UNIVERZA V LJUBLJANI
FAKULTETA ZA MATEMATIKO IN FIZIKO

Finančna matematika – 1. stopnja

Sara Radej

Kopule za večrazsežno Studentovo porazdelitev

Delo diplomskega seminarja

Mentor: doc. dr. Aljoša Peperko

Ljubljana, 2014

Kazalo

Kopule za večrazsežno Studentovo porazdelitev

Povzetek

V delu diplomskega seminarja je predstavljena Studentova T porazdelitev in njene osnovne lastnosti. Prav tako je podana definicija kopule z osnovnimi lastnostmi. V teoriji kopul je temeljnega pomena Sklarov izrek, tako teoretično, kot tudi aplikativno v statističnem modeliranju. Obravnavamo različne korelacijske koeficiente: Pearsonov korelacijski koeficient, Spearmanov korelacijski koeficient ranga in Kendallov tau.

Prikazani sta Studentova in normalna kopula. Ti dve kopuli sta pogosto uporabljeni, ker so simulacije na njih dovolj enostavne. Obravnavani so različni pristopi konstruiranja Studentove večrazsežne porazdelitve in posledično Studentove kopule, kjer so uporabljene enake marginalne porazdelitve in enake prostostne stopnje, nato pa še marginalne porazdelitve z različnimi prostostnimi stopnjami. Opisana je tudi konstrukcija Student-normalne porazdelitve. Studentova in normalna kopula sta primerjani na zgledu konkretnih finančnih podatkov.

Copulas for multivariate Student distribution

Abstract

In diploma thesis the Student T distribution with its basic characteristics is shown. The definition of copula with its basic characteristics is added as well. The Sklar's theorem is of immense importance in statistical modeling-theoretically and applicatively. Different correlation coefficients, namely Pearson's correlation coefficient, Spearman's rank correlation coefficient and Kendall tau are discussed.

Student and normal copula are shown. These two copulas are often used due to simple simulations that can be performed on them. Different approaches to Student multivariate distribution and consequently Student copula are discussed. Equal marginals and equal degrees of freedom, as well as marginal distributions with unequal degrees of freedom are used. The construction of Student-normal distribution is also described. Actual financial data set is used as an example for comparison of Student and normal copula.

Math. Subj. Class. (2010): 60E05, 62P20.

Ključne besede: kopule, Sklarov izrek, Studentova kopula, normalna kopula, mera konkordance, statistično modeliranje

Keywords: copulas, Sklar's theorem, Student copula, normal copula, concordance, statistical modeling

1. Osnovni pojmi in definicije

1.1. Studentova porazdelitev. Studentova T porazdelitev je zvezna verjetnostna porazdelitev. Odkril jo je William Sealy Gosset $(1876 - 1937)$ v letu 1908. Njeno odkritje je objavil pod psevdonimom Študent (Student). Gosset je bil pivovar v pivovarni pri Guinnessu. Porazdelitev je odkril med raziskavo vpliva kvasovk na kakovost piva. Pozneje je ameriški statistik in ekonomski teoretik Harold Hotelling $(1895-1973)$ razvil ekvivalentno T porazdelitev. Tako sedaj večkrat to porazdelitev poimenujemo Studentova T porazdelitev.

Studentova T porazdelitev se pojavlja v Bayesovi analizi podatkov. T porazdelitev je simetrična in v obliki zvona, kot normalna porazdelitev, vendar ima težje repe, kar pomeni, da je bolj dovzetna za proizvodnjo vrednosti, ki so daleč od povprečja. T porazdelitev je poseben primer eliptične porazdelitve.

1.2. Osnovna ideja kopul. V teoriji verjetnosti in statistike je kopula danega slučajnega vektorja porazdelitvena funkcija primerno pridruženega slučajnega vektorja. Kopule se uporabljajo za opis odvisnosti med slučajnimi spremenljivkami. Kumulativno porazdelitveno funkcijo slučajnega vektorja lahko opišemo s pomočjo marginalnih porazdelitvenih funkcij in kopule. Marginalne porazdelitvene funkcije opisujejo marginalno porazdelitev vsake komponente slučajnega vektorja, kopula pa opisuje odvisnost med komponentami.

Osnovna ideja: Naj bo $(X_1, X_2, ..., X_n)$ slučajni vektor in naj bodo porazdelitvene funkcije $F_i(x) = P(X_i \leq x)$ zvezne. Pogosto jih imenujemo marginalne, mejne ali robne porazdelitvene funkcije. Z uporabo verjetnostne transformacije za vsako komponento, dobimo slučajni vektor $(U_1, U_2, ..., U_n) = (F_1(X_1), F_2(X_2), ..., F_n(X_n))$, ki ima enakomerno porazdeljene marginalne porazdelitvene funkcije. Kopula slučajnega vektorja $(X_1, X_2, ..., X_n)$ je definirana kot skupna kumulativna porazdelitvena funkcija vektorja $(U_1, U_2, ..., U_n)$:

$$C(u_1, u_2, ..., u_n) = P(U_1 \leq u_1, U_2 \leq u_2, ..., U_n \leq u_n).$$

Kopula C vsebuje vse podatke o odvisnostni strukturi med komponentami vektorja $(X_1, X_2, ..., X_n)$, ker mejne kumulativne porazdelitvene funkcije F_i vsebujejo vse podatke mejnih porazdelitev. Glede na postopek za generiranje vzorca $(U_1, U_2, ..., U_n)$ iz porazdelitve kopul, je lahko vzorec zgrajen kot:

$$(X_1, X_2, ..., X_n) = (F_1^{-1}(U_1), F_2^{-1}(U_2), ..., F_n^{-1}(U_n)).$$

Inverzi F_i^{-1} niso problematični, če so F_i zvezni. Zgornjo formulo kopule lahko v tem primeru zapišemo kot:

$$C(u_1, u_2, ..., u_n) = P(X_1 \leq F_1^{-1}(u_1), X_2 \leq F_2^{-1}(u_2), ..., X_n \leq F_n^{-1}(u_n)).$$

V nadaljevanju bomo podali ekvivalentno analitično definicijo kopule (ekvivalenco utemeljuje Sklarov izrek, ki ga bomo podali v nadaljevanju).

Poglejmo si še *odvisnostno strukturo*. Naj bodo $X_1, X_2, ..., X_n$ slučajni vektorji, s skupno porazdelitveno funkcijo F. Vedenje posamezne marginalne porazdelitve in pogojne verjetnosti je mogoče oceniti. Potem je odvisnostna struktura slučajnih spremenljivk vsebovana v F in iz nje razberemo, kako se "obnašajo" skupaj. Za kopulo se lahko šteje, da obravnava odvisnostno strukturo v porazdelitvi ali pa v strukturi podatkov. Pomembno je poznati primerno odvisnost. To dobimo z uporabo ustrezne kopule na slučajnih spremenljivkah.

Na primer, da imamo dve pozitivni tveganji X in Y z marginalnima porazdelitvama F_1 in F_2. Naj tveganje predstavlja potencialno izgubo podjetja. Predvidevamo, da bo podjetje sprejelo izgube, ki ne bodo večje od x in y. Vendar, če sta meji x in y preseženi, mora podjetje plačati odškodnino. To pomeni, da je vrednost pogodbe odvisna od takšnih sočasnih presežkov. Podjetje bo moralo poiskati kopulo, ki opisuje pravo odvisnostno strukturo čim bolje, ker imajo različne kopule povsem drugačno strukturo odvisnosti.

Osnovne kopule izhajajo iz porazdelitvenih funkcij in številni parametri kopul nam pomagajo izbrati ravno določeno kopulo glede na problem, ki ga rešujemo.

1.3. **Pojmi iz verjetnosti.** Spomnimo se nekaj osnovnih pojmov iz verjetnosti, ki jih bomo potrebovali v nadaljevanju. Model slučajnega eksperimenta je vselej neka množica, ki jo označimo z Ω. Družina podmnožic $\mathcal{F}$ neprazne množice Ω je σ-algebra, če velja:

- $\Omega \in \mathcal{F}$,
- $\mathcal{F}$ je zaprta za komplementiranje: $A \in \mathcal{F} \Rightarrow A^C \in \mathcal{F}$,
- $\mathcal{F}$ je zaprta za števne unije: $A_1, A_2, ... \in \mathcal{F} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

Elementi σ-algebre $\mathcal{F}$ se imenujejo $\mathcal{F}$-merljive množice. Urejenemu paru $(\Omega, \mathcal{F})$ pravimo *merljiv prostor*. Funkciji $\mu : \mathcal{F} \to [0, \infty]$ pravimo *mera* na merljivem prostoru $(\Omega, \mathcal{F})$, če velja:

- $\mu(\emptyset) = 0$.
- Za poljubno števno družino $\{A_n\}_n$ paroma disjunktnih merljivih množic velja $\mu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$.

Naj bo P *verjetnostna mera* na $\mathcal{F}$ (P je mera na $\mathcal{F}$, za katero velja $P(\Omega) = 1$). Potem trojico $(\Omega, \mathcal{F}, P)$ imenujemo *verjetnostni prostor*.

Slučajna spremenljivka je količina, ki nastopi kot rezultat poskusa (dogodka), kjer je možnih več izidov. Pri tem pa pojavitev katerekoli vrednosti iz danega območja predstavlja slučajno vrednost.

Naj bo $(\Omega, \mathcal{F}, P)$ verjetnostni prostor in $(\mathcal{M}, d)$ metrični prostor. Slučajna spremenljivka X je merljiva funkcija $X : \Omega \to \mathcal{M}$, to je

$$X^{-1}(\mathcal{B}) = \{\omega \in \Omega : X(\omega) \in \mathcal{B}\} \in \mathcal{F}$$

za vsako Borelovo množico $\mathcal{B} \subseteq \mathcal{M}$.

Spomnimo se, da so Borelove množice $\mathcal{B}$ natanko elementi Borelove σ-algebre, ki je najmanjša σ-algebra, ki vsebuje vse odprte podmnožice prostora $\mathcal{M}$. Množice $\{\omega \in \Omega : X(\omega) \in \mathcal{B}\}$ pogosto označimo kar z $\{X \in \mathcal{B}\}$. V primeru, ko je $\mathcal{M} = \mathbb{R}^n$ imenujemo pogosto X slučajni vektor. V posebnem primeru, ko je $\mathcal{M} = \mathbb{R}$, pa imenujemo X (realna) slučajna spremenljivka. V tem primeru, so Borelove množice intervali ali števne unije intervalov ali števni preseki števnih unij intervalov.

Porazdelitev P_X slučajne spremenljivke $X : \Omega \to \mathcal{M}$ je definirana z

$$P_X(\mathcal{B}) = P(X \in \mathcal{B}).$$

Spomnimo se, da je P_X verjetnostna mera glede na Borelovo $\sigma-$algebro. Naj bo $X = (X_1, X_2, ..., X_n)$ slučajni vektor, kjer so $X_i : \Omega \to \mathbb{R}, i = 1, ..., n$ (realne) slučajne spremenljivke.

Porazdelitvena funkcija F_X slučajnega vektorja $X = (X_1, X_2, ..., X_n)$, je funkcija definirana z

$$F_X(x_1, x_2, ..., x_n) = P(X_1 \leq x_1, X_2 \leq x_2, ..., X_n \leq x_n)$$

za $\forall x_i \in \mathbb{R}, i = 1, 2, ..., n$. Imenujemo jo tudi kumulativna porazdelitvena funkcija. Porazdelitvena funkcija F_X natanko določa porazdelitev P_X. Pogosto bomo namesto oznake F_{X_i} uporabljali oznako F_i.

V primeru, ko je $\mathcal{M} = \mathbb{R}$ ima porazdelitvena funkcija $F_X : \mathbb{R} \to [0, 1]$ slučajne spremenljivke naslednje lastnosti:

- $\lim_{x\to\infty} F_X(x) = 1$,
- $\lim_{x\to-\infty} F_X(x) = 0$,
- je nepadajoča funkcija,
- zveznost z leve: $\lim_{t\uparrow x} F(t) = F(x)$.

