y) \in (1, \infty] \times [0, \infty], \\ 0, & \text{sicer.} \end{cases}$$

Preverimo, da je H res porazdelitvena funkcija. H je v vsaki spremenljivki zase naraščajoča in zvezna z desne. Velja $\lim_{x,y\to\infty} H(x,y) = 1$ ter tudi $\lim_{x\to-\infty} H(x,y) = 0$ in $\lim_{y\to-\infty} H(x,y) = 0$.

Funkciji $F(x) = H(x,\infty)$ in $G(y) = H(\infty,y)$ podani z

$$F(x) = \begin{cases} 0, & x < 0, \\ x, & x \in [0,1], \\ 1, & x > 1, \end{cases} \qquad G(y) = \begin{cases} 0, & y < 0, \\ 1 - e^{-y}, & y \geq 0. \end{cases}$$

Izračunamo, da sta kvazi inverza enaka $F^{(-1)}(u) = u$ in $G^{(-1)}(v) = -\ln(1-v)$ za $u, v \in \mathbf{I}$. Ker sta Z_F in Z_G ravno enaki $\mathbf{I}$, s pomočjo zgornje posledice dobimo kopulo C z enačbo

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

saj velja

$$C(u,v) = H\big(F^{(-1)}(u), G^{(-1)}(v)\big) = H\big(u, -\ln(1-v)\big) = \frac{u(e^{-\ln(1-v)} - 1)}{u + (e^{-\ln(1-v)}) - 1} =$$

$$= \frac{u(\frac{1}{1-v} - 1)}{u + (\frac{1}{1-v} - 1)} = \frac{\frac{uv}{1-v}}{\frac{u(1-v)+v}{1-v}} = \frac{uv}{u+v-uv}.$$

$\diamondsuit$

2.4. **Kopule in slučajne spremenljivke.**
S kopulami bi si rabi pomagali pri obravnavanju odvisnosti slučajnih spremenljivk, zato si poglejmo, kakšne lastnosti veljajo za kopule v primeru slučajnih spremenljivk.

Izrek 2.25. *Naj bosta X in Y slučajni spremenljivki s porazdelitvenima funkcijama F oziroma G in skupno porazdelitveno funkcijo H. Potem obstaja kopula C, da velja enakost $H(x,y) = C(F(x), G(x))$. Če sta F in G zvezni, je enolična, sicer pa je enolično določena na $Z_F \times Z_G$.*

Bralec lahko podrobnejšo razlago najde v [3] na strani 25.

Izrek 2.26. *Naj bosta X in Y zvezni slučajni spremenljivki in $\Pi(u,v) = uv$. Potem sta X in Y neodvisni natanko tedaj, ko je $C_{XY} = \Pi$.*

Dokaz. Naj bo F porazdelitvena funkcija slučajne spremenljivke X in G porazdelitvena funkcija slučajne spremenljivke Y. Iz teorije verjetnosti vemo, da sta slučajni spremenljivki X in Y neodvisni natanko tedaj, ko velja $H(x,y) = P(X \leq x, Y \leq y) = P(X \leq x)P(Y \leq y) = F(x)G(y)$ za vse $x, y \in \overline{\mathbb{R}}$. Po prejšnjem izreku obstaja enolična kopula C, da drži enakost $H(x,y) = C(F(x), G(y))$. Ker je $H(x,y) = F(x)G(y)$, velja $C_{XY}(u,v) = uv$. $\square$

Ena od pomembnih lastnosti kopul je, da se v primeru strogo monotonih transformacij slučajnih spremenljivk kopula ali ne spremeni ali pa se spremeni na vnaprej znan način.

Izrek 2.27. *Naj bosta X in Y zvezni slučajni spremenljivki s kopulo C_{XY} ter α in β strogo monotoni funkciji na $\mathbf{I}$.*

- *Če sta α in β strogo naraščajoči funkciji, potem je*
$$C_{\alpha(X)\beta(Y)} = C_{XY},$$
torej se kopula ne spremeni.
- *Če je α strogo naraščajoča in β strogo padajoča, potem je*
$$C_{\alpha(X)\beta(Y)}(u,v) = u - C_{XY}(u, 1-v).$$

- *Če je α strogo padajoča in β strogo naraščajoča, potem je*
$$C_{\alpha(X)\beta(Y)}(u,v) = v - C_{XY}(1-u,v).$$
- *Če sta α in β obe strogo padajoči, potem je*
$$C_{\alpha(X)\beta(Y)}(u,v) = u + v - 1 + C_{XY}(1-u,1-v).$$

Dokaz. Dokažimo izrek v primeru, ko sta α in β strogo naraščajoči funkciji. Naj bo F_1 porazdelitvena funkcija slučajne spremenljivke X in G_1 porazdelitvena funkcija slučajne spremenljivke Y. Naj bosta še F_2 porazdelitvena funkcija slučajne spremenljivke $\alpha(X)$ in G_2 porazdelitvena funkcija slučajne spremenljivke $\beta(Y)$. Ker sta α in β strogo naraščajoči, je $F_2(x) = P\big(\alpha(X) \le x\big) = P\big(X \le \alpha^{-1}(x)\big) = F_1\big(\alpha^{-1}(x)\big)$ in analogno $G_2(y) = G_1\big(\beta^{-1}(y)\big)$. Posledično za poljubna $x, y \in \overline{\mathbb{R}}$ velja

$$\begin{aligned}
C_{\alpha(X)\beta(Y)}\big(F_2(x), G_2(y)\big) &= P\big(\alpha(X) \le x, \beta(Y) \le y\big) = \\
&= P\big(X \le \alpha^{-1}(x), Y \le \beta^{-1}(y)\big) = \\
&= C_{XY}\big(F_1(\alpha^{-1}(x)), G_1(\beta^{-1}(y))\big) = \\
&= C_{XY}\big(F_2(x), G_2(y)\big).
\end{aligned}$$

Ker sta X in Y zvezni, je $Z_{F_2} = Z_{G_2} = \mathbf{I}$, od koder sledi $C_{\alpha(X),\beta(Y)} = C_{XY}$.

Dokaze za ostale tri primere lahko najdemo v [3] na strani 26. □

2.5. **Mere odvisnosti.**

Eno od glavnih področji, s katerimi se ukvarjajo statistiki, je odvisnost med slučajnimi spremenljivkami. Eden od načinov, kako lahko ugotovimo odvisnost med njimi, je s pomočjo skladnosti (angl. *concordance*). Skladnost najpogosteje izračunamo s pomočjo mer skladnosti, denimo Kendallovega τ ali Spearmanovega ρ.

Definicija 2.28 (skladnost)**.** Naj bosta para (x_1, y_1) in (x_2, y_2) dve realizaciji zveznega slučajnega vektorja (X, Y). Rečemo, da sta (x_1, y_1) in (x_2, y_2) *skladna*, če velja $x_1 < x_2$ in $y_1 < y_2$ ali $x_1 > x_2$ in $y_1 > y_2$. V nasprotnem primeru, ko je $x_1 < x_2$ in $y_1 > y_2$ ali $x_1 > x_2$ in $y_1 < y_2$, pa pravimo, da sta (x_1, y_1) in (x_2, y_2) neskladna.

Definicija 2.29 (Kendallov τ)**.** *Kendallov τ* lahko definiramo na dva načina, s pomočjo slučajnega vzorca in s pomočjo verjetnosti skladnosti.

Denimo, da s pari $(x_1, y_1), \ldots, (x_n, y_n)$ označimo n realizacij zveznega slučajnega vektorja (X, Y). Predpostavimo še, da je $x_i \ne x_j$ in $y_i \ne y_j$ za vsaka $i \ne j$. Imamo $\binom{n}{2}$ kombinacij parov (x_i, y_i) in (x_j, y_j), od katerih jih je c skladnih in d neskladnih. *Vzorčni Kendallov τ* je definiran kot
$$\tau = \frac{c-d}{c+d} = \frac{2\,(c-d)}{n\,(n-1)}.$$

Naj bo H skupna porazdelitvena funkcija zveznega slučajnega vektorja (X, Y). Naj bosta (X_1, Y_1) in (X_2, Y_2) neodvisna enako porazdeljena slučajna vektorja s skupno porazdelitveno funkcijo H. *Kendallov τ* je definiran z enačbo
$$\tau = \tau_{X,Y} = P\Big((X_1 - X_2)(Y_1 - Y_2) > 0\Big) - P\Big((X_1 - X_2)(Y_1 - Y_2) < 0\Big).$$

Izrek 2.30. *Naj bosta (X_1, Y_1) in (X_2, Y_2) neodvisna slučajna vektorja s skupnima porazdelitvenima funkcijama H_1 in H_2 in z enakima robnima porazdelitvama F (za X_1 in X_2) in G (za Y_1 in Y_2). S C_1 in C_2 označimo kopuli (X_1, Y_1) in (X_2, Y_2), torej je $H_1 = C_1(F(x), G(y))$ in $H_2 = C_2(F(x), G(y))$. S Q označimo razliko med verjetnostjo skladnosti in verjetnostjo neskladnosti,*
$$Q(C_1, C_2) = P\Big((X_1 - X_2)(Y_1 - Y_2) > 0\Big) - P\Big((X_1 - X_2)(Y_1 - Y_2) < 0\Big).$$

Potem je

$$Q = Q(C_1, C_2) = -1 + 4 \iint_{\mathbf{I}^2} C_2(u, v)\ dC_1(u, v).$$

Dokaz tega izreka je v [3] na strani 159.