Pravimo, da je slučajni vektor $X : \Omega \to \mathbb{R}^n$ *zvezen*, če obstaja taka funkcija $f_X \geq 0$, da je

$$P_X(\mathcal{B}) = \int_{\mathcal{B}} f_X(x)dx$$

za vse Borelove množice $\mathcal{B}$ (z dx smo označili integracijo glede na Lebesguovo mero). Taki funkciji f_X pravimo gostota slučajnega vektorja X. Za gostoto f_X realne slučajne spremenljivke X je

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(t)dt.$$

Tedaj je $f_X(x) = F'_X(x)$, če odvod F_X v točki x obstaja. V točkah x, kjer odvod ne obstaja, pa je gostota enaka $f_X(x) = 0$.

Osnovni lastnosti gostote sta:

- pozitivnost: $f_X(x) \geq 0, \forall x \in \mathbb{R}$,
- normiranost: $\int_{-\infty}^{\infty} f_X(x)dx = 1$.

Za porazdelitev funkcije F_X zveznega slučajnega vektorja $X = (X_1, X_2, ..., X_n)$ velja

$$F_X(x_1, x_2, ..., x_n) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} ... \int_{-\infty}^{x_n} f_X(x_1, x_2, ..., x_n)dx_1dx_2...dx_n,$$

in tudi

$$\frac{\partial^n F_X(x_1, x_2, ..., x_n)}{\partial x_1 \partial x_2 ... \partial x_n} = f_X(x_1, x_2, ..., x_n),$$

če zapisani parcialni odvod v $(x_1, ..., x_n)$ obstaja. Gostoto robne slučajne spremenljivke dobimo iz gostote slučajnega vektorja po naslednji formuli:

$$f_{X_1}(x_1) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} f_X(x_1, x_2, ..., x_n)dx_2...dx_n$$

in analogno za ostale $f_{X_i}(x_i)$.

Naj bo X realna slučajna spremenljivka in $0 < p < 1$. Število q_p je *kvantil* slučajne spremenljivke za verjetnost p (ali tudi kar njen p-ti kvantil), če velja:

$$P(X < q_p) \leq p \quad \text{in} \quad P(X \leq q_p) \geq p.$$

Kvantilna funkcija slučajne spremenljivke X je poljubna funkcija $f : (0, 1) \to \mathbb{R}$, ki dano število p preslika v ustrezni kvantil.

Naj bodo $X_1, X_2, ..., X_n$ zvezne slučajne spremenljivke in $X = (X_1, X_2, ..., X_n)$ slučajni vektor. Pogojni gostoti sta podani z

$$f_{X_1|X_2,X_3,...,X_n}(x_1|x_2, x_3, ..., x_n) = \frac{f_X(x_1, x_2, ..., x_n)}{f_{X_2,X_3,...,X_n}(x_2, x_3, ..., x_n)},$$

$$f_{X_1,X_2,...,X_{n-1}|X_n}(x_1,x_2,...,x_{n-1}|x_n) = \frac{f_X(x_1,x_2,...,x_n)}{f_{X_n}(x_n)}.$$

Velja:

- $f_{X_1|X_2,X_3,...,X_n}(x_1|x_2,x_3,...,x_n) \geq 0$,
- $f_{X_1,X_2,...,X_{n-1}|X_n}(x_1,x_2,...,x_{n-1}|x_n) \geq 0$,
- $\int_{-\infty}^{\infty} f_{X_1|X_2,X_3,...,X_n}(x_1|x_2,x_3,...,x_n)dx_1 = 1$, če $f_{X_2,X_3,...,X_n}(x_2,x_3,...,x_n) \neq 0$,
- $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}...\int_{-\infty}^{\infty} f_{X_1,X_2,...,X_{n-1}|X_n}(x_1,x_2,...,x_{n-1}|x_n)dx_1dx_2...dx_{n-1} = 1$, če $f_{X_n}(x_n) \neq 0$.

Matematično upanje ali pričakovana vrednost $E[X]$ (realne) slučajne spremenljivke $X : \Omega \to \mathbb{R}$ je definirana kot Lebesgueov integral

$$E[X] = \int_{\Omega} X(\omega)dP(\omega).$$

Pravimo, da E[X] obstaja natanko tedaj, ko je integral

$$\int_{\Omega} |X(\omega)|dP(\omega) < \infty.$$

V nadaljevanju bomo potrebovali naslednji rezultat o vpeljavi nove spremenljivke, ki ga lahko pokažemo s pomočjo izreka o monotoni konvergenci.

Lema 1.1. *Naj bosta $(S,\mathcal{S})$ in $(T,\mathcal{T})$ merljiva prostora in $\Phi : S \to T$ merljiva funkcija. Naj bo λ mera na S in μ mera na T definirana z $\mu(A) = \lambda(\Phi^{-1}(A))$. Če je $g : (T,\mathcal{T}) \to \mathbb{R}$ merljiva, je*

$$\int_T g(x)\mu(dx) = \int_S g(\Phi(s))\lambda(ds).$$

Če je X (realna) slučajna spremenljivka, potem iz leme 1.1. sledi

$$E[X] = \int_{-\infty}^{\infty} xP_X(dx),$$

če $E(X)$ obstaja.

V zveznem primeru je matematično upanje enako

$$E[X] = \int_{-\infty}^{\infty} xf_X(x)dx,$$

kjer je $f_X(x)$ gostota slučajne spremenljivke X. Matematično upanje slučajnega vektorja $X = (X_1,...,X_n)$ je definirano z $E[X] = (E[X_1],...,E[X_n])$.

Varianca $Var(X)$ (realne) slučajne spremenljivke X je v statistiki in verjetnostni teoriji mera statistične razpršenosti določene spremenljivke. Prikazuje, kako so dejanske vrednosti razporejene okoli matematičnega upanja. Definirana je z

$$Var(X) = E[X^2] - E[X]^2 = E[(X - E[X])^2].$$

Kvadratni koren variance je imenovan *standardni odklon* $\sigma(X) = \sqrt{Var(X)}$.

Kovarianca je merilo, s katerim določamo, kako sta dve slučajni spremenljivki povezani. Kovarianca med dvema realnima slučajnima spremenljivkama X in Y, je enaka $Cov(X,Y) = E[(X - E(X))(Y - E(Y))] = E[XY] - E[X]E[Y]$.

Iz Cauchy-Schwarzove neenakosti sledi $|Cov(X,Y)| \leq \sigma(X)\sigma(Y)$. Za slučajna vektorja $X = (X_1,...,X_n)$ in $Y = (Y_1,...,Y_m)$ je njuna kovariančna matrika definirana z $Cov(X,Y) = [c_{ij}]$, kjer je $c_{ij} = Cov(X_i,Y_j)$ za $i = 1,...,n, j = 1,...,m$. Označimo $Var(X) = Cov(X,X)$.

Poglejmo si še definicijo *neodvisnosti* slučajnih spremenljivk. Če so $X_1, X_2, ..., X_n$ (realne) slučajne spremenljivke, potem pravimo, da so neodvisne, če za poljubne Borelove množice $A_1, A_2, ..., A_n \subset \mathbb{R}$ velja, da so dogodki $\{X_1 \in A_1\}$, $\{X_2 \in A_2\}$, ..., $\{X_n \in A_n\}$ neodvisni. Drugače povedano:

$$P(X_1 \in A_1, X_2 \in A_2, ..., X_n \in A_n) = P(X_1 \in A_1)P(X_2 \in A_2)...P(X_n \in A_n).$$

Za neodvisne slučajne spremenljivke $X_1, ..., X_n$ velja

$$P(X_1 \le x_1, X_2 \le x_2, ..., X_n \le x_n) = P(X_1 \le x_1)P(X_2 \le x_2)...P(X_n \le x_n)$$

oziroma v jeziku porazdelitvene funkcije

$$F_{X_1,X_2,...,X_n}(x_1, x_2, ..., x_n) = F_{X_1}(x_1)F_{X_2}(x_2)...F_{X_n}(x_n).$$

Podobno se pravilo neodvisnosti odraža na gostoti neodvisnih slučajnih spremenljivk $X_1, ..., X_n$:

$$f_{X_1,X_2,...,X_n}(x_1, x_2, ..., x_n) = f_{X_1}(x_1)f_{X_2}(x_2)...f_{X_n}(x_n)$$

in na matematičnem upanju:

$$E[X_1X_2...X_n] = E[X_1]E[X_2]...E[X_n].$$

Enakomerna zvezna porazdelitev je v statistiki in teoriji verjetnosti takšna porazdelitev, definirana na $[a, b]$, ki ima na $[a, b]$ konstantno gostoto. Porazdelitev označujemo z $X \sim U(a, b)$. Gostota enakomerne zvezne porazdelitve je enaka

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{za } a \le x \le b, \\ 0 & \text{za } x < a \text{ ali } x > b. \end{cases}$$

Njena pričakovana vrednost oziroma matematično upanje je enako $E[X] = \frac{a+b}{2}$ in varianca $Var(X) = \frac{(b-a)^2}{12}$.

Gaussova (normalna) porazdelitev je verjetnostna porazdelitev vrednosti statističnih enot v statistični populaciji, ki je v grafični predstavitvi oblikovana v obliki zvona. Zapišemo kot $X \sim N(\mu, \sigma^2)$, kjer je μ matematično upanje in σ standardni odklon. Funkcija gostote verjetnosti za normalno porazdelitev je

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

Standardna normalna porazdelitev je porazdelitev vrednosti z aritmetično sredino 0 in standardnim odklonom 1.

Vektor slučajnih spremenljivk $X = [X_1, X_2, ..., X_n]^T$ ima *večrazsežno Gaussovo (normalno) porazdelitev* z vektorjem aritmetične sredine $\mu = (\mu_1, ..., \mu_n)$ in $n \times n$ kovariančno matriko $\Sigma = [Cov(X_i, X_j)]$, $i = 1, ..., n$ in $j = 1, ..., n$. Gostota:

$$f_X(x_1, ..., x_n) = \frac{1}{\sqrt{(2\pi)^k det(\Sigma)}} e^{-\frac{1}{2}(X-\mu)^T\Sigma^{-1}(X-\mu)}.$$

Porazdelitveno funkcijo označimo z $\Phi(X) = (\Phi(X_1), \Phi(X_2), ..., \Phi(X_n))$, pri čemer je $\Phi(X_i)$ porazdelitvena funkcija normalno porazdeljene spremenljivke X_i.

Formuli, ki jih bomo potrebovali v nadaljevanju:

- $X \sim N(\mu, \sigma^2)$, $Y = aX + b$, $\Rightarrow Y \sim N(a\mu + b, (a\sigma)^2)$;
- $X \sim N(\mu_X, \sigma_X^2)$, $Y \sim N(\mu_Y, \sigma_Y^2)$, $Z = X + Y \Rightarrow Z \sim N(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2)$.

Porazdelitev hi-kvadrat je družina zveznih verjetnostnih porazdelitev vsot kvadratov k neodvisnih normalno porazdeljenih slučajnih spremenljivk. Naj bodo $X_1, X_2, ..., X_k$ neodvisne slučajne spremenljivke, ki so porazdeljene standardno normalno. Potem pravimo, da je slučajna spremenljivka $Q = \sum_{i=1}^{k} X_i^2$ porazdeljena po porazdelitvi hi kvadrat s k prostostnimi stopnjami. To zapišemo kot: $Q \sim \chi^2(k)$. Funkcija gostote verjetnosti hi-kvadrat porazdelitve za $x > 0$ je:

$$f(x;k) = \frac{1}{2\Gamma(\frac{k}{2})} e^{-\frac{x}{2}} \left(\frac{x}{2}\right)^{\frac{k}{2}-1}.$$

Pričakovana vrednost te porazdelitve je $E[Q] = k$, varianca pa $Var(Q) = 2k$.

1.4. **Pojmi iz analize.** V tem razdelku se bomo spomnili nekaj matematičnih orodij iz analize, ki jih bomo potrebovali v nadaljevanju.

Definicija 1.2. *Eulerjeva* Γ *- funkcija* je definirana za $x > 0$ s predpisom

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt.$$

Naslednjo formulo lahko izpeljemo s pomočjo vpeljane nove spremenljivke $t = bx$: za $\forall a > 0$ in $\forall b > 0$ je

$$\int_0^\infty x^a e^{-bx} dx = b^{-a-1}\Gamma(a+1). \tag{1}$$

Lastnosti Γ funkcije:

- $\Gamma(x+1) = x\Gamma(x)$,
- $\Gamma(n) = (n-1)!$ za $n \in \mathbf{N}$,
- $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

Definicija 1.3. *Eulerjeva* B *- funkcija* je definirana za $x, y > 0$ s predpisom

$$B(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1} dt.$$

Funkciji Γ in B sta med seboj povezani:

$$B(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}, \quad \textit{za } x, y > 0.$$

Definicija 1.4. *Hipergeometrijska funkcija* ${}_2F_1$ je definirana kot

$${}_2F_1 = F(a,b;c;x) = \sum_{n=0}^\infty \frac{(a)_n(b)_n}{(c)_n} \frac{x^n}{n!}.$$

Hipergeometrijska funkcija ${}_1F_1$ je definirana kot

$${}_1F_1 = F(a;b;z) = \sum_{n=0}^\infty \frac{(a)_n}{(b)_n} \frac{z^n}{n!}.$$

Pochamerjev simbol $(q)_n$ je definiran kot

$$(q)_n = \begin{cases} 1 & \text{za } n = 0, \\ q(q+1)...(q+n-1) & \text{za } n > 0. \end{cases}$$

Posebni primeri hipergeometrijske funkcije:

- $F(-p,1;1;-x) = (1+x)^p$,
- $F(1,p;p;x) = \frac{1}{1-x}$,

- $F(\frac{1}{2}, -\frac{1}{2}; \frac{1}{2}; \sin^2 x) = \cos x$,
- $\lim_{n\to\infty} F(1, n; 1; \frac{x}{n}) = e^x$,
- $F(a; b; 0) = 1$,
- $F(a; a; z) = e^z$.

Več o hipergeometrijskih funkcijah si lahko preberemo v knjigi [11].

Riemann-Stieltjesov integral je posplošitev Riemannovega integrala na intervalu $[a, b]$. Riemann-Stieltjesov integral funkcije h po funkciji g je enak:

$$\int_a^b h(x)dg(x) = \lim_{\delta\to 0} \sum_{i=1}^{n} h(\xi_i)(g(x_i) - g(x_{i-1})),$$

kjer je $\xi_i \in [x_{i-1}, x_i]$, $\delta = \max_{i=1,2,...,n}(x_i - x_{i-1})$ in velja $a = x_0 < x_1 < ... < x_{n-1} < x_n = b$. Spomnimo se, da je definicija integrala neodvisna od načina delitve. Če je h zvezna, g pa zvezno odvedljiva na $[a, b]$, se lahko Riemann-Stieltjesov integral izrazi kar z Riemannovim integralom:

$$\int_a^b h(x)dg(x) = \int_a^b h(x)g'(x)dx.$$

Splošneje, naj bo h zvezna, g pa odsekoma zvezno odvedljiva, t.j. zvezno odvedljiva na celotnem intervalu (a, b) (ki je lahko tudi neomejen), razen morda v nekaj končno točkah $c_0 = a < c_1 < ... < c_{n-1} < c_n = b$. Označimo z $\delta_0, \delta_1, ..., \delta_n$ skoke funkcije g v točkah $c_0, c_1, ..., c_n$. Tedaj je

$$\int_a^b h(x)dg(x) = \int_{c_0}^{c_1} h(x)g'(x)dx + ... + \int_{c_{n-1}}^{c_n} h(x)g'(x)dx +$$
$$+ h(c_0)\delta_0 + h(c_1)\delta_1 + ... + h(c_n)\delta_n.$$

Če je F kumulativna porazdelitvena funkcija slučajne spremenljivke X in h zvezna na $(-\infty, \infty)$, velja:

$$E[h(X)] = \int_{-\infty}^{\infty} h(x)dF(x).$$

Če X zavzame vrednosti le na intervalu $[a, b]$ in je h zvezna na $[a, b]$, pa velja tudi:

$$E[h(X)] = h(a)F(a) + \int_a^b h(x)dF(x) + h(b)(1 - F(b)).$$

Posplošitev Riemann-Stieltjesovega integrala za funkcije večih spremenljivk je podobno tehnično zahtevna kot analogna večdimenzionalna posplošitev Riemannovega integrala (glej [14]). V delu bomo potrebovali naslednji rezultat (Theorem 5, [14]).

Izrek 1.5. *Naj bo $X = (X_1, ..., X_n)$ slučajni vektor in $g : \mathbb{R}^n \to \mathbb{R}$ taka funkcija, da je $Y = g(X_1, ..., X_n)$ slučajna spremenljivka. Potem je*

$$E[Y] = \int_{\mathbb{R}^n} g(x_1, ..., x_n)dF_X(x_1, ..., x_n).$$

2. Studentova porazdelitev

2.1. **Definicija.** Studentova T porazdelitev je verjetnostna porazdelitev razmerja

$$\frac{Z}{\sqrt{\frac{V}{\nu}}},$$

kjer sta Z in V neodvisni slučajni spremenljivki, pri čemer ima

- Z normalno porazdelitev s pričakovano vrednostjo 0 ter varianco 1 in
- V porazdelitev hi-kvadrat, z ν prostostnimi stopnjami.

Gostota Studentove T porazdelitve za $\nu = n \in \mathbf{Z}$ je enaka:

$$f_n(t) = \frac{1}{\sqrt{n\pi}} \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}}. \tag{2}$$

To formulo izpeljemo na naslednji način. Naj bodo $Z_0, Z_1, ..., Z_n$ neodvisne slučajne spremenljivke porazdeljene standardno normalno in naj bo $\chi_n^2 = Z_1^2 + Z_2^2 + ... + Z_n^2$. Gostota slučajne spremenljivke χ_n^2 je

$$f(z; n) = \frac{1}{2\Gamma(\frac{n}{2})} e^{-\frac{z}{2}} \left(\frac{z}{2}\right)^{\frac{n}{2}-1},$$

s pričakovano vrednostjo n in varianco $2n$.

Zanima nas gostota $f_n(t)$ spremenljivke T.

$$T = \frac{Z_0}{\sqrt{\frac{\chi_n^2}{n}}}.$$

Najprej zapišemo:

$$f_n(t|\chi_n^2 = \nu) = \sqrt{\frac{\nu}{2\pi n}} e^{-\frac{t^2\nu}{2n}}.$$

Gostota $f_n(t)$ je torej enaka

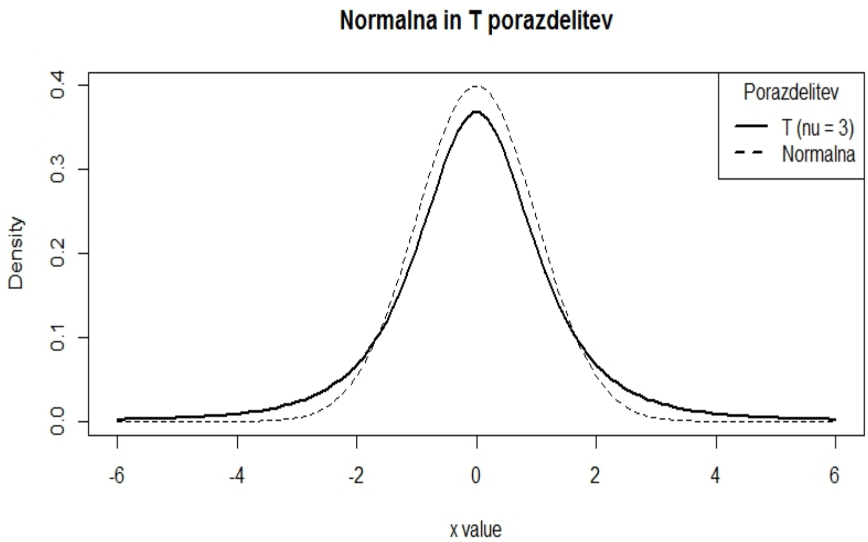
$$\begin{aligned}
f_n(t) &= \int_0^\infty f_{T,\chi_n^2}(t,\nu) d\nu = \\
&= \int_0^\infty f_n(t|\chi_n^2 = \nu) f(\nu; n) d\nu \\
&= \int_0^\infty \sqrt{\frac{\nu}{2\pi n}} e^{-\frac{t^2\nu}{2n}} \frac{1}{2\Gamma(\frac{n}{2})} e^{-\frac{\nu}{2}} \left(\frac{\nu}{2}\right)^{\frac{n}{2}-1} d\nu = \\
&= \frac{1}{2\Gamma(\frac{n}{2})} \int_0^\infty \sqrt{\frac{\nu}{2\pi n}} \left(\frac{\nu}{2}\right)^{\frac{n}{2}-1} e^{-(\frac{\nu}{2}+\frac{t^2\nu}{2n})} d\nu = \\
&= \frac{1}{2\Gamma(\frac{n}{2})} \int_0^\infty \sqrt{\frac{z}{\pi n}} z^{\frac{n}{2}-1} e^{-z(1+\frac{t^2}{n})} 2dz = \qquad (3) \\
&= \frac{1}{\Gamma(\frac{n}{2})} \frac{1}{\sqrt{\pi n}} \int_0^\infty z^{\frac{n}{2}-\frac{1}{2}} e^{-z(1+\frac{t^2}{n})} 2dz = \\
&= \frac{1}{\Gamma(\frac{n}{2})} \frac{1}{\sqrt{\pi n}} \left(1 + \frac{t^2}{n}\right)^{-(\frac{n}{2}-\frac{1}{2})-1} \Gamma\left(\frac{n}{2} - \frac{1}{2} + 1\right) = \qquad (4) \\
&= \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})} \frac{1}{\sqrt{\pi n}} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}}
\end{aligned}$$

V koraku (3) uvedemo novo spremenljivko $z = \frac{\nu}{2}$,v koraku (4) pa uporabimo lastnost (1).

Poglejmo si nekaj posebnih primerov T porazdelitvene funkcije za majhne n:

- n=1 (Cauchyjeva porazdelitev):
$$F(x) = \frac{1}{2} + \frac{1}{\pi} arctan(x),$$
- n=2:
$$F(x) = \frac{1}{2}[1 + \frac{x}{\sqrt{2+x^2}}],$$
- n=3:
$$F(x) = \frac{1}{2} + \frac{1}{\pi} arctan\left(\frac{x}{\sqrt{3}}\right) + \frac{x\sqrt{3}}{\pi(3+x^2)},$$
- n=4:
$$F(x) = \frac{1}{2} + \frac{x(x^2+6)}{2(4+x^2)^{3/2}},$$
- n=5:
$$F(x) = \frac{1}{2} + \frac{1}{\pi} arctan\left(\frac{x}{\sqrt{5}}\right) + \frac{x\sqrt{5}(3x^2+25)}{3\pi(5+x^2)^2},$$
- n=6:
$$F(x) = \frac{1}{2} + \frac{x(2x^4+30x^2+135)}{4(6+x^2)^{5/2}}.$$

Ko pošljemo n $\to \infty$ v (2), dobimo gostoto standardne normalne porazdelitve
$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$



Slika 1. Gostota Normalne in Studentove T porazdelitve.

Na sliki 1 je s črtkano črto narisana gostota normalne porazdelitve, z neprekinjeno črto pa gostota T porazdelitve s 3 prostostnimi stopnjami. S povečevanjem prostostnih stopenj se oblika vedno bolj prilega obliki normalne porazdelitve.

Povemo lahko še, da se porazdelitvena funkcija z eno prostostno stopnjo, imenovana Cauchyjeva porazdelitev, pogosto uporablja pri poučevanju, predvsem zaradi njenih težkih repov. Porazdelitvena funkcija s 4 prostostnimi stopnjami, pa je dober model za izračun indeksov donosnosti v globalnem okolju.

3. Kopule

3.1. **Analitična definicija kopule.** Kopula je večrazsežna porazdelitvena funkcija, definirana na $[0,1]^n$, ki ima enakomerno porazdeljene marginalne porazdelitve. V naslednji definiciji podajamo analitično definicijo kopule.