Posledica 2.31. *Naj bodo C_1, C_2 in Q definirani kot v izreku 2.30. Potem velja, da je Q simetrična v svojih argumentih, $Q(C_1, C_2) = Q(C_2, C_1)$.*

Dokaz. Velja $Q(C_1, C_2) = P\big((X_2 - X_1)(Y_2 - Y_1) > 0\big) - P\big((X_2 - X_1)(Y_2 - Y_1) > 0\big)$ in $Q(C_2, C_1) = P\big((X_1 - X_2)(Y_1 - Y_2) > 0\big) - P\big((X_1 - X_2)(Y_2 - Y_1) > 0\big)$.

Ker je $(X_2 - X_1)(Y_2 - Y_1) = (X_1 - X_2)(Y_1 - Y_2)$, enakost sledi. □

Definicija 2.32. Naj bo $X : \Omega \to \mathbb{R}$ slučajna spremenljvka na verjetnostnem prostoru (Ω, F, P). Potem je *nosilec (angl. support)* slučajne spremenljivke X zaprtje najmanjše množice $R_X \subset \mathbb{R}$, da je $P(X \in \mathbb{R}_X) = 1$. V primeru diskretne slučajne spremenljivke je tako nosilec R_X enak zaprtju množice $\{x \in \mathbb{R} : P(X = x) > 0\}$, v primeru zvezne pa je enak zaprju množice $\{x \in \mathbb{R} : f_X(x) > 0\}$, pri čemer je $f_X(x)$ gostota slučajne spremenljivke X v točki x.

Podobno velja tudi v primeru, ko delamo z večimi slučajnimi spremenljivkami. Naj bo $(X, Y) : \Omega \to \mathbb{R}^2$ slučajni vektor na verjetnostnem prostoru (Ω, F, P). Potem je *nosilec* slučajnega vektorja (X, Y) zaprtje najmanjše množice $R_{(X,Y)} \subset \mathbb{R}^2$, da je $P((X, Y) \in \mathbb{R}_{(X,Y)}) = 1$. V primeru diskretnega slučajnega vektorja je tako nosilec $R_{(X,Y)}$ enak zaprtju množice $\{(x, y) \in \mathbb{R}^2 : P(X = x, Y = y) > 0\}$, v zveznem primeru pa je nosilec enak zaprtju množice $\{(x, y) \in \mathbb{R}^2 : f_{(X,Y)}(x, y) > 0\}$, pri čemer je $f_{(X,Y)}(x, y)$ gostota slučajnega vektorja (X, Y) v točki (x, y).

Nosilec za kopulo C je kar nosilec slučajnega vektorja, katerega porazdelitvena funkcija je C.

V primeru kopule M je njen nosilec diagonala $u = v$, torej $R_M = \{(u, u) : u \in \mathbf{I}\}$. Nosilec za kopulo W je diagonala $u = 1 - v$, kar lahko označimo $R_W = \{(u, 1-u) : u \in \mathbf{I}\}$, nosilec za kopulo Π pa je kar celoten kvadrat $\mathbf{I}^2$.

Zgled 2.33. V primeru, ko delamo s kopulami M, Π in W, lahko funkcijo Q izračunamo dokaj preprosto. Nosilec za funkcijo M je diagonala $v = u$ za $v \in \mathbf{I}$, za W pa $v = 1 - u$ za $v \in \mathbf{I}$. Ker ima M enakomerno zvezni robni porazdelitveni funkciji na intervalu $(0, 1)$, sledi, da če je g integrabilna funkcija, definirana na $\mathbf{I}^2$, potem je

$$\iint_{\mathbf{I}^2} g(u, v) dM(u, v) = \int_0^1 g(u, u) du.$$

Od tod sledi

$$\begin{aligned}
Q(M, M) &= -1 + 4 \iint_{\mathbf{I}^2} \min\{u, v\}\ dM(u, v) = -1 + 4 \int_0^1 u\ du = 1, \\
Q(M, \Pi) &= -1 + 4 \iint_{\mathbf{I}^2} uv\ dM(u, v) = -1 + 4 \int_0^1 u^2\ du = 1/3, \\
Q(M, W) &= -1 + 4 \iint_{\mathbf{I}^2} \max\{u + v - 1, 0\}\ dM(u, v) = -1 + 4 \int_{1/2}^1 (2u - 1)\ du = 0.
\end{aligned}$$

Podobno dobimo tudi

$$\iint_{\mathbf{I}^2} g(u, v)\ dW(u, v) = \int_0^1 g(u, 1 - u)\ du$$

in posledično

$$Q(W,\Pi) = -1 + 4\iint_{\mathbf{I}^2} uv\ dW(u,v) = -1 + 4\int_0^1 u(1-u)\ du = -1/3,$$

$$Q(W,W) = -1 + 4\iint_{\mathbf{I}^2} \max\{u+v-1,0\}\ dW(u,v) = -1 + 4\int_0^1 0\ du = -1.$$

Ker velja $d\Pi(u,v) = dudv$, dobimo še

$$Q(\Pi,\Pi) = -1 + 4\iint_{\mathbf{I}^2} uv\ d\Pi(u,v) = -1 + 4\int_0^1\int_0^1 uv\ dudv = 0.$$

$\diamondsuit$

Zgled 2.34. Naj bo parameter $\theta \in [-1,1]$. Naj bo C_θ iz Farlie-Gumbel-Morgensternove družine kopul, torej $C_\theta(u,v) = uv + \theta uv(1-u)(1-v)$. Funkcija C_θ je dvakrat zvezno odvedljiva, zato je

$$dC_\theta(u,v) = \frac{\partial^2 C_\theta(u,v)}{\partial u \partial v} dudv = \Big(1 + \theta(1-2u)(1-2v)\Big) dudv.$$

Sledi, da je

$$\iint_{\mathbf{I}^2} C_\theta(u,v) dC_\theta(u,v) = \frac{1}{4} + \frac{\theta}{18},$$

od koder dobimo $\tau_\theta = \frac{2\theta}{9}$. $\diamondsuit$

Zgled 2.35. Naj bosta parametra $\alpha \geq 0$, $\beta \geq 0$ in $\alpha + \beta \leq 1$. Naj bo $C_{\alpha,\beta}$ iz Fréchetove družine kopul, torej $C_{\alpha,\beta} = \alpha M + (1-\alpha-\beta)\Pi + \beta W$. Opazimo, da je kopula C konveksna kombinacija kopul M, W in Π.

Izračunamo, da je $dC_{\alpha,\beta} = \alpha dM + (1-\alpha-\beta)d\Pi + \beta dW$, torej je

$$\tau_{\alpha,\beta} = -1 + 4\iint_{\mathbf{I}^2} \big(\alpha M + (1-\alpha-\beta)\Pi + \beta W\big)\big(\alpha dM + (1-\alpha-\beta)d\Pi + \beta dW\big).$$

Uporabimo linearnost integrala in rezultate iz zgleda 2.33 in dobimo

$$\tau_{\alpha,\beta} = \frac{(\alpha-\beta)(\alpha+\beta+2)}{3}.$$

$\diamondsuit$

Definicija 2.36 (Spearmanov ρ)**.** Druga pomembna mera odvisnosti, ki je prav tako definirana s pomočjo skladnosti, je *Spearmanov ρ*.

Naj bodo (X_1,Y_1), (X_2,Y_2) in (X_3,Y_3) neodvisni slučajni vektorji s skupno porazdelitveno funkcijo H in kopulo C. *Vzorčna vrednost* $\rho_{X,Y}$ Spearmanovega ρ je definirana z enačbo

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

Porazdelitvena funkcija za vektor (X_1,Y_1) je enaka $H(x,y)$, za vektor (X_2,Y_3) pa je $F(x)G(y)$, saj sta slučajni spremenljivki X_2 in Y_3 neodvisni (in ju lahko modeliramo s kopulo Π).