Definicija 3.1. *Večrazsežna kopula* je taka funkcija $C : [0,1]^n \to [0,1]$, da za $\forall u, u_1, ..., u_i \in I = [0,1]$ velja:

(i) $C(u_1, ..., u_{i-1}, 0, u_{i+1}, ..., u_n) = 0$,

(ii) $C(1, ..., 1, u, 1, ..., 1) = u$,

(iii) C je n-naraščajoča, kar pomeni, da je za vsak hiperpravokotnik

$$B = \prod_{i=1}^{n} [x_i, y_i] \subseteq [0,1]^n,$$

C-volumen B-ja nenegativen:

$$\int_B dC(u) = \sum_{z \in \times_{i=1}^{n} \{x_i, y_i\}} (-1)^{N(z)} C(z) \geq 0,$$

kjer je $N(z) = \#\{k : z_k = x_k\}$.

V primeru $n = 2$ se pogoj (iii) glasi: $C(y_1, y_2) - C(x_1, y_2) - C(y_1, x_2) + C(x_1, x_2) \geq 0$ za vse $[x_1, y_1] \times [x_2, y_2] \subset [0,1] \times [0,1]$. Sledi:

- C nepadajoča po vsaki spremenljivki,
- C zadostuje Lipshitzovemu pogoju za vsak $u_1, u_2, v_1, v_2 \in I$:

$$|C(u_2, v_2) - C(u_1, v_1)| \leq |u_2 - u_1| + |v_2 - v_1|.$$

Od tod sledi, da so kopule enakomerno zvezne funkcije.

3.2. **Sklarov izrek.** Kopule so izjemno pomembne v statističnem modeliranju, kar Sklarov izrek delno pojasnjuje. Pred kratkim je pridobil popularnost pri kopulah, pojavil pa se je že leta 1959, ko ga je v svojih matematičnih raziskavah omenil Abe Sklar.

Izrek 3.2. *Naj bodo $X_1, X_2, ..., X_n$ slučajne spremenljivke s porazdelitvenimi funkcijami $F_1(x_1), ..., F_n(x_n)$ in skupno porazdelitveno funkcijo $F(x_1, ..., x_n)$. Potem obstaja taka n-razsežna kopula $C(u_1, ..., u_n)$, da je*

(5)
$$F(x_1, ..., x_n) = C(F_1(x_1), ..., F_n(x_n)) = P(F_1(X_1) \leq F_1(x_1), ..., F_n(X_n) \leq F_n(x_n))$$

za vsak $x_i \in \mathbb{R}, i = 1, 2, ..., n$. Če so vse porazdelitvene funkcije $F_i(x_i)$ zvezne, potem je taka kopula ena sama, izračunamo jo z enačbo

$$C(u_1, ..., u_n) = F(F_1^{-1}(u_1), ..., F_n^{-1}(u_n)) = P(X_1 \leq F_1^{-1}(u_1), ..., X_n \leq F_n^{-1}(u_n)).$$

Velja tudi obratno. Če je $C(u_1, ..., u_n)$ kopula in so $F_1(x_1), ..., F_n(x_n)$ porazdelitvene funkcije, potem je tudi funkcija $F(x_1, ..., x_n)$, ki je definirana z enačbo (5), porazdelitvena funkcija z robnimi porazdelitvenimi funkcijami $F_1(x_1), ..., F_n(x_n)$.

Dokaz Sklarovega izreka je precej dolg in razmeroma zahteven. Bralec ga lahko najde v [8].

Sklarov izrek je zelo pomemben, ker lahko z njegovo pomočjo analiziramo odvisnostno strukturo večrazsežnih porazdelitev brez, da bi preučili marginalne porazdelitve. Oglejmo si primer. Imamo Gumbelovo dvorazsežno porazdelitveno funkcijo $F(x_1, x_2) = (1 + e^{-x_1} + e^{-x_2})^{-1}$, $F(x_1, x_2) \in [0, 1]$. Lahko pokažemo, da sta marginalni porazdelitvi enaki $F_1(x_1) = \int_{\mathbb{R}} F(x_1, x_2)dx_2 = (1 + e^{-x_1})^{-1}$ in $F_2(x_2) = (1 + e^{-x2})^{-1}$. Za izračun kopule si pomagajmo z razmerjem

$$\frac{1 - F(x_1, x_2)}{F(x_1, x_2)} = e^{-x_1} + e^{-x_2} = \frac{1 - F(x_1)}{F(x_1)} + \frac{1 - F(x_2)}{F(x_2)}.$$

Označimo $u_1 = F(x_1)$ in $u_2 = F(x_2)$. Sledi:

$$\frac{1 - C(u_1, u_2)}{C(u_1, u_2)} = \frac{1 - u_1}{u_1} + \frac{1 - u_2}{u_2}.$$

Izrazimo in dobimo

$$C(u_1, u_2) = \frac{u_1 u_2}{u_1 + u_2 - u_1 u_2}.$$

Vendar pa ni vedno očitno identificirati kopulo. Za veliko finančnih aplikacij ni problem v tem, da bi uporabili dano večrazsežno porazdelitev, ampak, da iskana priročna porazdelitev, ki opiše nekatera splošna dejstva, na primer odnos med različnimi donosi sredstev, obstaja. V večini aplikacij se je predpostavljalo, da je porazdelitev Gaussova. Kopule so v financah dobro orodje, ker lahko problem modeliranja razdelimo na dva koraka:

- na prvem koraku se ukvarja z identifikacijo marginalnih porazdelitev;
- drugi korak pa sestoji pri določanju primerne kopule, da dobro zastopa odvisnostno strukturo.

3.3. **Lastnosti kopul.** Naslednje lastnosti so opisane v [7] v poglavjih 1.3.3, 1.3.4, 1.4.2 in 1.4.3.

3.3.1. *Diferenciabilnost.* Naj bo C kopula. Za vsak $i = 1, \ldots, n$, obstaja parcialni odvod $\frac{\partial C}{\partial u_i}$ in velja:

$$0 \leq \frac{\partial}{\partial u_i} C(u_1, ..., u_i, ..., u_n) \leq 1$$

Funkcija $\frac{\partial}{\partial u_i} C(u_1, ..., u_i, ..., u_n)$ je definirana in nepadajoča povsod na območju $[0, 1]$.

3.3.2. *Gostota kopula.* Naj bo $X = (X_1, \ldots, X_n)$ slučajni vektor s pripadajočo kopulo C. Če ima X gostoto f_X, potem je gostota kopule C enaka

$$c(u_1, \ldots, u_n) = \frac{f_X(F_1^{-1}(u_1), \ldots, F_n^{-1}(u_n))}{\Pi_{i=1}^n f_i(F_i^{-1}(u_i))}.$$

$$c(u_1, \ldots, u_n) = \frac{\partial^n C(u_1, \ldots, u_n)}{\partial u_1 \ldots \partial u_n}$$

Velja, če zgornji parcialni odvodi obstajajo.

Če je $g : \mathbb{R}_n \to \mathbb{R}$ taka funkcija, da je $Y = g(X_1, \ldots, X_n)$ slučajna spremenljivka, potem je po izreku 1.5, lemi 1.1 in Sklarovem izreku

$$E[g(X_1, \ldots, X_n)] = \int_{[0,1]^n} g(F_1^{-1}(u_1), \ldots, F_n^{-1}(u_n))dC(u_1, \ldots, u_n).$$

Če ima kopula C gostoto c, potem je

$$E[g(X_1, \ldots, X_n)] = \int_{[0,1]^n} g(F_1^{-1}(u_1), \ldots, F_n^{-1}(u_n))c(u_1, \ldots, u_n)du_1 \ldots du_n.$$

3.3.3. *Invariantnost.* Naslednji izrek pove, da so kopule invariantne pod strogo monotonimi transformacijami slučajnih spremenljivk.

Izrek 3.3. *Naj bodo $X_1, X_2, ..., X_n$ zvezno porazdeljene slučajne spremenljivke s kopulo C in naj bodo $T_1, T_2, ..., T_n$ strogo naraščajoče funkcije. Potem velja, da imajo $T_1(X_1), T_2(X_2), ..., T_n(X_n)$ enako kopulo C.*

Dokaz. Omejili se bomo na primer $n = 2$. Dokaz za splošen n je analogen. Naj bosta F_1 in F_2 porazdelitveni funkciji slučajnih spremenljivk X_1 in X_2 in naj bosta T_1^{-1} in T_2^{-1} inverza transformacijskih funkcij T_1 in T_2. Dalje, naj bosta G_1, G_2 porazdelitveni funkciji za $T_1(X_1), T_2(X_2)$ in naj bo C_T njuna kopula. G_1 in G_2 sta zvezni. Tako imamo za $i = 1, 2$ in poljubna X_1, X_2

$$G_i(x_i) = P(T_i(X_i) \leq x_i) = P(X_i \leq T_i^{-1}(x_i)) = F_i(T_i^{-1}(x_i))$$

$$\begin{aligned} C_T(G_1(x_1), G_2(x_2)) &= P[T_1(X_1) \leq x_1, T_2(X_2) \leq x_2], \\ &= P[X_1 \leq T_1^{-1}(x_1), X_2 \leq T_2^{-1}(x_2)], \\ &= C[F_1(T_1^{-1}(x_1)), F_2(T_2^{-1}(x_2))], \\ &= C[G_1(x_1), G_2(x_2)]. \end{aligned}$$

Torej je $C_T(G_1(x_1), G_2(x_2)) = C(G_1(x_1), G_2(x_2))$ in zato je $C_T = C$ na I^2. □

3.3.4. *Produktna kopula.* Struktura neodvisnosti slučajnih spremenljivk in njihovih porazdelitvenih funkcij je pomembna za aplikacije. Če sta X in Y neodvisni slučajni spremenljivki, potem je produkt njunih marginalnih porazdelitev F in G enak njuni skupni porazdelitveni funkciji H:

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

za vse $x, y \in \mathbb{R}$. Dalje, kopulo neodvisnih slučajnih spremenljivk označimo kot $C = \Pi$, torej je $\Pi(u, v) = uv$. Za n neodvisnih spremenljvk: $C^{ind}(u_1, ..., u_n) = \Pi(u_1, ..., u_n) = \Pi_{i=1}^n u_i$.

3.3.5. *Preživetvena kopula.* Verjetnost posameznika, da živi dlje od časa x, je podana s preživetveno funkcijo $\bar{F}(x) = P(X > x) = 1 - F(x)$, kjer je F porazdelitvena funkcija x-a. Če sta X in Y zvezni slučajni spremenljivki s skupno porazdelitveno funkcijo H in kopulo C, je skupna preživetvena funkcija dana kot: $\bar{H}(x, y) = P(X > x, Y > y)$. Povezava med preživetveno funkcijo posameznika in skupno preživetveno funkcijo je:

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

Sedaj definiramo *preživetveno kopulo* $\tilde{C} : I^2 \to I$ z $\tilde{C}(u, v) = u + v - 1 + C(1 - u, 1 - v)$. Tako je $\bar{H}(x, y) = \tilde{C}(\bar{F}(x), \bar{G}(y))$. Za $U = F(X)$ in $V = F(Y)$ je

$$\bar{C}(u, v) = P(U > u, V > v) = 1 - u - v + C(u, v) = \tilde{C}(1 - u, 1 - v),$$

kjer je $\bar{C}$ skupna preživetvena funkcija in $\tilde{C}$ preživetvena kopula enakomerno porazdeljenih slučajnih spremenljivk U in V na $(0, 1)$, ki imata po izreku 3.3 pripadajočo kopulo C.

3.4. **Fréchet-Hoeffdingove meje.** Naj bo $C(u, v)$ kopula. Potem je

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

Če je H bivariatna porazdelitvena funkcija, z zveznima marginalnima porazdelitvama F in G, potem je po Sklarovem izreku $H(x, y) = P(X \leq x, Y \leq y) = C(F(x), G(y))$, kjer je $C(u, v) = H(F^{-1}(u), G^{-1}(v))$ in $u = F(x), v = G(y)$.

Imamo

$$max\{F(x) + G(y) - 1, 0\} \leq H(x, y) \leq min\{F(x), G(y)\}.$$

M in W sta zgornja in spodnja Fréchet-Hoeffdingova meja. Sta ravno tako kopuli. Funkcije M^n, W^n in Π^n so dane kot:

$$\begin{aligned} M^n(u) &= min(u_1, u_2, ..., u_n), \\ W^n(u) &= max(u_1 + u_2 + ... + u_n - n + 1, 0), \\ \Pi^n(u) &= u_1 u_2 ... u_n. \end{aligned}$$

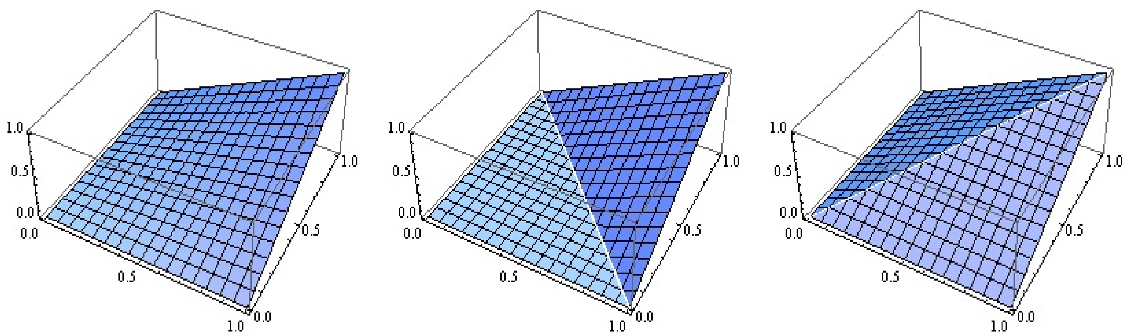
Funkciji M^n in Π^n sta n-kopuli za vse $n \geq 2$, medtem ko funkcija W^n ni n-kopula za $n > 2$. Kopule zgornje in spodnje Fréchetove meje (M in W) imata pomembne interpretacije v statistiki.

Trditev 3.4. *Naj bosta X in Y zvezna slučajna vektorja. Potem:*

- *Kopula spremenljivk X in Y je enaka $M(u, v)$ natanko tedaj, ko je vsaka izmed X in Y strogo naraščajoča funkcija druge.*
- *Kopula spremenljivk X in Y je enaka $W(u, v)$ natanko tedaj, ko je vsaka izmed X in Y strogo padajoča funkcija druge.*

Dokaz zgornjih lastnosti lahko bralec najde v [7] v poglavju 1.5.

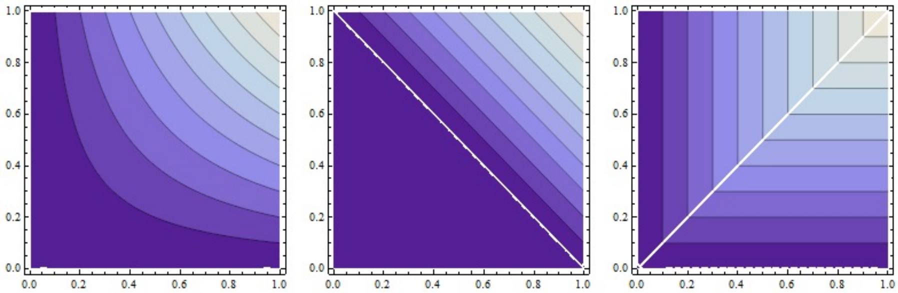
Na sliki 2 so narisane kopule Π, W in M (v tem vrstnem redu), na sliki 3 pa so njeni obrisi.



Slika 2. Produktna Π, spodnja W in zgornja M kopula.

3.5. **Mere odvisnosti.**

3.5.1. *Korelacija.* Korelacija meri moč povezave dveh spremenljivk. Če obstaja močna povezava, je tam velika korelacija. Korelacijo je moč meriti z več različnimi koeficienti, prilagojenimi za različne tipe podatkov, ki so na voljo. Izmed korelacijskih koeficientov je najbolj znan Pearsonov korelacijski koeficient, ki je računan na podlagi kovariance in standardnih odklonov obeh spremenljivk. Ogledali pa si bomo



SLIKA 3. Obrisi produktne, spodnje in zgornje kopule.

še Spearmanov korelacijski koeficient ranga, ki je posebna oblika Pearsonovega koeficienta, v kateri so podatki pred izračunom koeficientov preoblikovani v range, ter Kendallou tau.

3.5.2. *Pearsonov korelacijski koeficient.* Imamo slučajni spremenljivki X in Y, ki v finančnih aplikacijah lahko opisujeta določene dejavnike tveganj. Pearsonov korelacijski koeficient je podan kot:

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}},$$

kjer je $Var(X)$ in $Var(Y)$ varianca spremenljivk X in Y, ter $Cov(X,Y)$ njuna kovarianca. Vrednost dobljenega koeficienta je v mejah med -1 do +1. Pozitiven koeficient izraža pozitivno povezanost, kar pomeni, da so vrednosti pozitivno linearno povezane (pri rasti enega dejavnika tveganja raste tudi vrednost drugega dejavnika tveganja). Negativen koeficient pa označuje negativno korelacijo med vrednostmi dejavnikov tveganja.

3.5.3. *Mera konkordance.* Kot drugo mero odvisnosti omenimo *mero konkordance.* Z njo bi radi merili, kako se velike oziroma majhne vrednosti ene slučajne spremenljivke pojavljajo hkrati z velikimi oziroma majhnimi vrednostmi druge slučajne spremenljivke.

Definicija 3.5. Naj bosta (x_i, y_i) in (x_j, y_j) opazovanji slučajnega zveznega vektorja (X,Y). Pravimo, da opazovanji *imata mero konkordance*, če $x_i < x_j$ in $y_i < y_j$, ali $x_i > x_j$ in $y_i > y_j$, torej: $(x_i - x_j)(y_i - y_j) > 0$. Podobno pravimo, da opazovanji *nimata mere konkordance*, če $x_i < x_j$ in $y_i > y_j$, ali $x_i > x_j$ in $y_i < y_j$, torej: $(x_i - x_j)(y_i - y_j) < 0$.

Geometrično: dve točki v ravnini (x_i, y_i) in (x_j, y_j) imata mero konkordance natanko tedaj, ko ima daljica, ki ju povezuje, pozitiven naklon. Ta daljica ima negativen naklon natanko tedaj, ko točki nimata mere konkordance.

Trditev 3.6. *Naj bosta (X_1, Y_1), (X_2, Y_2) zvezna slučajna vektorja s skupnima porazdelitvama H_1 in H_2. Naj bosta X_1 in X_2 enako porazdeljena (s porazdelitvijo F), ter Y_1 in Y_2 enako porazdeljena (s porazdelitvijo G). Naj bo C_1 kopula vektorja (X_1, Y_1) in C_2 kopula vektorja (X_2, Y_2), torej je $H_1(x,y) = C_1(F(x), G(y))$ in $H_2(x,y) = C_2(F(x), G(y))$. Označimo s κ razliko med verjetnostima, ko slučajna vektorja (X_1, Y_1) in (X_2, Y_2) imata in nimata mero konkordance:*

$$\kappa = P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0]$$

Potem velja:

$$\kappa = \kappa(C_1, C_2) = 4 \int\int_{I^2} C_2(u,v) dC_1(u,v) - 1. \tag{6}$$

Dokaz. Velja

$$\begin{aligned}
P[(X_1 - X_2)(Y_1 - Y_2) < 0] &= 1 - P[(X_1 - X_2)(Y_1 - Y_2) > 0] \\
\kappa &= 2P[(X_1 - X_2)(Y_1 - Y_2) > 0] - 1 \\
P[(X_1 - X_2)(Y_1 - Y_2) > 0] &= P[X_1 > X_2, Y_1 > Y_2] + P[X_1 < X_2, Y_1 < Y_2].
\end{aligned}$$

Te verjetnosti lahko ocenimo z integriranjem porazdelitve enega izmed vektorjev (X_1, Y_1) ali (X_2, Y_2)

$$\begin{aligned}
P[X_1 > X_2, Y_1 > Y_2] &= P[X_2 < X_1, Y_2 < Y_1] \\
&= \int\int_{I^2} P[X_2 \leq x, Y_2 \leq y] dC_1(F(x), G(y)) \\
&= \int\int_{I^2} C_2(F(x), G(y)) dC_1(F(x), G(y)) \\
&= \int\int_{I^2} C_2(u,v) dC_1(u,v).
\end{aligned}$$

Zadnji izraz je dosežen z verjetnostno transformacijo $u = F(x)$ in $v = G(y)$.

Ker je C_1 skupna porazdelitvena funkcija para $(U, V) = (F(X_1), G(Y_1))$, kjer sta spremenljivki enakomerno porazdeljeni na $(0,1)$, je $E(U) = E(V) = \frac{1}{2}$ in velja

$$\begin{aligned}
P[X_1 < X_2, Y_1 < Y_2] &= \int\int_{I^2} P[X_2 > x, Y_2 > y] dC_1(F(x), G(y)) \\
&= \int\int_{I^2} [1 - F(x) - G(y) + C_2(F(x), G(y))] dC_1(F(x), G(y)) \\
&= \int\int_{I^2} [1 - u - v + C_2(u,v)] dC_1(u,v).
\end{aligned}$$

$$\begin{aligned}
P[X_1 < X_2, Y_1 < Y_2] &= 1 - \frac{1}{2} - \frac{1}{2} + \int\int_{I^2} C_2(u,v) dC_1(u,v) \\
&= \int\int_{I^2} C_2(u,v) dC_1(u,v)
\end{aligned}$$

in zato je

$$P[(X_1 - X_2)(Y_1 - Y_2) > 0] = 2 \int\int_{I^2} C_2(u,v) dC_1(u,v)$$

$$\kappa = 2P[(X_1 - X_2)(Y_1 - Y_2) > 0] - 1 = 4 \int\int_{I^2} C_2(u,v) dC_1(u,v) - 1,$$

s čimer je trditev dokazana. □

Definicija 3.7. Odvisnostni ukrep $\kappa = \kappa_{X,Y} = \kappa_C$, med dvema zveznima slučajnima spremenljivkama X in Y s kopulo C, je mera konkordance, če izpolnjuje naslednje lastnosti.

(i) κ je definirana za vsak par X, Y slučajnih zveznih spremenljivk.

(ii) $-1 \leq \kappa_{X,Y} \leq 1, \kappa_{X,X} = 1$ in $\kappa_{X,-X} = -1$.

(iii) $\kappa_{X,Y} = \kappa_{Y,X}$.

(iv) Če sta X in Y neodvisni, potem velja: $\kappa_{X,Y} = \kappa_\Pi = 0$.
(v) $\kappa_{-X,Y} = \kappa_{X,-Y} = -\kappa_{X,Y}$.
(vi) Če sta C_1 in C_2 kopuli za kateri velja: $C_1 \leq C_2$ (po točkah), potem velja: $\kappa_{C_1} \leq \kappa_{C_2}$.
(vii) Naj bo $\{(X_n, Y_n)\}$ zaporedje parov zveznih slučajnih spremenljivk, povezanih s kopulami C_n. Naj zaporedje $\{C_n\}$ konvergira po točkah proti C. Potem velja: $\lim_{n\to\infty} \kappa_{C_n} = \kappa_C$.

Trditev 3.8. *Naj bo κ mera konkordance za zvezni slučajni spremenljivki X in Y.*
(i) *Če je Y strogo naraščajoča funkcija X-a, potem: $\kappa_{X,Y} = \kappa_M = 1$.*
(ii) *Če je Y strogo padajoča funkcija X-a, potem: $\kappa_{X,Y} = \kappa_W = -1$.*

Dokaz. (i) Po trditvi 3.4 in trditvi 3.6 je

$$\begin{aligned}\kappa_{X,Y} = \kappa(M,M) = \kappa_M &= 4\int\int_{I^2} min(u,v) dM(u,v) - 1 \\ &= 4\int_0^1 u du - 1 = 1.\end{aligned}$$

(ii) Podobno je

$$\begin{aligned}\kappa_{X,Y} = \kappa(W,W) = \kappa_W &= 4\int\int_{I^2} max(u+v-1,0) dW(u,v) - 1 \\ &= 4\int_0^1 0 du - 1 = -1.\end{aligned}$$

□

V nadaljevanju bomo definirali populacijski verziji Kendallovega taua in Spearmanovega korelacijskega koeficienta.