Izrek 2.37. *Naj bosta X in Y zvezni slučajni spremenljivki s kopulo C. Potem je vzorčni Spearmanov ρ za X in Y podan z enačbo*

$$\begin{aligned}\rho_{X,Y} = \rho_C = {} & 3\ Q(C,\Pi) = \\ & = -3 + 12\iint_{\mathbf{I}^2} uv\ dC(u,v) = \\ & = -3 + 12\iint_{\mathbf{I}^2} C(u,v)\ dudv.\end{aligned}$$

Dokaz. Naj bo C kopula za slučajni vektor (X_1, Y_1). Slučajni spremenljivki X_2 in Y_3 sta neodvisni, zaradi česar je njuna kopula ravno Π. Uporabimo izrek 2.30, ki pove, da je

$$Q(C, \Pi) = -1 + 4 \iint_{\mathbf{I}^2} \Pi(u, v) dC_1(u, v) = -1 + 4 \iint_{\mathbf{I}^2} uv \;\; dC_1(u, v).$$

Uporabimo posledico 2.31 in dobimo $Q = -1 + 4 \iint_{\mathbf{I}^2} C_1(u, v) dudv$. Torej je

$$3Q(C, \Pi) = -3 + 12 \iint_{\mathbf{I}^2} C_1(u, v) dudv.$$

□

Zgled 2.38. Naj bo $C_{\alpha,\beta}$ iz Fréchetove družine kopul s parametroma $\alpha \geq 0$, $\beta \geq 0$ in $\alpha + \beta \leq 1$. Potem je $C_{\alpha,\beta} = \alpha M + (1 - \alpha - \beta)\Pi + \beta W$ in

$$\begin{aligned} Q(C_{\alpha,\beta}, \Pi) &= \alpha \; Q(M, \Pi) + (1 - \alpha - \beta) \; Q(\Pi, \Pi) + \beta \; Q(W, \Pi) \\ &= \alpha \cdot \Big(\frac{1}{3}\Big) + \Big(1 - \alpha - \beta\Big) \cdot 0 + \beta \cdot \Big(-\frac{1}{3}\Big) \\ &= \frac{\alpha - \beta}{3}. \end{aligned}$$

in sledi $\rho_{\alpha,\beta} = \alpha - \beta$. ♢

Zgled 2.39. Vzemimo parameter θ z intervala $[-1, 1]$. Naj bo C_θ iz Farlie-Gumbel-Morgensternove družine kopul, torej $C_\theta(u, v) = uv + \theta uv(1 - u)(1 - v)$. Sledi

$$\iint_{\mathbf{I}^2} C_\theta(u, v) \; dudv = \frac{1}{4} + \frac{\theta}{36},$$

od koder dobimo $\rho_\theta = \frac{\theta}{3}$. ♢

Naslednjo definicijo lahko najdemo v [3] na strani 168.

Definicija 2.40 (Scarsini)**.** Naj bosta X in Y slučajni spremenljivki, ki ju povezuje kopula C. *Mera odvisnost*i κ je funkcija, ki slučajnemu vektorju (X, Y) priredi število $\kappa_{X,Y}$ in je odvisno samo od kopule C, ne pa od robnih porazdelitev. Mera odvisnosti je mera skladnosti, če zadošča naslednjim lastnostim:

(1) κ je definirana za vsak par X, Y zveznih slučajnih spremenljivk;
(2) $-1 \leq \kappa_{X,Y} \leq 1$, $\kappa_{X,X} = 1$, $\kappa_{X,-X} = -1$;
(3) $\kappa_{X,Y} = \kappa_{Y,X}$;
(4) če sta X in Y neodvisni, potem je $\kappa_{X,Y} = \kappa_\Pi = 0$;
(5) $\kappa_{-X,Y} = \kappa_{X,-Y} = -\kappa_{X,Y}$;
(6) če sta C_1 in C_2 taki kopuli, da velja $C_1 \leq C_2$, potem je $\kappa_{C_1} \leq \kappa_{C_2}$;
(7) če je $\big((X_n, Y_n)\big)_n$ zaporedje zveznih slučajnih vektorjev s kopulami C_n in če zaporedje $(C_n)_n$ konvergira po točkah proti C, potem $\lim_{n\to\infty} \kappa_{C_n} = \kappa_C$.

Izrek 2.41. *Naj bo κ mera skladnosti za zvezni slučajni spremenljivki X in Y.*

(1) Če je Y skoraj gotovo naračajoča funkcija X, potem velja $\kappa_{X,Y} = \kappa_M = 1$.
(2) Če je Y skoraj gotovo padajoča funkcija X, potem velja $\kappa_{X,Y} = \kappa_W = -1$.
(3) Če sta funkciji α in β strogo monotoni na velja $\mathbf{I}$, potem $\kappa_{\alpha(X),\beta(Y)} = \kappa_{X,Y}$.

Bralec lahko dodatno razlago najde v [3] *na strani 169.*

Izrek 2.42. *Naj bosta X in Y zvezni slučajni spremenljivki s kopulo C. Potem τ in ρ zadoščata lastnostim za mero skladnosti iz zgornje definicije in posledično za njiju velja prejšnji izrek.*

Pri dokazu te trditve bi potreboval definicije in leme, ki jih v to delo nisem vključil. Bralec lahko dokaz najde v [3] na strani 169.

3. Kopule pri diskretnih slučajnih spremenljivkah

Podobno kot v zveznem bi si tudi v diskretnem primeru pri ocenjevanju odvisnosti med spremenljivkami radi pomagali s kopulami. V naslednjih poglavjih bomo prikazali, na kaj je treba biti pri tem pazljiv in kakšne so omejitve. Glavni vzrok, da pride do težav, je ta, da v diskretnem primeru s Sklarovim izrekom kopula ni enolično določena. To poglavje je povzeto po članku C. Genest in J.Nešlehová, A Primer on Copulas for Count Data, [2].

3.1. **Odsotnost enoličnosti in njene posledice.** V tem delu naj (X, Y) označuje par slučajnih spremenljivk s skupno porazdelitveno funkcijo H z robnima porazdelitvama F in G, torej $H(x, y) = P(X \leq x, Y \leq y)$, $F(x) = P(X \leq x)$ in $G(y) = P(Y \leq y)$. Kvazi inverza funkcij F in G sta definirana kot običajno, tj. $F^{(-1)}(u) = \inf\{x \in \mathbb{R} : F(x) \geq u\}$, $G^{(-1)}(v) = \inf\{y \in \mathbb{R} : G(y) \geq v\}$ in sta zvezna z leve za vsak $u, v \in (0, 1]$. Ker nista nujno zvezna z desne, definiramo še vrednosti

$$F^{(-1)}(u_{\leftarrow}) = \lim_{s \downarrow u} F^{(-1)}(s), \quad G^{(-1)}(v_{\leftarrow}) = \lim_{t \downarrow v} G^{(-1)}(t).$$

Podobno definiramo tudi levi limiti nezveznih porazdelitvenih funkcij F in G,

$$F(x_{\rightarrow}) = \lim_{s \uparrow x} F(s), \quad G(y_{\rightarrow}) = \lim_{t \uparrow y} G(t).$$

Sledeča definicija je uporabna za ponazoritev razlike med primerom, ko sta F in G zvezni funkciji, in primerom, ko nista. Povezavo med porazdelitvenimi funkcijami F, G in H lahko definiramo na različne načine.

Definicija 3.1. Naj bo (X, Y) slučajnih vektor, H njegova porazdelitvena funkcija ter F in G porazdelitveni funkciji za slučajni spremenljivki X in Y.

(i) $\mathcal{A}$ je množica vseh funkcij $A : \mathbf{I}^2 \to \mathbf{I}$, za katere velja, da je $A|_{Z_F \times Z_G}$ enaka podkopuli C', definirani s porazdelitveno funkcijo H, torej velja
$$H(x, y) = A(F(x), G(y)).$$
(ii) $B : [0, 1]^2 \to [0, 1]$ je funkcija, definirana za vsak $x, y \in [0, 1]$ s predpisom
$$B(u, v) = H(F^{(-1)}(u), G^{(-1)}(v)).$$
(iii) $C : [0, 1]^2 \to [0, 1]$ je funkcija, definirana za vsak $x, y \in [0, 1]$ s predpisom
$$C(u, v) = H(F^{(-1)}(u_{\leftarrow}), G^{(-1)}(v_{\leftarrow})).$$
(iv) D je porazdelitvena funkcija slučajnega vektorja $(F(X), G(Y))$.

(v) E je porazdelitvena funkcija slučajnega vektorja $(F(X_{\rightarrow}), G(Y_{\rightarrow}))$.

Ko sta tako F kot G zvezni, iz Sklarovega izreka vidimo, da so vsi zgoraj definirani objekti sorodni,

$$\mathcal{A} = \{B\} = \{C\} = \{D\}. \tag{1}$$

V tem primeru množica $\mathcal{A}$ vsebuje samo enolično kopulo, povezano s funkcijo H.

To prinaša pomembne posledice in je temelj modeliranja s kopulami:

(a) Naj bo tip funkcije H dan. Potem na C lahko gledamo kot na enolično podan (funkcijski) parameter, ki ga lahko ocenimo dosledno iz slučajnega vzorca $(X_1, Y_1), \ldots, (X_n, Y_n)$, ki ga dobimo s pomočjo funkcije H.
(b) Če poznamo F in G in imamo slučajni vzorec (X_i, Y_i) za $i = 1, \ldots, n$, lahko definiramo $U_i = F(X_i)$ in $V_i = G(Y_i)$ za vsak $i \in \{1, \ldots, n\}$. Na ta način dobimo slučajen vzorec $(U_1, V_1), \ldots, (U_n, V_n)$ za kopulo C.
(c) Če F in G ne poznamo, lahko še vedno sklepamo o C in sicer na podlagi psevdo opazovanj $\hat{U}_i = \hat{F}(X_i), \hat{V}_i = \hat{G}(Y_i)$, ki jih dobimo iz cenilk $\hat{F}$ in $\hat{G}$ za F in G.