3.5.4. *Spearmanov korelacijski koeficient ranga.* Definiran je glede na mero konkordance. Naj bodo (X_1, Y_1), (X_2, Y_2) in (X_3, Y_3) neodvisni zvezni slučajni vektorji s skupno porazdelitveno funkcijo H. Naj bo F marginalna porazdelitev slučajnih spremeljivk X_1, X_2, X_3, G marginalna porazdelitev slučajnih spremenljivk Y_1, Y_2, Y_3 in naj bo C kopula vektorja (X_1, Y_1). Trikratnik razlike med verjetnostjo, ko imamo in verjetnostjo, ko nimamo mere konkordance vektorjev $(X_1, Y_1), (X_2, Y_3)$ označimo z:

$$\rho = \rho_C = 3(P[(X_1 - X_2)(Y_1 - Y_3) > 0] - P[(X_1 - X_2)(Y_1 - Y_3) < 0])$$

in ga imenujemo Spearmanov korelacijski koeficient.

Par (X_2, Y_3) je neodvisen in zato kopulo spremeljivk X_2 in Y_3 označimo s Π; skupna porazdelitvena funkcija vektorja (X_2, Y_3) je $F(x)G(y)$.
Tako je $\rho_C = \rho_{X,Y} = 3\kappa(C, \Pi)$. Iz enačbe (6) sledi

$$\begin{aligned}\rho_C &= 12\int\int_{I^2} uv dC(u,v) - 3 \\ &= 12E(UV) - 3 \\ &= \frac{E(UV) - E(U)E(V)}{\sqrt{Var(U)}\sqrt{Var(V)}} = \frac{cov(U,V)}{\sqrt{Var(U)}\sqrt{Var(V)}},\end{aligned}$$

saj imata enako porazdeljeni slučajni spremenljivki $U = F(X_1)$ in $V = G(Y_1)$ povprečje $\frac{1}{2}$ in varianco $\frac{1}{12}$. Torej je ρ_C enak Pearsonovemu korelacijskemu koeficientu spremenljivk U in V.

3.5.5. *Kendallov tau.* Kendallou tau je ravno tako kot Spearmanov korelacijski koeficient ranga definiran glede na mero konkordance.

Naj bo $\{(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)\}$ slučajni vzorec n opazovanj vektorja (X, Y), kjer sta X in Y zvezni slučajni spremenljivki. Vseh različnih parov (x_i, y_i), (x_j, y_j) opazovanj je $\binom{n}{2}$, kjer vsak par ali je ali pa ni konkordanten. Naj c označuje število parov, ki imajo mero konkordance in d število parov, ki nimajo mere konkordance. Potem je *Kendallov tau* za dan slučajen vzorec definiran kot:

$$t = \frac{c-d}{c+d} = \frac{c-d}{\binom{n}{2}}.$$

Sedaj definirajmo Kendallov tau populacije za vektor (X, Y) slučajnih zveznih spremenljivk s skupno porazdelitveno funkcijo H in kopulo C. Naj bosta (X_1, Y_1) in (X_2, Y_2) neodvisna in enako porazdeljena slučajna vektorja (s porazdelitvijo H). Kendallov tau populacije je enak

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

Torej je po trditvi 3.5

$$\tau_{X,Y} = \tau_C = \kappa(C, C) = 4 \int\int_{I^2} C(u, v) dC(u, v) - 1, \tag{7}$$

kjer je C kopula. Integral v (7) je pričakovana vrednost slučajne spremenljivke $C(U, V)$, za enakomerno porazdeljeni slučajni spremenljivki $U = F(X_1)$ in $V = G(Y_1)$ na $(0, 1)$ in skupno porazdelitvijo funkcij, ki pripada C. Torej je:

$$\tau_C = \tau(X, Y) = 4E(C(U, V)) - 1.$$

Kendallov tau in Spearmanov korelacijski koeficient sta obe simetrični odvisnostni meri, ki zavzamata vrednost na intervalu $[-1, 1]$. Ko imajo tveganja zvezne marginalne porazdelitve, sta obe korelaciji odvisni le od kopule tveganja.

Več o merah odvisnosti lahko najdemo v [7] v 2. poglavju.

4. Večrazsežna Studentova T porazdelitev

4.1. **Studentova T kopula.** Naj bo P simetrična, pozitivno definitna matrika, ki ima na diagonalah enice. Dalje, naj bo $T_{P,\nu}$ večrazsežna T porazdelitev z ν prostostnimi stopnjami in korelacijsko matriko P.

$$T_{P,\nu}(x_1, ..., x_n) = \int_{-\infty}^{x_1} ... \int_{-\infty}^{x_n} \frac{\Gamma(\frac{\nu+n}{2})|P|^{-\frac{1}{2}}}{\Gamma(\frac{\nu}{2})(\nu\pi)^{\frac{n}{2}}} \left(1 + \frac{x^T P^{-1} x}{\nu}\right)^{-\frac{\nu+n}{2}} dx.$$

Pripadajoča kopula je enaka

$$C(u_1, u_2, ..., u_n; P, \nu) = C^t_{P,\nu}(u) = T_{P,\nu}(t_\nu^{-1}(u_1), t_\nu^{-1}(u_2), ..., t_\nu^{-1}(u_n)),$$

kjer je t_ν^{-1} kvantilna funkcija T porazdelitve v eni spremenljivki.

Gostota T kopule je enaka

$$c(u_1, u_2, ..., u_n; P, \nu) = c^t_{P,\nu}(u) = \frac{f_{P,\nu}(t_\nu^{-1}(u_1), ..., t_\nu^{-1}(u_n))}{\Pi_{i=1}^n f_\nu(t_\nu^{-1}(u_i))} =$$

$$= |P|^{-1} \frac{\Gamma(\frac{\nu+n}{2})[\Gamma(\frac{\nu}{2})]^n}{[\Gamma(\frac{\nu+1}{2})]^n \Gamma(\frac{\nu}{2})} \frac{(1 + \frac{1}{\nu}\zeta^T P^{-1}\zeta)^{-\frac{\nu+n}{2}}}{\Pi_{k=1}^n \left(1 + \frac{\zeta_k^2}{\nu}\right)^{-\frac{\nu+1}{2}}},$$

kjer je $\zeta = t_\nu^{-1}(u_n)$.

Velikokrat T kopulo primerjamo s kopulo večrazsežne Gaussove porazdelitve, za katero pogosto rečemo, da je limitni primer T kopule, ko pošljemo $\nu \to \infty$. Poglejmo si še večrazsežno Gaussovo kopulo. Naj bo P enako definirana kot pri T kopuli. Naj bo Φ_P standardna večrazsežna normana porazdelitev s korelacijsko matriko P. Večrazsežna Gaussova kopula je enaka

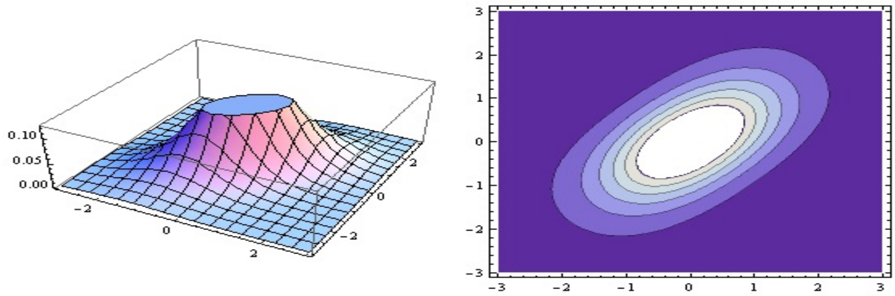
$$C(u_1, u_2, ..., u_n; P) = C_P^{Ga}(u) = \Phi_P(\Phi^{-1}(u_1), \Phi^{-1}(u_2), ..., \Phi^{-1}(u_n)).$$

Gostota kopule pa je enaka:

$$c(u_1, ..., u_n; P) = c_P^{Ga}(u) = |P|^{-\frac{1}{2}} e^{-\frac{1}{2}\zeta^T (P^{-1} - I)\zeta},$$

kjer je $\zeta_n = \Phi^{-1}(u_n)$, I pa identična matrika.

Poglejmo si primer T kopule. Imamo 10 prostostnih stopenj ($\nu = 10$) in korelacijsko matriko $\mathbf{P} = \begin{bmatrix} 1 & 0,6 \\ 0,6 & 1 \end{bmatrix}$.



Slika 4. Gostota T kopule in njen obris.

Na sliki 4 je narisana gostota T kopule in njen obris pri danemu ν in matriki P.

Gaussova in T kopula sta posebna primera splošnejših eliptičnih kopul. Zanju veljata naslednji odvisnosti (glej npr. [13]):

- med Spearmanovim in Pearsonovim korelacijskim koeficientom:

$$\rho_C = \frac{6}{\pi} \sin^{-1}\left(\frac{\rho}{2}\right),$$

- med Kendallovim tau in Pearsonovim korelacijskim koeficientom:

$$\tau = \frac{2}{\pi} \sin^{-1}(\rho).$$

Pogledati si moramo še odvisnost zgornjih repov.

Definicija 4.1. Če je dvorazsežna kopula C taka, da obstaja limita

$$\lim_{u \to 1} \frac{\bar{C}(u, u)}{1 - u} = \lambda,$$

potem pravimo, da ima C odvisnost zgornjih repov, če je $\lambda \in (0, 1]$ in nima odvisnosti zgornjih repov, če je $\lambda = 0$.

Odvisnost λ veliko uporabljajo v teoriji ekstremnih vrednosti. Predstavlja verjetnost, da je ena spremenljivka ekstremna, pri pogoju, da je tudi druga spremenljivka ekstremna. Naj bo $\lambda(u) = P(U_1 > u | U_2 > u) = \frac{\bar{C}(u,u)}{1-u}$. $\lambda(u)$ lahko predstavlja kvantilno-odvisno mero odvisnosti.

Ekstremi Gaussove kopule so asimptotično neodvisni za $\rho \neq 1$, to je $\lambda = 0$ za $\rho < 1$. Studentova kopula pa je zanimiv kontrast Gaussove kopule, saj so njeni ekstremi asimptotično odvisni za $\rho \neq 1$. Ko pa pošljemo stopnje odvisnosti proti neskončnosti, Studentova kopula ustreza Gaussovi.

V matematičnih financah in drugih aplikacijah statistike pogosto ni jasno, katera večrazsežna porazdelitev bi najbolje simulirala dane podatke. V zadnjem času je pri tem postala popularna uporaba kopul, saj lahko z njihovo pomočjo (Sklarov izrek) s komponiranjem primernih marginalnih porazdelitev obvladamo zgornjo težavo. V nadaljevanju si bomo ogledali konstrukcijo kandidatov za večrazsežno Studentovo T porazdelitev (in posledično kopulo) v primeru, ko marginalne porazdelitve nimajo nujno enakih prostostnih stopenj (pri tem sledimo članku [16]). Omejili se bomo na dvorazsežen primer. Najprej pa si podrobneje oglejmo primer, ko imajo marginalne porazdelitve enake prostostne stopnje.

4.2. **Studentova T porazdelitev z enakimi marginalnimi porazdelitvami (z enakimi prostostnimi stopnjami).** Studentovo T kopulo z enakimi marginalnimi porazdelitvami razumemo kot simulacijo linearne kombinacije Gaussovih spremenljivk, ki so definirane s strukturo korelacije razcepa Choleskega, in njihovo delitvijo glede na isto χ^2 spremenljivko. Naj bodo W_i neodvisne standardne Gaussove spremenljivke in naj bosta:

$$Z_{01} = \alpha W_1 + \beta W_2, \quad Z_{02} = \gamma W_1 + \delta W_2,$$

transformirani spremenljivki z varianco 1:

$$\alpha^2 + \beta^2 = 1 = \gamma^2 + \delta^2.$$

Nato konstruiramo

$$T_1 = Z_{01}\sqrt{\frac{n}{C^2}}, \quad T_2 = Z_{02}\sqrt{\frac{n}{C^2}},$$

kjer je C spremenljivka s porazdelitvijo χ^2 in parametrom n. Ker imata T_1 in T_2 isti imenovalec, nista neodvisni. Torej njuna skupna gostota ne bo produkt njunih mejnih gostot. Za izpeljavo gostote uvedemo 'mešani kot' θ ($\theta = \arcsin \rho$). Ideja je v tem, da je Kendallov $\tau = \frac{2\theta}{\pi}$.