(d) Naj bo za C dana parametrična družina (C_θ). Potem lahko ocenimo parameter θ ali naredimo test skladnosti (angl. *goodness-of-fit*), tako da uporabimo psevdo opazovanja $(\hat{U}_1, \hat{V}_1), \ldots, (\hat{U}_n, \hat{V}_n)$.

V diskretnem primeru pa naletimo na težave. Če imata F in G skoke, imata namreč njuna kvazi inverza intervale, kjer sta konstantna. V tem primeru Sklarov izrek še vedno zagotavlja obstoj kopule v množici $\mathcal{A}$, s katero lahko predstavimo H, a ta ni enolično določena. Posledično pride do težave z določenostjo in zgornji princip ne deluje več, saj identiteta 1, iz katere izhaja, ne velja več. Funkcije B, C in D namreč niso kopule, in samo nekatere izmed njih, a ne vse, pripadajo množici $\mathcal{A}$. Za ilustracijo si poglejmo naslednji primer.

Zgled 3.2. Naj bosta X in Y Bernoullijevi slučajni spremenljivki, $P(X=0)=p$, $P(X=1)=1-p$ in $P(Y=0)=q$, $P(Y=1)=1-q$. Naj velja $P(X=0, Y=0)=r$, kjer je $r \in [\max\{0, p+q-1\}, \min\{p, q\}]$.
Množica $\mathcal{A}$ vsebuje vse funkcije $A : \mathbf{I}^2 \to \mathbf{I}$, za katere velja

$$\begin{array}{lll} A(0,0)=0 & A(0,q)=0 & A(0,1)=0 \\ A(p,0)=0 & A(p,q)=r & A(p,1)=p \\ A(1,0)=0 & A(1,q)=q & A(1,1)=1. \end{array}$$

Pokažimo te enakosti za $A(0,0)$, $A(p,q)$ in $A(p,1)$. Vemo, da je $A(F(x), G(y)) = H(x,y)$.

- Da izračunamo vrednost funkcije A v točke $(0,0)$, moramo izbrati tako točko (x,y), da je $F(x)=0$ in $G(y)=0$. Ker sta X in Y porazdeljena Bernoullijevo, to pomeni, da sta x in y strogo manjša od 0. Denimo, da je $x=-1$ in $y=-2$. Potem je $H(-1,-2)=P(X \le -1, Y \le -2)=0$.
- Da izračunamo vrednost $A(p,q)$, moramo izbrati tako točko (x,y), da je $F(x)=p$ in $G(y)=q$. To pomeni, da sta x in y na intervalu $[0,1)$. V tem primeru je $H(x,y)=P(X \le x, Y \le y)=P(X=0, Y=0)$, to pa smo v navodilu tega zgleda označili z r.
- Da izračunamo vrednost $A(p,1)$, moramo izbrati tako točko (x,y), da je $F(x)=p$ in $G(y)=1$. To pomeni, da je x iz intervalu $[0,1)$, y pa je vsaj 1. V tem primeru je $H(x,y)=P(X \le x, Y \le y)=P(X=0,Y=0)+P(X=0,Y=1)=P(X=0)=p$.

Podobno izračunamo še ostale vrednosti.

Spomnimo se, da je funkcija B definirana kot $B(u,v)=H(F^{(-1)}(u), G^{(-1)}(v))$ in kvazi inverz funkcije F kot $F^{(-1)}(u)=\inf\{x \in \mathbb{R} : F(x) \ge u\}$.

- Če je $u=0$, potem je $F^{(-1)}(u)=F^{(-1)}(0)=-\infty$. Sledi
$$B(0,v)=H(-\infty, G^{(-1)}(v))=P(X \le -\infty, Y \le G^{(-1)}(v))=0.$$
Simetrično velja, če je $v=0$.
- Če je $(u,v) \in (0,p] \times (0,q]$, potem je $F^{(-1)}(u)=0$ in $G^{(-1)}(v)=0$. Zato je $H(0,0)=P(X=0,Y=0)=r$.
- Če je $(u,v) \in (p,1] \times (0,q]$, potem je $F^{(-1)}(u)=1$ in $G^{(-1)}(v)=0$. Zato je $H(1,0)=P(X \le 1, Y=0)=P(Y=0)=q$.

Podobno naredimo še za preostali možnosti in dobimo

$$B(u,v)=\begin{cases} 0, & \text{če } u=0 \text{ ali } v=0 \\ r, & \text{če } (u,v) \in (0,p] \times (0,q], \\ q, & \text{če } (u,v) \in (p,1] \times (0,q], \\ p, & \text{če } (u,v) \in (0,p] \times (q,1], \\ 1, & \text{če } (u,v) \in (p,1] \times (q,1]. \end{cases}$$

Funkcija C je definirana kot $C(u,v) = H(F^{(-1)}(u_\leftarrow), G^{(-1)}(v_\leftarrow))$.

- Če je $u \in [0,p)$ in $v \in [0,q)$, potem je $\lim_{s\downarrow u} F^{(-1)}(s) = 0$ in $\lim_{t\downarrow v} G^{(-1)}(t) = 0$, torej je $C(u,v) = H(0,0) = P(X=0, Y=0) = r$.
- Če je $u \in [0,p)$ in $v \in [q,1]$, potem je $\lim_{s\downarrow u} F^{(-1)}(s) = 0$ in $\lim_{t\downarrow v} G^{(-1)}(t) = 1$, torej je $C(u,v) = H(0,1) = P(X=0, Y\le 1) = P(X=0) = p$.
- Če je $u \in [p,1]$ in $v \in [0,q)$, potem je $\lim_{s\downarrow u} F^{(-1)}(s) = 1$ in $\lim_{t\downarrow v} G^{(-1)}(t) = 0$, torej je $C(u,v) = H(1,0) = P(X\le 1, Y=0) = P(Y=0) = q$.
- Če je $u \in [p,1]$ in $v \in [q,1]$, potem je $\lim_{s\downarrow u} F^{(-1)}(s) = 1$ in $\lim_{t\downarrow v} G^{(-1)}(t) = 1$, torej je $C(u,v) = H(1,1) = P(X\le 1, Y\le 1) = 1$.

Torej je

$$C(u,v) = \begin{cases} r, & \text{če } (u,v) \in [0,p) \times [0,q), \\ q, & \text{če } (u,v) \in [p,1] \times [0,q), \\ p, & \text{če } (u,v) \in [0,p) \times [q,1], \\ 1, & \text{če } (u,v) \in [p,1] \times [q,1]. \end{cases}$$

Funkcija D je porazdelitvena funkcija slučajnega vektroja $(F(X), G(Y))$, torej $D(u,v) = P\big(F(X) \le u, G(Y) \le v\big)$.

Slučajni spremenljivki X in Y sta porazdeljeni z zakonoma

$$X \sim \begin{pmatrix} 0 & 1 \\ p & 1-p \end{pmatrix}, \qquad Y \sim \begin{pmatrix} 0 & 1 \\ q & 1-q \end{pmatrix}.$$

Posledično sta porazdelitvi $F(X)$ in $G(Y)$ določeni z

$$F(X) \sim \begin{pmatrix} p & 1 \\ p & 1-p \end{pmatrix}, \qquad G(Y) \sim \begin{pmatrix} q & 1 \\ q & 1-q \end{pmatrix}.$$

- Če je $u \in [0,p)$, potem je funkcija $F(X)$ zagotovo večja od u, torej je $D(u,v) = 0$.
- Če je $v \in [0,q)$, potem je funkcija $G(Y)$ zagotovo večja od v, torej je $D(u,v) = 0$.
- Če je $(u,v) \in [p,1) \times [q,1)$, potem morata X in Y pokazati 0, torej je $D(u,v) = P(X=0, Y=0) = r$.
- Če je $u = 1$ in $v \in [q,1)$, potem mora Y pokazati 0, torej je $D(u,v) = q$.
- Če je $u \in [p,1)$ in $v = 1$, potem mora X pokazati 0, torej je $D(u,v) = p$.
- Če je $u = 1$ in $v = 1$, potem bo $F(X)$ v vsakem primeru manj od u in $G(Y)$ v vsakem primeru manj od v, torej je $D(u,v) = 1$.

$$D(u,v) = \begin{cases} 0, & \text{če } (u,v) \in [0,p) \times [0,1], \\ 0, & \text{če } (u,v) \in [0,1] \times [0,q), \\ r, & \text{če } (u,v) \in [p,1) \times [q,1), \\ p, & \text{če } (u,v) \in [p,1) \times \{1\}, \\ q, & \text{če } (u,v) \in \{1\} \times [q,1), \\ 1, & \text{če } (u,v) = (1,1). \end{cases}$$

Opazimo lahko, da je funkcija C porazdelitvena funkcija, medtem ko funkcija B ni, ker ni zvezna z desne. Poleg tega je B vsebovana v množici $\mathcal{A}$, C pa ni. V $\mathcal{A}$ je vsebovana tudi porazdelitvena funkcija D. $\diamondsuit$

Opažanja iz zgornjega primera veljajo v splošnem.

Trditev 3.3. *Naj bodo $\mathcal{A}$, B, C in D definirani kot zgoraj. Če je skupna porazdelitev slučajnega vektorja (X,Y) diskretna, potem velja*

(i) Funkcija $B \in \mathcal{A}$ ni porazdelitvena funkcija, zato tudi ni kopula.
(ii) Funkcija $D \in \mathcal{A}$ je porazdelitvena funkcija, ampak ni kopula.
(iii) Funkcija $C \notin \mathcal{A}$ je porazdelitvena funkcija, ampak prav tako ni kopula.

Dokaz. Iz Sklarovega izreka dobimo, da je $B \in \mathcal{A}$. Poleg tega B ni porazdelitvena funkcija, saj ni zvezna z desne, ker funkciji $F^{(-1)}$ in $G^{(-1)}$ nista zvezni z desne.
Dokažimo, da je $C = E$. Iz definicije funkcije $F^{(-1)}$ vidimo, da ekvivalenca

$$u \geq F(x_{\to}) \Leftrightarrow F^{(-1)}(u_{\leftarrow}) \geq x$$

drži za vse $u \in (0,1)$ in $x \in \mathbb{R}$. Podobno velja tudi za funkcijo G. Sedaj izračunamo

$$\begin{aligned} C(u,v) &= H(F^{(-1)}(u_{\leftarrow}), G^{(-1)}(v_{\leftarrow})) = \\ &= P(X \leq F^{(-1)}(u_{\leftarrow}), Y \leq G^{(-1)}(v_{\leftarrow})) = \\ &= P(u \geq F(X_{\to}), v \geq G(Y_{\to})) = \\ &= P(F(X_{\to}) \leq u, G(Y_{\to}) \leq v) = \\ &= E(u,v). \end{aligned}$$

D in E sta po definiciji porazdelitveni funkciji. Ker sta X in Y diskretni slučajni spremenljivki, so take tudi funkcije $F(X)$, $F(X_{\to})$, $G(Y)$ in $G(Y_{\to})$. Posledično niso enakomerno porazdeljene na intervalu $[0,1]$, tako da D in E ne moreta biti kopuli.
Pokažimo še, da je $D \in \mathcal{A}$. Pokazati moramo, da je

$$D(F(x), G(y)) = P\big(F(X) \leq F(x), G(Y) \leq G(y)\big) = H(x,y).$$

To velja, ker za vsak $x \in \mathbb{R}$ velja $P\big(F(X) \leq F(x)\big) = P(X \leq x)$ (kljub temu, da ekvivalenca $u \leq x \Leftrightarrow F(u) \leq F(x)$ ne velja). Na enak način naredimo še za funkcijo G.
Za konec še pokažimo, da funkcija C ni element razreda $\mathcal{A}$. Da bi bila, bi moralo veljati, $H(x,y) = C(F(x), G(y))$. Iz definicije funkcije C dobimo

$$C(F(x), G(y)) = H(F^{(-1)}(F(x)_{\leftarrow}), G^{(-1)}(G(y)_{\leftarrow})).$$

Sedaj si pomagamo z vrednostmi, definiranimi v začetku poglavja in dobimo

$$C(F(x), G(y)) = H(\lim_{s \downarrow F(x)} F^{(-1)}(s), \lim_{t \downarrow G(y)} G^{(-1)}(t)),$$

kar pa ni enako $H(x,y)$. □

Kot smo videli, v primeru, ko je vsaj ena izmed porazdelitvenih funkcij F in G diskretna, nastopijo težave. Obseg le teh si bomo pogledali v naslednjem poglavju.

4. Kako huda težava je nedoločenost?

Nedoločenost je lastnost, da dobljenega parametričnega modela ne moremo opisati z enoličnim naborom parametrov.

Naj bo H skupna porazdelitvena funkcija slučajnega vektorja (X,Y) ter $C_H \subset \mathcal{A}$ razred kopul, za katere velja enakost $H(x,y) = C(F(x), G(y))$. Če je ta razred dovolj majhen, potem lahko s katerokoli kopulo, ki pripada H, precej dobro opišemo lastnosti odvisnosti funkcije H. Da ugotovimo pomembnost težav z nedoločenostjo, moramo najprej izračunati velikost razreda C_H. To naredimo tako, da poiščemo najmanjši in največji element v C_H, potem pa izračunamo ustrezno mero odvisnosti za obe meji ter ju primerjamo. Uporabili bomo meri Kendallov τ ali Spearmanov ρ, ki smo ju definirali zgoraj.

Kot bomo videli, je razred C_H v splošnem razmeroma velik, zaradi česar je nedoločenost huda težava. To poglavje je povzeto po C. Genest in J. Nešlehová, A Primer on Copulas for Count Data, [2].

4.1. **Carleyini meji za C_H.** Kadar spremenljivki X in Y zavzameta le končno mnogo vrednosti, lahko najdemo spodnjo in zgornjo mejo za razred C_H, t.j. kopuli C_H^- in C_H^+, za kateri velja, da $C_H^-(x,y) \leq C(x,y) \leq C_H^+(x,y)$ za vsako funkcijo $C \in C_H$ in vsak par $(x,y) \in \mathbf{I}^2$.

Brez škode za splošnost tako predpostavimo, da X in Y zavzameta le celoštevilske nenegativne vrednosti. Naj bo

$$h_{ij} = P(X = i, Y = j), \quad H(i,j) = P(X \leq i, Y \leq j),$$

torej je h_{ij} masa, ki jo C dodeli pravokotniku $(F(i-1), F(i)] \times (G(j-1), G(j)]$, pri čemer je $F(-1) = G(-1) = 0$. Označimo še robni verjetnostni funkciji spremenljivk X in Y,

$$h_{i+} = \sum_{j=0}^{\infty} h_{ij} = P(X = i), \quad h_{+j} = \sum_{i=0}^{\infty} h_{ij} = P(Y = j).$$

Trditev 4.1. *Naj bo C_H razred kopul, ki zadoščajo enačbi $H(x,y) = C(F(x), G(y))$ za diskretno porazdelitev H na $\mathbb{N}^2$. Zgornja Carleyina meja za razred C_H je podana z enačbo*

$$C_H^+(u,v) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \max\{0, \min\{u - \alpha_{ij}, v - \beta_{ij}, h_{ij}\}\},$$

spodnja pa z enačbo

$$C_H^-(u,v) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \max\{0, -h_{ij} + \min\{u - \gamma_{ij}, h_{ij}\} + \min\{v - \delta_{ij}, h_{ij}\}\}.$$

Tukaj so $\alpha_{ij}, \beta_{ij}, \gamma_{ij}$ in δ_{ij} definirani kot

$$\alpha_{ij} = \sum_{k=0}^{i-1} h_{k+} + \sum_{l=0}^{j-1} h_{il}, \quad \beta_{ij} = \sum_{l=0}^{j-1} h_{+l} + \sum_{k=0}^{i-1} h_{kj},$$

$$\gamma_{ij} = \sum_{k=0}^{i-1} h_{k+} + \sum_{l=j+1}^{\infty} h_{il}, \quad \delta_{ij} = \sum_{l=0}^{j-1} h_{+l} + \sum_{k=i+1}^{\infty} h_{kj}$$

in velja, da je prazna vsota enaka 0.

Dokaz. Dokaz je v [1]. □

Za lažjo predstavo zgoraj definiranih koeficientov si poglejmo naslednji zgled.

Zgled 4.2. Naj bo (X,Y) slučajni vektor, ki zavzeme vrednosti na $\{0,1,2,3\} \times \{0,1,2\}$ in je porazdeljen kot je prikazano v sledeči tabeli.

TABELA 1. Verjetnosti, da slučajni vektor (X,Y) zavzame posamezne vrednosti.

	$Y=0$	$Y=1$	$Y=2$	skupaj
$X=0$	$\frac{1}{24}$	$\frac{1}{24}$	$\frac{1}{24}$	$\frac{3}{24}$
$X=1$	$\frac{1}{24}$	$\frac{3}{24}$	$\frac{3}{24}$	$\frac{6}{24}$
$X=2$	$\frac{3}{24}$	$\frac{6}{24}$	$\frac{3}{24}$	$\frac{12}{24}$
$X=3$	$\frac{1}{24}$	$\frac{2}{24}$	0	$\frac{3}{24}$
skupaj	$\frac{6}{24}$	$\frac{12}{24}$	$\frac{6}{24}$	$\frac{24}{24}$

V naslednjih tabelah so izračunani α_{ij}, β_{ij}, γ_{ij} in δ_{ij}.

TABELA 2. Izračun α_{ij}, β_{ij}, γ_{ij} in δ_{ij} glede na porazdelitev slučajnega vektorja (X, Y), ki je podana v tabeli 1.