Postavimo

$$Z_{01} = W_1, \quad Z_{02} = W_1 \sin\theta + W_2 \cos\theta,$$

$$T_1 = \sqrt{\frac{n}{C^2}} W_1, \quad T_2 = \sqrt{\frac{n}{C^2}} (W_1 \sin\theta + W_2 \cos\theta).$$

Izrazimo

$$W_1 = \sqrt{\frac{C^2}{n}} T_1, \quad W_2 = \sqrt{\frac{C^2}{n}} \frac{1}{\cos\theta}(T_2 - T_1 \sin\theta).$$

Poznamo standardno normalno gostoto $W_1 + W_2$ v obliki

$$\frac{1}{2\pi} exp\{-\frac{1}{2}(w_1^2 + w_2^2)\}$$

in tako tudi pogojno gostoto T_i glede na $C^2 = z$

$$f(t_1, t_2 | C^2 = z) = \frac{z}{2\pi n \cos\theta} exp\{-\frac{z}{2n\cos^2\theta}(t_1^2 + t_2^2 - 2t_1 t_2 \sin\theta)\}.$$

Sedaj integriramo naslednji produkt, kjer za rezultat dobimo skupno gostoto.

$$\begin{aligned}
f(t_1,t_2) &= \int_0^\infty f(t_1,t_2|C^2=z)f(z;n)dz \\
&= \int_0^\infty \frac{z}{2\pi n\cos\theta} e^{\{-\frac{z}{2n\cos^2\theta}(t_1^2+t_2^2-2t_1t_2\sin\theta)\}} \frac{1}{2\Gamma(\frac{n}{2})} e^{-\frac{z}{2}} \left(\frac{z}{2}\right)^{\frac{n}{2}-1} dz \\
&= \frac{1}{\pi n\cos\theta\Gamma(\frac{n}{2})} \int_0^\infty k^{\frac{n}{2}} e^{-k(1+\Delta)} dk \qquad \left(k=\frac{z}{2}\right) \\
&= \frac{1}{\pi n\cos\theta\Gamma(\frac{n}{2})} \int_0^\infty \left(\frac{m}{1+\Delta}\right)^{n/2} e^{-m}\frac{1}{1+\Delta} dm \qquad \Big(m=k(1+\Delta)\Big) \\
&= \frac{1}{\pi n\cos\theta\Gamma(\frac{n}{2})} \left(\frac{1}{1+\Delta}\right)^{n/2+1} \int_0^\infty m^{n/2}e^{-m}dm \\
&= \frac{1}{2\pi\cos\theta}\frac{1}{(1+\Delta)^{n/2+1}}
\end{aligned}$$

kjer je $\Delta = \frac{t_1^2+t_2^2-2t_1t_2\sin\theta}{n\cos^2\theta}$. Ta skupna porazdelitev ima mejni porazdelitvi, ki sta T-porazdeljeni z n prostostnimi stopnjami. Pri $\theta = 0$ dobimo

$$f(t_1,t_2) = \frac{1}{2\pi}\frac{1}{(1+(t_1^2+t_2^2)/n)^{n/2+1}},$$

ki pa ni produkt dveh mejnih gostot slučajnih spremenljivk, razen če pošljemo $n \to \infty$ in imamo spet normalen primer.

4.2.1. *Alternativna T kopula z enakimi prostostnimi stopnjami.* Sedaj bomo uporabili neodvisne imenovalce. Kot prej imamo

$$T_1 = \sqrt{\frac{n}{C_1^2}}W_1, \quad T_2 = \sqrt{\frac{n}{C_2^2}}(W_1\sin\theta + W_2\cos\theta).$$

Preoblikujemo in dobimo

$$W_1 = \sqrt{\frac{C_1^2}{n}}T_1, \quad W_2 = \frac{1}{\cos\theta}(\sqrt{\frac{C_2^2}{n}}T_2 - \sqrt{\frac{C_1^2}{n}}T_1\sin\theta).$$

Nadaljujemo kot prej

$$f(t_1,t_2|C_1^2=z_1;C_2^2=z_2) = \frac{\sqrt{z_1z_2}}{2\pi n\cos\theta} exp\{-\frac{1}{2n\cos^2\theta}(z_1t_1^2+z_2t_2^2-2t_1t_2\sin\theta\sqrt{z_1z_2})\},$$

ki se zreducira na prejšnjo pogojno gostoto, če je $z_1 = z_2$.

Produktna gostota spremenljivke C_1C_2 je enaka

$$\frac{1}{2^n\Gamma^2(n/2)}(z_1z_2)^{(n/2-1)}e^{-\frac{1}{2}(z_1+z_2)}.$$

Sedaj dobimo skupno gostoto $f(t_1,t_2)$ z integriranjem produkta pogojne gostote $(f(t_1,t_2|C_1^2=z_1;C_2^2=z_2))$ in produktne gostote kot prej. Rezultat je mogoče izraziti s hipergeometričnimi funkcijami tako, kot sledi. Uvedemo naslednje vmesne spremenljivke:

$$\alpha_1 = 1+\frac{t_1^2}{n\cos^2(\theta)}, \quad \alpha_2 = 1+\frac{t_2^2}{n\cos^2(\theta)}, \quad \gamma = \frac{2t_1t_2\sin(\theta)}{n\cos^2(\theta)},$$

in normalizirano konstanto

$$C' = \frac{1}{cos(\theta)\pi n\Gamma(n/2)^2}.$$

Gostota je tako dana kot

$$f(t_1,t_2) = C'(\alpha_1\alpha_2)^{-\frac{n}{2}-1}\Bigg[\Gamma\left(\frac{n+1}{2}\right)^2 {}_2F_1\left(\frac{n+1}{2}, \frac{n+1}{2}; \frac{1}{2}; \frac{\gamma^2}{4\alpha_1\alpha_2}\right)\sqrt{\alpha_1}\sqrt{\alpha_2}+$$

$$+\gamma\Gamma\left(\frac{n}{2}+1\right)^2 {}_2F_1\left(\frac{n}{2}+1, \frac{n}{2}+1; \frac{3}{2}; \frac{\gamma^2}{4\alpha_1\alpha_2}\right)\Bigg].$$

Pošljemo $\theta \to 0$ in dobimo

$$f(t_1,t_2) = \frac{\Gamma[\frac{n+1}{2}]^2}{\Gamma[\frac{n}{2}]^2 a\pi}\left(\frac{1}{1+t_1^2/n}\right)^{(n+1)/2}\left(\frac{1}{1+t_2^2/n}\right)^{(n+1)/2}.$$

V primeru neodvisnosti lahko seveda obnovimo produktno strukturo, a s težjo dvorazsežno funkcijo gostote. Dobili smo produkt dveh gostot neodvisnih slučajnih spremenljivk.

4.3. **Studentova T porazdelitev z različnimi marginalnimi porazdelitvami (z različnimi prostostnimi stopnjami).** S pomočjo zgornjega pristopa si bomo ogledali definicije večrazsežne T porazdelitve, kjer imajo marginalne T porazdelitve različne prostostne stopnje. Za boljše razumevanje se bomo ponovno omejili le na dvorazsežen primer. V vseh primerih bomo imeli stopnji prostosti n_1 in n_2, ki nista nujno enaki, niti ne nujno celi.

4.3.1. *Skupinski pristop.* Najprej bomo opisali pristop, ki je poimenovan skupinski pristop.

Imamo dve neodvisni standardni normalni slučajni spremenljivki W_1 in W_2 in transformaciji

$$Z_{01} = \alpha W_1 + \beta W_2, \qquad Z_{02} = \gamma W_1 + \delta W_2$$

z varianco 1: $\alpha^2 + \beta^2 = 1 = \gamma^2 + \delta^2$. Nato konstruiramo

$$T_1 = Z_{01}\sqrt{\frac{n_1}{C_1^2}}, \qquad T_2 = Z_{02}\sqrt{\frac{n_2}{C_2^2}},$$

kjer sta C_i^2 odvisni χ^2 spremenljivki s prostostno stopnjo n_i. Simuliramo jo tako, da vzamemo enakomerno zvezno porazdeljen vektor U in postavimo

$$C_i^2 = G_{n_i}^{-1}(U),$$

kjer je G_{n_i} porazdelitvena funkcija χ^2 s parametrom n_i. Skupinski pristop se nanaša na to, da v večrazsežnem primeru zberemo skupaj vse marginalne porazdelitve z enako prostostno stopnjo in uporabimo isti χ^2 imenovalec, da spremenimo normalno spremenljivko v Studentovo T spremenljivko znotraj vsake skupine. Rečemo lahko, da je to "tesen" skupinski pristop, ker v vseh skupinah uporabimo isti U. Torej imamo

$$(t_{n_1}(T_1), ..., t_{n_1}(T_{s_1}), t_{n_2}(T_{s_1+1}), ..., t_{n_2}(T_{s_1+s_2})),$$

kjer sta s_1 in s_2 velikosti skupin. Velja, da ima vektor $(\sqrt{\frac{n_1}{C_1^2}}Z_{01}, ..., \sqrt{\frac{n_1}{C_1^2}}Z_{s_1})$ s_1-razsežno T porazdelitev, z n_1 prostostnimi stopnjami.

Vendar pa obstajata vsaj še dve drugi poti. Prva in dokaj očitna možnost bi bila, da vzamemo, da so C_i neodvisne. Temu rečemo "ohlapni" skupinski pristop. Ta pristop je dovolj pregleden. Vendar pa ima značilnost, na katero je treba opozoriti. Dokler je $n_2 \neq n_1$ in $n_2 \to n_1$, skupine ostajajo različne. Druga možnost je posplošitev Jonesove metode, kar si bomo pogledali spodaj.

Ogledali smo si simulacijski del konstrukcije. Konstrukcija kopule sledi z aplikacijo Sklarovega izreka in marginalnih porazdelitev, kjer ima spremenljivka T_i stopnjo prostosti n_i. Pri tem lahko uporabimo spodnje formule.

Spomnimo se, da je marginalna porazdelitvena funkcija T spremenljivke s stopnjo n podana z

$$F_n(x) = \int_{-\infty}^{x} f_n(t)dt = \frac{1}{\sqrt{n\pi}} \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})} \int_{-\infty}^{x} \frac{1}{(1+\frac{t^2}{n})^{\frac{n+1}{2}}} dt.$$

Za splošen n lahko zapišemo $F_n(x)$ s pomočjo hipergeometričnih funkcij. Integracija v programu Mathematica nam vrne:

$$F_n(x) = \frac{1}{2} + \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\Gamma(\frac{n}{2})} x \; {}_2F_1\left(\frac{1}{2}, \frac{n+1}{2}; \frac{3}{2}; -\frac{x^2}{n}\right).$$

Zapis pa je mogoč tudi s pomočjo beta funkcije:

$$F_n(x) = \frac{1}{2}\left(1 + sgn(x)\left(1 - I_{(\frac{n}{x^2+n})}\left(\frac{n}{2}, \frac{1}{2}\right)\right)\right),$$

pri čemer je $sgn(x)$ enak $+1$ za $x > 0$ in -1 za $x < 0$. $I_x(a,b)$ pa je definiran:

$$I_x(a,b) = \frac{B_x(a,b)}{B(a,b)}.$$

Pri tem je $B(a,b)$ običajna beta funkcija, $B_x(a,b)$ pa je enaka:

$$B_x(a,b) = \int_0^x t^{a-1}(1-t)^{b-1} dt.$$

Za soda števila n je mogoče F_n zapisati le z algebraičnimi izrazi:

$$F_n(x) = \frac{1}{2} + x\left(\frac{x^2}{n} + 1\right)^{\frac{1-n}{2}} \left(\sum_{k=0}^{\frac{n}{2}-1} x^{2k} a(k,n)\right),$$

kjer velja

$$a(0,n) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\Gamma(\frac{n}{2})}, \qquad a(k,n) = \frac{(n-2k)a(k-1,n)}{(2k+1)n}.$$

4.3.2. *Jonesov pristop.* Naslednji eleganten pristop je posplošitev Jonesovega pristopa. Razvrstimo stopnje prostosti po naraščajočem vrstnem redu in uporabimo aditivnost χ^2 porazdelitve. Naj bo $n_1 \leq n_2$ in označimo $a = n_1$ in $b = n_2 - n_1$. V prvotni Jonesovi specifikaciji je odvisnost le preko χ^2 spremenljivke. Torej v našem zapisu je njegov model v obliki

$$T_1 = W_1 \sqrt{\frac{a}{C_1^2}}, \qquad T_2 = W_2 \sqrt{\frac{a+b}{C_1^2 + C_2^2}},$$

kjer ima C_1^2 hi-kvadrat porazdelitev z a prostostnimi stopnjami in C_2 hi-kvadrat porazdelitev z b prostostnimi stopnjam. Tako ima model ničelno korelacijo, kljub odvisnosti v imenovalcih. Vendar pa lahko posplošimo tako, da vključimo odvisnost v števcu z uporabo mešanih spremenljivk

$$T_1 = Z_{01}\sqrt{\frac{a}{C_1^2}}, \qquad T_2 = Z_{02}\sqrt{\frac{a+b}{C_1^2+C_2^2}}.$$

Kopule, ki izhajajo iz posplošene Jonesove metode, dobimo z uporabo marginalnih porazdelitvenih funkcij spremenljivk T_1 s stopnjo prostosti a in T_2 s stopnjo prostosti $a+b$.