α_{ij}	$j=0$	$j=1$	$j=2$
$i=0$	0	$\frac{1}{24}$	$\frac{2}{24}$
$i=1$	$\frac{3}{24}$	$\frac{4}{24}$	$\frac{7}{24}$
$i=2$	$\frac{9}{24}$	$\frac{12}{24}$	$\frac{18}{24}$
$i=3$	$\frac{21}{24}$	$\frac{22}{24}$	$\frac{24}{24}$

β_{ij}	$j=0$	$j=1$	$j=2$
$i=0$	0	$\frac{6}{24}$	$\frac{18}{24}$
$i=1$	$\frac{1}{24}$	$\frac{7}{24}$	$\frac{19}{24}$
$i=2$	$\frac{2}{24}$	$\frac{10}{24}$	$\frac{21}{24}$
$i=3$	$\frac{5}{24}$	$\frac{16}{24}$	$\frac{24}{24}$

γ_{ij}	$j=0$	$j=1$	$j=2$
$i=0$	$\frac{2}{24}$	$\frac{1}{24}$	0
$i=1$	$\frac{9}{24}$	$\frac{5}{24}$	$\frac{3}{24}$
$i=2$	$\frac{18}{24}$	$\frac{12}{24}$	$\frac{9}{24}$
$i=3$	$\frac{23}{24}$	$\frac{21}{24}$	$\frac{21}{24}$

δ_{ij}	$j=0$	$j=1$	$j=2$
$i=0$	$\frac{5}{24}$	$\frac{17}{24}$	$\frac{23}{24}$
$i=1$	$\frac{4}{24}$	$\frac{14}{24}$	$\frac{21}{24}$
$i=2$	$\frac{1}{24}$	$\frac{8}{24}$	$\frac{18}{24}$
$i=3$	0	$\frac{6}{24}$	$\frac{18}{24}$

$\diamondsuit$

Sedaj, ko imamo Carleyini meji, lahko izrazimo razpon mer odvisnosti, ki ga določa razred C_H.

4.2. **Ekstremi za τ in ρ v C_H.** Če je κ mera skladnosti, po točki 6 v definiciji 2.40 velja

$$\kappa(C_H^-) \leq \kappa(C) \leq \kappa(C_H^+).$$

Ker kopuli C_H^- in C_H^+ pripadata C_H, sta ti meji najboljši možni. Poleg tega so vse točke v intervalu $[\kappa(C_H^-), \kappa(C_H^+)]$ možne vrednosti za $\kappa(C)$, saj za vsak $\theta \in [0,1]$ velja

$$C_\theta = \theta C_H^- + (1-\theta)C_H^+ \in C_H,$$

$\kappa(C_\theta)$ pa je zvezna funkcija parametra θ (sledi iz sedme točke definicije Scarsinija).

Kendallov τ in Spearmanov ρ izračunamo s pomočjo naslednjih dveh trditev. Koeficienti α_{ij}, β_{ij}, γ_{ij} in δ_{ij} naj bodo definirani tako kot v trditvi 4.1.

Trditev 4.3. *Naj bo H porazdelitvena funkcija na $\mathbb{N}^2$ z verjetnostmi h_{ij} za vsak $i, j \in \mathbb{N}$. Potem velja*

$$\tau(C_H^-) = -1 + 4\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{i-1}\sum_{l=0}^{j-1} h_{ij}h_{kl}$$

in

$$\tau(C_H^+) = 1 - 4\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{i-1}\sum_{l=j+1}^{\infty} h_{ij}h_{kl}.$$

Dokaz. Preverimo trditev najprej za $\tau(C_H^-)$. Za začetek moramo ugotoviti, kaj je nosilec funkcije C_H^-. Najprej ugotovimo, da točke nimajo pozitivne mase, saj v tem primeru robna porazdelitev ne bi bila enakomerna, kar je v nasprotju z definicijo kopule. Iz podobnega razloga sledi, da tudi vodoravne in navpične daljice nimajo pozitivne mase. Vemo še, da je funkcija C_H^- odsekoma linearna, zato je, v točkah, kjer obstaja, mešani odvod od C_H^-, t.j. $(C_H^-)_{uv}$, enak 0. Posledično nosilec ne vsebuje množice s pozitivno Lebesguovo mero.

Torej je nosilec kopule C vsebovan v množici, kjer C_H^- ni odvedljiva. To je tam, kjer se min ali max prelomita. Sedaj pogledamo funkcijo v točkah, kjer se funkcija C prelomi, v našem primeru v točkah $u-\gamma_{ij}=h_{ij}$ in $v-\delta_{ij}=h_{ij}$. Dobimo $u=h_{ij}+\gamma_{ij}$, torej navpično daljico, oziroma $v=h_{ij}+\delta_{ij}$, kar je vodoravna daljica. Zato tu C_H^- nima mase. Sedaj preverimo možnost, da je $\min\{u-\gamma_{ij},h_{ij}\}+\min\{v-\delta_{ij},h_{ij}\}=h_{ij}$. Če je $u>\gamma_{ij}+h_{ij}$, potem sledi $v=\delta_{ij}$. Zopet dobimo vodoravno daljico, torej na njej ni mase. Če je $u<\gamma_{ij}$, enakost ni mogoča. Podobno velja tudi za spremenljivko v. Preverimo še primer, ko je $\gamma_{ij}\leq u\leq\gamma_{ij}+h_{ij}$ in $\delta_{ij}\leq u\delta_{ij}+h_{ij}$. Iz enačbe $\min\{u-\gamma_{ij},h_{ij}\}+\min\{v-\delta_{ij},h_{ij}\}=h_{ij}$ dobimo $u-\gamma_{ij}+v-\delta_{ij}=h_{ij}$. Ko uredimo, vidimo, da je spremenljivka v linearna funkcija spremenljivke u, podana z enačbo $v=h_{ij}+\delta_{ij}+\gamma_{ij}-u$. Torej je masa za posamezen par (i,j) zbrana na daljici med točkama $(\gamma_{ij},\delta_{ij}+h_{ij})$ in $(\gamma_{ij}+h_{ij},\delta_{ij})$. Ko i,j tečeta, te daljice ravno pokrijejo interval $\mathbf{I}$.

Posledično sem nam poenostavi integral v enačbi za Q,

$$\int_0^1\int_0^1 C_H^-(u,v)dC_H^-(u,v)=\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\int_{\gamma_{ij}}^{\gamma_{ij}+h_{ij}} C_H^-(u,h_{ij}\delta_{ij}+\gamma_{ij}-u)du.$$

Poleg tega za vsak $u\in[\gamma_{ij},\gamma_{ij}+h_{ij}]$ velja tudi enačba

$$C_H^-(u,h_{ij}+\delta_{ij}+\gamma_{ij}-u)=\sum_{k=0}^{i-1}\sum_{l=0}^{j-1}h_{kl}.$$

To uporabimo v integralu in dobimo

$$\iint_{\mathbf{I}^2} C_H^-(u,v)dC_H^-(u,v)=\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\int_{\gamma_{ij}}^{\gamma_{ij}+h_{ij}}\sum_{k=0}^{i-1}\sum_{l=0}^{j-1}h_{kl}du=\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{i-1}\sum_{l=0}^{j-1}h_{ij}h_{kl}.$$

Ko dobljen rezultat ustavimo v formlu za τ, dobimo

$$\tau(C_H^-)=-1+4\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{i-1}\sum_{l=0}^{j-1}h_{ij}h_{kl}.$$

Podobno postopamo tudi v primeru dokazovanja neenakosti za $\tau(C_H^+)$. □

Trditev 4.4. *Naj bo H porazdelitvena funkcija na $\mathbb{N}^2$ z verjetnostmi h_{ij} za vsak $i,j\in\mathbb{N}$. Potem velja*

$$\rho(C_H^-)=-1-6\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}h_{ij}\Big(\delta_{ij}+h_{ij}-(1-\gamma_{ij})\Big)\Big(2(1-\gamma_{ij})-h_{ij}\Big)$$

in

$$\rho(C_H^+)=1+6\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}h_{ij}(\beta_{ij}-\alpha_{ij})(2\alpha_{ij}+h_{ij}).$$

Dokaz te trditve je podoben dokazu trditve 4.3, zato ga v delo nisem vključil. Bralec ga lahko najde v [2] na strani 509.

Kot prikazuje naslednji primer, je lahko razlika med Carleyinima mejama na Kendallovem ρ in Spearmanovem τ precej velika.

Zgled 4.5. Naj bosta X in Y Bernoullijevi slučajni spremenljivki, $P(X=0)=p$, $P(Y=0)=q$. Naj bo še $r=P(X=0,Y=0)\in[\max\{0,p+q-1\},\min\{p,q\}]$. Potem lahko s kopulo C modeliramo odvisnost med X in Y natanko tedaj, ko je $C(p,q)=r$.

Izračunamo h_{ij} za $i = 0, 1$ in $j = 0, 1$.

$$\begin{aligned}
h_{00} &= P(X = 0, Y = 0) = r \\
h_{01} &= P(X = 0, Y = 1) = P(X = 0) - P(X = 0, Y = 0) = p - r \\
h_{10} &= P(X = 1, Y = 0) = P(Y = 0) - P(X = 0, Y = 0) = q - r \\
h_{11} &= P(X = 1, Y = 1) = 1 - (p - r) - (q - r) - r = 1 - p - q + r
\end{aligned}$$

Izračunamo še $\tau(C_H^-)$ in $\tau(C_H^+)$.

$$\begin{aligned}
\tau(C_H^-) &= -1 + 4 \sum_{i=0}^{1} \sum_{j=0}^{1} \sum_{k=0}^{i-1} \sum_{l=0}^{j-1} h_{ij} h_{kl} = -1 + 4\, h_{11} h_{00} \\
\tau(C_H^+) &= 1 - 4 \sum_{i=0}^{1} \sum_{j=0}^{1} \sum_{k=0}^{i-1} \sum_{l=j+1}^{1} h_{ij} h_{kl} = 1 - 4\, h_{10} h_{01}
\end{aligned}$$

Iz trditve 4.3. tako dobimo

$$4rs - 1 \le \tau(C) \le 1 - 4(p - r)(q - r),$$

kjer je $s = 1 - p - q + r = P(X = 1, Y = 1)$.

Podobno izračunamo še meji za $\rho(C)$. S pomočjo formul iz trditve 4.1 izračunamo α_{ij}, β_{ij}, γ_{ij} in δ_{ij}.

$$\begin{aligned}
\alpha_{00} &= 0 \\
\alpha_{01} &= h_{00} = r \\
\alpha_{10} &= h_{00} + h_{01} = r + p - r = p \\
\alpha_{11} &= h_{00} + h_{01} + h_{10} = r + p - r + q - r = p + q - r
\end{aligned}$$

$$\begin{aligned}
\beta_{00} &= 0 \\
\beta_{01} &= h_{00} + h_{10} = r + q - r = q \\
\beta_{10} &= h_{00} = r \\
\beta_{11} &= h_{00} + h_{01} + h_{10} = r + p - r + q - r = p + q - r
\end{aligned}$$

$$\begin{aligned}
\gamma_{00} &= h_{01} = p - r \\
\gamma_{01} &= 0 \\
\gamma_{10} &= h_{00} + h_{01} + h_{11} = r + p - r + s = p + s \\
\gamma_{11} &= h_{00} + h_{01} = r + p - r = p
\end{aligned}$$

$$\begin{aligned}
\delta_{00} &= h_{10} = q - r \\
\delta_{01} &= h_{00} + h_{10} + h_{11} = r + q - r + s = q + s \\
\delta_{10} &= 0 \\
\delta_{11} &= h_{00} + h_{10} = r + q - r = q
\end{aligned}$$

S pomočjo formule iz trditve 4.4 izračunamo $\rho(C_H^+)$ in $\rho(C_H^-)$.

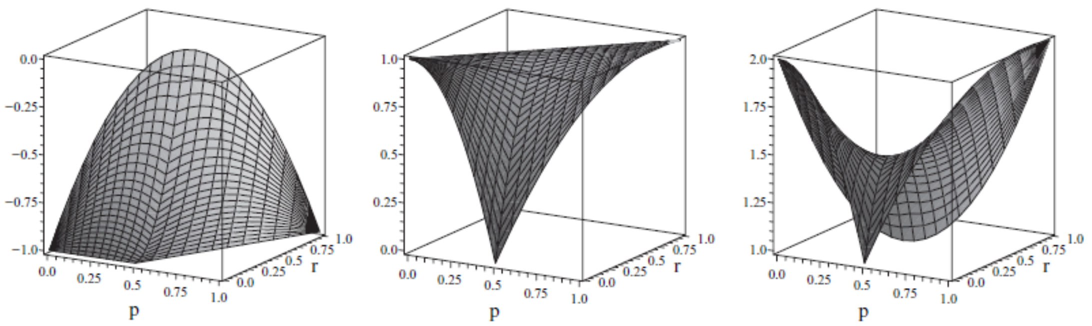
$$
\begin{aligned}
\rho(C_H^+) &= 1 + 6\big(0 + (p-r)(q-r)(2r+p-r) + (q-r)(r-p)(2p+q-r) + 0\big) = \\
&= 1 + 6(p-r)(q-r)(r+p-2p-q+r) = \\
&= 1 + 6(p-r)(q-r)(2r-p-q) = \\
&= 1 - 6(p-r)(q-r)(1-r-s)
\end{aligned}
$$

$$
\begin{aligned}
\rho(C_H^-) = -1 - 6 \cdot \Big(& r(q-r+r-1+p-r)(2-2p+2r-r) + \\
& + (p-r)(q+s+p-r-1)(2-p+r) + \\
& + (q-r)(q-r-1+p+s)(2-2p-2s-q+r) + \\
& + s(q+s-1+p)(2-2p-s)\Big) = \\
= -1 - 6\Big(& r(-1+p+q-r)(2-2p+r) + 0 + 0 + sr(2-2p-s)\Big) = \\
= -1 - 6\Big(& rs(-2+2p-r) + rs(2-2p-s)\Big) = \\
= -1 - 6&rs(-r-s) = \\
= -1 + 6&rs(r+s)
\end{aligned}
$$

Iz trditve 4.4 tako dobimo

$$6rs(r+s) - 1 \le \rho(C) \le 1 - 6(p-r)(q-r)(1-r-s).$$

Na spodnji sliki so za poseben primer, ko je $p = q$, prikazani meji za τ kot funkcija p in r, skupaj z razliko med zgornjo in spodnjo mejo. V tem primeru velja, da r leži v intervalu $[\max\{0, 2p-1\}, p]$. Spodnja meja je enako $4r(1-2p+r) - 1$, zgornja meja je enako $1 - 4(p-r)^2$, njuna razlika pa znaša $2 - 4p^2 + 16pr - 8r^2$.



Slika 1. Graf spodnje meje $\tau(C_H^-)$, zgornje meje $\tau(C_H^+)$ in razlike med njima, ko sta slučajni spremenljivki X in Y porazdeljeni Bernoulijevo, $P(X = 0) = p$, $P(Y = 0) = q$ in je $p = q$.

Iz grafov opazimo:

(i) Spodnja meja je enako 0 natanko tedaj, ko je $p = q = r = \frac{1}{2}$, torej ko je $X = Y$ skoraj gotovo.

(ii) Zgornja meja je enaka 0 natanko tedaj, ko je $p = q = \frac{1}{2}$ in $r = 0$, torej ko je $Y = 1-X$ skoraj gotovo.
(iii) Razlika med obema mejama je vedno vsaj 1.

Grafi za ρ so podobne oblike, a z večjim razponom zaloge vrednosti. Spodnja meja ima vrednosti na intervalu $[-1, 0.5]$, zgornja na intervalu $[0.5, 2]$, tako da je razlika med njima med 0.5 in 2. $\diamondsuit$

Kot vidimo, se lahko zgornja in spodnja meja za τ (ali ρ) zelo razlikujeta, zaradi česar je nedoločenost kopule hud problem.

Opomba 4.6. Geometrijska porazdelitev je porazdelitev števila poskusov do vključno prvega uspelega, če izvajamo Bernoullijevo zaporedje poskusov, pri katerih vsak poskus uspe z verjetnostjo p. Če je slučajna spremenljvika X porazdeljena geometrijsko z verjetnostjo uspeha p, potem velja $P(X = k) = p(1-p)^k$ za k iz množice $\{0, 1, 2, 3 \dots\}$.

Zgled 4.7. Naj bosta X in Y neodvisni geometrijsko porazdeljeni slučajni spremenljivki na $\mathbb{N}$ z verjetnostjo uspeha $1-p$ oziroma $1-q$. Produktna kopula Π je kopula za slučajni vektor (X, Y), a ni edina. Naj bo C možna kopula za par (X, Y). Če označimo še $h_{ij} = P(X = i, Y = j) = (1-p)p^i(1-q)q^j$, dobimo iz trditve 4.3 neenakost

$$|\tau(C)| \leq 1 - 4\frac{p}{1+p}\frac{q}{1+q}.$$

Pokažimo neenakost najprej za $\tau(C^-)$.

$$\begin{aligned}
\tau(C^-) &= -1 + 4\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{i-1}\sum_{l=0}^{j-1} h_{ij}h_{kl} = \\
&= -1 + 4\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{i-1}\sum_{l=0}^{j-1}(1-p)p^i(1-q)q^j(1-p)p^k(1-q)q^l = \\
&= -1 + 4(1-p)^2(1-q)^2\sum_{i=0}^{\infty}p^i\sum_{j=0}^{\infty}q^j\sum_{k=0}^{i-1}p^k\sum_{l=0}^{j-1}q^l = \\
&= -1 + 4(1-p)^2(1-q)^2\sum_{i=0}^{\infty}p^i\sum_{j=0}^{\infty}q^j\frac{1-p^i}{1-p}\frac{1-q^j}{1-q} = \\
&= -1 + 4(1-p)(1-q)\sum_{i=0}^{\infty}p^i(1-p^i)\sum_{j=0}^{\infty}q^j(1-q^j) = \\
&= -1 + 4(1-p)(1-q)\sum_{i=0}^{\infty}(p^i - p^{2i})\sum_{j=0}^{\infty}(q^j - q^{2j}) = \\
&= -1 + 4(1-p)(1-q)(\frac{1}{1-p} - \frac{1}{1-p^2})(\frac{1}{1-q} - \frac{1}{1-q^2}) = \\
&= -1 + 4(1-p)(1-q)(\frac{1+p}{(1-p)(1+p)} - \frac{1}{(1-p)(1+p)})\cdot \\
&\qquad \cdot(\frac{1+q}{(1-q)(1+q)} - \frac{1}{(1-q)(1+q)}) = \\
&= -1 + 4(1-p)(1-q)\frac{p}{(1-p)(1+p)}\frac{q}{(1-q)(1+q)} = \\
&= -1 + 4\frac{p}{1+p}\frac{q}{1+q}
\end{aligned}$$

Na enak način naredimo še za $\tau(C^+)$.

$$\begin{aligned}
\tau(C^+) &= 1-4\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{i-1}\sum_{l=j-1}^{\infty} h_{ij}h_{kl} = \\
&= 1-4\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{i-1}\sum_{l=j-1}^{\infty} (1-p)p^i(1-q)q^j(1-p)p^k(1-q)q^l = \\
&= 1-4(1-p)^2(1-q)^2\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{i-1}\sum_{l=j-1}^{\infty} p^iq^jp^kq^l = \\
&= 1-4(1-p)^2(1-q)^2\sum_{i=0}^{\infty}p^i\sum_{j=0}^{\infty}q^j\sum_{k=0}^{i-1}p^k\sum_{l=j-1}^{\infty}q^l = \\
&= 1-4(1-p)^2(1-q)^2\sum_{i=0}^{\infty}p^i\sum_{j=0}^{\infty}q^j\frac{1-p^i}{1-p}q^{j-1}\frac{1}{1-q} = \\
&= 1-4(1-p)(1-q)\sum_{i=0}^{\infty}p^i(1-p^i)\sum_{j=0}^{\infty}q^jq^{j-1} = \\
&= 1-4(1-p)(1-q)\sum_{i=0}^{\infty}(p^i-p^{2i})\sum_{j=0}^{\infty}q^{2j}q = \\
&= 1-4(1-p)(1-q)\left(\frac{1}{1-p}-\frac{1}{1-p^2}\right)\frac{1}{1-q^2}q = \\
&= 1-4(1-p)(1-q)\frac{1+p-1}{(1-p)(1+p)}\frac{1}{(1-q)(1+q)}q = \\
&= 1-4\frac{p}{1+p}\frac{q}{1+q} =
\end{aligned}$$

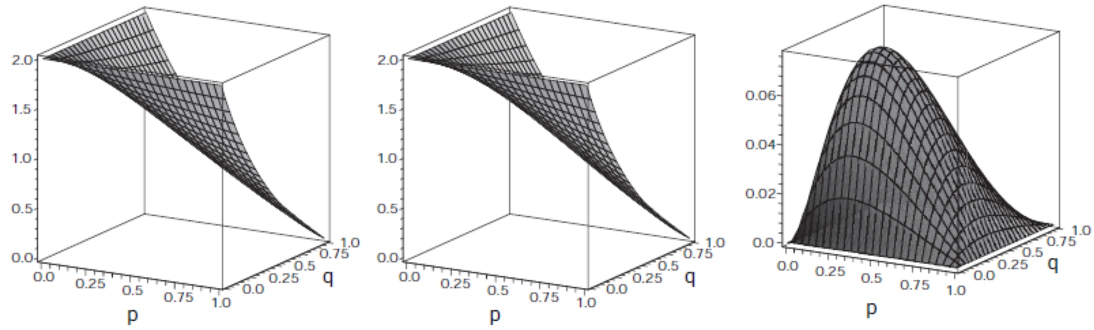
Podobno iz trditve 4.4. dobimo neenakosti še za ρ:

$$\rho(C) \leq 1-6pq\frac{p^2q+pq^2+p^2+q^2+p+q}{(1+p)(1+q)(p^2+p+1)(q^2+q+1)}$$

in

$$\rho(C) \geq -1+6pq\frac{p^2q+pq^2+p^2+q^2+pq+1}{(1+p)(1+q)(p^2+p+1)(q^2+q+1)}.$$

Vpeljemo oznaki $T_H = \tau(C_H^+)-\tau(C_H^-) = 2-8\dfrac{pq}{(1+p)(1+q)}$ in
$R_H = \rho(C_H^+)-\rho(C_H^-) = 2-6\dfrac{pq(p+q-pq-1)}{(1+p)(1+q)(p^2+p+1)(q^2+q+1)}$.



SLIKA 2. Grafi za (od leve proti desni) T_H, R_H in za razliko $T_H - R_H$ v primeru neodvisnih geometrijsko porazdeljenih slučajnih spremenljivk.

Opazimo, da sta funkciji T_H in R_H padajoči v p in q. Ko je verjetnost uspeha enaka 0, torej ko sta p in q enaka 1, sta tako T_H kot R_H enaka 0. Iz desnega grafa vidimo, da je T_H vedno večji od R_H, od koder sledi, da sta meji za τ ožji od mej za ρ za vse vrednosti p in q. $\diamondsuit$

5. Zaključek

Kot smo videli v prejšnjih poglavjih, si lahko s kopulami v primeru zveznih porazdelitvenih funkcij veliko pomagamo pri ugotavljanju odvisnosti med slučajnimi spremenljivkami. V tem primeru je namreč kopula, s katero lahko predstavimo skupno porazdelitveno funkcijo, enolična. Poleg tega smo za primer 2-dimenzionalne kopule videli, da lahko kopulo konstruiramo iz skupne porazdelitvene funkcije s pomočjo inverzov porazdelitvenih funkcij.

Ugotovili smo, da v primeru, ko je vsaj ena od porazdelitvenih funkcij slučajnih spremenljivk diskretna, veliko lepih lastnosti kopul izgine in da nastanejo težave. Velikost teh je odvisna od velikosti razreda C_H, ki jo izračunamo s pomočjo Carleyine spodnje in zgornje meje ter z merami skladnosti. V zadnjih dveh zgledih pa smo videli še, da je lahko razlika med Carleyinima mejama precej velika.

Velja, da mere skladnosti v diskretnem primeru postanejo odvisne od robnih porazdelitev. Izkaže se namreč, da je edina mera skladnosti, ki je neodvisna od robnih porazdelitev, skalar, ta pa ni posebej uporabna. V zveznem primeru velja, da v primeru popolne monotone odvisnosti velja $|\tau| = |\rho| = 1$. Kot kaže primer v [2] na strani 492, to v diskretnem primeru ne velja več. Vidimo torej, da sta poleg tega, da sta od robnih porazdelitev odvisna $\tau(X,Y)$ in $\rho(X,Y)$, od njih odvisni tudi njuni zalogi vrednosti.

Da bi se izognili tem problemom, bi lahko τ in ρ normalizirali v τ_b in ρ_s (definirani sta v [2] na strani 492) in na ta način dobili primernejši meri. S tem bi njuno zalogo vrednosti povečali, a bi se še vedno lahko zgodilo, da bi bili X in Y v popolni monotoni odvisnosti, vseeno pa bi veljajo $|\tau_b(X,Y)| < 1$ in $|\rho_s(X,Y)| < 1$.

Kot smo videli, so lahko težave z nedoločenostjo velike, zaradi česar je potrebno biti pri uporabi kopul v diskretnem primeru previden.

Slovar strokovnih izrazov

copula kopula
subcopula podkopula
H-volume H-ploščina
grounded prizemljen
concordance skladnost
concordance measures mere skladnosti
unidentifiability nedoločenost
inference sklepanje
uniqueness enoličnost

Literatura

[1] H. Carley, *Maximum and minimum extensions of finite subcopulas*, Communications in Statistics - Theory and Methods **Volume 31 Issue 12** (2002) 2151-2166; dostopno tudi na `https://www.tandfonline.com/doi/full/10.1081/STA-120017218?scroll=top&needAccess=true`.

[2] C. Genest in J. Nešlehová, *A Primer on Copulas for Count Data*, ASTIN Bulletin **Volume 37 Issue 2** (2007) 475-515; dostopno tudi na `https://www.cambridge.org/core/services/aop-cambridge-core/content/view/B1857CBA9F03491F6CD41B85A553E977/S0515036100014963a.pdf/primer_on_copulas_for_count_data.pdf`.

[3] R. Nelsen, *An Introduction to Copulas*, Springer Series in Statistics, Springer-Verlag New York, New York, 2006.

[4] M. Scarsini, *On measures of concordance*, Stochastica **Volume 8 Issue 3** (1984) 201-218; dostopno tudi na `http://dmle.icmat.es/pdf/STOCHASTICA_1984_08_03_01.pdf`.