4.4. **Kanonična Studentova – Normalna porazdelitev.** Razvidno je, da so tesen skupinski, ohlapni skupinski in Jonesov pristop različni. Imajo pa skupno limito, ko pošljemo $n_2 \to \infty$. Opozoriti je treba na to, da je v skupinskem pristopu, ko $C_2^2 \to 1$, in v primeru Jonesa limitni primer enak, ko $b \to \infty$. Torej imamo enako *kanonično* limito za te pristope, kjer vzamemo

$$T_1 = Z_{01}\sqrt{\frac{n_1}{C_1^2}}, \qquad T_2 = Z_{02}.$$

To nam da recept za simulacijo Student-Normalne porazdelitve s korelacijo. Enakost limite ima pomembno vlogo pri kanonični Student-Normalni porazdelitvi, ki jo bomo v nadaljevanju podrobneje karakterizirali. Pristopa, ki smo si ju prej ogledali, se združita v eno, ko pošljemo $n_2 \to \infty$. Najprej parametriziramo povezave med dvema porazdelitvama z rotacijo kota θ. Z W_1, W_2 kot prej postavimo

$$T_1 = \frac{W_1}{\sqrt{C_1^2/a}}, \qquad T_2 = W_1 \sin\theta + W_2 \cos\theta.$$

Dobimo

$$W_1 = \sqrt{\frac{C_1^2}{a}}T_1, \qquad W_2 = \frac{1}{\cos\theta}\big(T_2 - \sqrt{\frac{C_1^2}{a}}T_1 \sin\theta\big).$$

Poznamo gostoto spremenljivke $W_1 + W_2$ v obliki

$$\frac{1}{2\pi}exp\{-\frac{1}{2}(w_1^2+w_2^2)\}.$$

Pogojna gostota vektorja (T_1, T_2) glede na $C_1^2 = z$ je tako dana kot

$$f(t_1,t_2|C_1^2=z) = \frac{1}{2\pi\cos\theta}\sqrt{\frac{z}{a}}exp\bigg\{-\frac{1}{2\cos^2\theta}\bigg(\frac{z}{a}t_1^2+t_2^2-2t_1t_2\sin\theta\sqrt{\frac{z}{a}}\bigg)\bigg\}.$$

Brezpogojna skupna gostota vektorja (T_1, T_2) je podana z integracijo pogojne gostote glede na gostoto hi-kvadrat po z.

$$\begin{aligned} f(t_1,t_2) &= \int_0^\infty f(t_1,t_2|\chi^2=z)\frac{1}{\Gamma(a/2)2^{a/2}}z^{a/2-1}e^{-z/2}dz \\ &= c\int_0^\infty z^{(a-1)/2}exp\bigg\{-\frac{z}{2}-\frac{1}{2\cos^2\theta}\bigg(\frac{z}{a}t_1^2+t_2^2-2t_1t_2\sin\theta\sqrt{\frac{z}{a}}\bigg)\bigg\}dz, \end{aligned}$$

kjer je $c = 1/\big(2\pi 2^{a/2}\Gamma(a/2)\sqrt{a}\cos\theta\big)$. Če malo poenostavimo in postavimo $z = q^2$ dobimo

$$f(t_1,t_2) = 2ce^{-t_2^2/(2\cos^2\theta)}\int_0^\infty q^a exp\bigg\{-\frac{q^2}{2}\bigg(1+\frac{t_1^2}{a\cos^2\theta}\bigg)+\frac{t_1t_2q\sin\theta}{\sqrt{a}\cos^2\theta}\bigg\}dq.$$

Zapišimo sedaj nekaj osnovnih momentov.

4.4.1. *Momenti.* Spremenljivka T_1 je porazdeljena Studentovo in je simetrična porazdelitev. Zato so vsi njeni lihi momenti enaki 0. Prepričamo se, da je prvi moment enak 0.

$$\begin{aligned}
E[T_1] &= \int_{-\infty}^{\infty} t \frac{1}{\sqrt{a\pi}} \frac{\Gamma(\frac{a+1}{2})}{\Gamma(\frac{a}{2})} \left(1 + \frac{t^2}{a}\right)^{-\frac{a+1}{2}} dt \\
&= \int_{0}^{\infty} t \frac{2}{\sqrt{a\pi}} \frac{\Gamma(\frac{a+1}{2})}{\Gamma(\frac{a}{2})} \left(1 + \frac{t^2}{a}\right)^{-\frac{a+1}{2}} dt \\
&= \frac{2}{\sqrt{a\pi}} \frac{\Gamma(\frac{a+1}{2})}{\Gamma(\frac{a}{2})} \frac{a}{2} \int_{0}^{\infty} z^{-\frac{a+1}{2}} dz = 0, \quad \forall a > 1.
\end{aligned}$$

Sedaj pa izračunajmo še drugi moment.

$$\begin{aligned}
E[T_1^2] &= \int_{-\infty}^{\infty} t^2 \frac{1}{\sqrt{a\pi}} \frac{\Gamma(\frac{a+1}{2})}{\Gamma(\frac{a}{2})} \left(1 + \frac{t^2}{a}\right)^{-\frac{a+1}{2}} dt \\
&= \frac{2}{\sqrt{a\pi}} \frac{\Gamma(\frac{a+1}{2})}{\Gamma(\frac{a}{2})} \int_{0}^{\infty} \frac{t^2}{(1+\frac{t^2}{a})^{(a+1)/2}} dt \\
&= \frac{2}{\sqrt{a\pi}} \frac{\Gamma(\frac{a+1}{2})}{\Gamma(\frac{a}{2})} \int_{0}^{\infty} \frac{\sqrt{ax}}{(1+x)^{(a+1)/2}} \frac{a}{2} dx \\
&= \frac{2a}{2\sqrt{\pi}} \frac{\Gamma(\frac{a+1}{2})}{\Gamma(\frac{a}{2})} \int_{0}^{\infty} \frac{\sqrt{x}}{(1+x)^{(a+1)/2}} dx \\
&= \frac{a}{\sqrt{\pi}} \frac{\Gamma(\frac{a+1}{2})}{\Gamma(\frac{a}{2})} B\left(\frac{3}{2}, \frac{a-2}{2}\right) \\
&= \frac{a}{\sqrt{\pi}} \frac{\Gamma(\frac{a+1}{2})}{\Gamma(\frac{a}{2})} \frac{\Gamma(\frac{3}{2})\Gamma(\frac{a-2}{2})}{\Gamma(\frac{a+1}{2})} \\
&= \frac{a}{a-2}, \quad \forall a > 2.
\end{aligned}$$

Pri izračunu smo si pomagali z naslednjo lastnostjo beta funkcije

$$B(\alpha, \beta) = \int_{0}^{\infty} \frac{t^{\alpha-1}}{(1+t)^{\alpha+\beta}} dt, \qquad za \quad \alpha, \beta > 0.$$

Spremenljivka T_2 je porazdeljena standardno normalno. Iz njene definicije tako njena prva dva momenta že poznamo.

$$E[T_2] = 0, \quad \forall a > 0$$

$$E[T_2^2] = 1, \quad \forall a > 0.$$

Zanima nas samo še skupni moment.

$$E[T_1 T_2] = \frac{\sqrt{a}\sin(\theta)\Gamma(\frac{a-1}{2})}{\sqrt{2}\Gamma(\frac{a}{2})}, \quad \forall a > 1.$$

Iz teh rezultatov lahko sklepamo, da Pearsonov korelacijski koeficient ρ obstaja za $a > 2$ in je enak

$$\rho = \sin(\theta) \frac{\Gamma(\frac{a-1}{2})}{\Gamma(\frac{a}{2})} \sqrt{\frac{a}{2} - 1}.$$

Ko pošljemo $a \to +\infty$, gre $\rho \to \sin(\theta)$, tako da dobimo običajni rezultat dvorazsežne normalne spremenljivke. Za končne a pa imamo $|\rho| < |\sin(\theta)|$, kjer z naraščajočim a, $\rho/\sin(\theta)$ monotono narašča.

4.4.2. *Nekoreliran primer.* Če ločimo spremenljivke tako, da postavimo $\theta = 0$, dobimo produktno gostoto:

$$f(t_1, t_2) = \frac{1}{\sqrt{a\pi}} \frac{\Gamma(\frac{a+1}{2})}{\Gamma(\frac{a}{2})} \left(1 + \frac{t_1^2}{a}\right)^{(a+1)/2} \times \frac{1}{\sqrt{2\pi}} e^{-t_2^2/2}.$$

4.4.3. *Koreliran primer.* V splošnem primeru imamo vmesne spremenljivke

$$\alpha = 1 + \frac{t_1^2}{a\cos^2\theta}, \qquad \beta = \frac{t_1 t_2 \sin\theta}{\sqrt{a}\cos^2\theta}, \qquad \gamma = \frac{\beta}{\sqrt{\alpha}}$$

s pomočjo katerih je podana gostota

$$f(t_1, t_2) = \frac{2^{\frac{a-1}{2} - \frac{a}{2}} e^{\frac{\gamma^2}{2} - \frac{1}{2} sec^2(\theta) t_2^2} \alpha^{-\frac{a}{2} - \frac{1}{2}}}{\sqrt{a}\pi\Gamma(\frac{a}{2})\cos\theta} \times$$

$$\times \left[\sqrt{2}\gamma\Gamma\left(\frac{a}{2} + 1\right) {}_1F_1\left(\frac{1}{2} - \frac{a}{2}; \frac{3}{2}; -\frac{\gamma^2}{2}\right) + \Gamma\left(\frac{a+1}{2}\right) {}_1F_1\left(-\frac{a}{2}; \frac{1}{2}; -\frac{\gamma^2}{2}\right)\right],$$

kjer je $sec(\theta) = \frac{1}{\cos(\theta)}$. Koristno je, da vidimo, kako se to poenostavi, ko je a majhno celo število. Takrat je hipergeometrijska funkcija lahko poenostavljena v smislu normalne porazdelitvene funkcije. Poglejmo si nekaj primerov za majhne a.

- a = 1:

$$\alpha = 1 + \frac{t_1^2}{\cos^2\theta}, \qquad \beta = \frac{t_1 t_2 \sin\theta}{\cos^2\theta}, \qquad \gamma = \frac{\beta}{\sqrt{\alpha}}$$

$$f(t_1, t_2) = \frac{1}{\cos\theta\sqrt{2\pi}} e^{-t_2^2/(2\cos^2\theta)} \frac{1}{\pi\alpha} \left[1 + e^{\gamma^2}\gamma\sqrt{2\pi}\Phi(\gamma)\right]$$

- a = 2:

$$\alpha = 1 + \frac{t_1^2}{2\cos^2\theta}, \qquad \beta = \frac{t_1 t_2 \sin\theta}{\sqrt{2}\cos^2\theta}, \qquad \gamma = \frac{\beta}{\sqrt{\alpha}}$$

$$f(t_1, t_2) = \frac{e^{-\frac{1}{2} sec^2(\theta) t_2^2}\left(\gamma + e^{\frac{\gamma^2}{2}}\sqrt{2\pi}(\gamma^2 + 1)\Phi(\gamma)\right)}{\cos(\theta) 2\sqrt{2}\pi\alpha^{3/2}}$$

- a = 3:

$$\alpha = 1 + \frac{t_1^2}{3\cos^2\theta}, \qquad \beta = \frac{t_1 t_2 \sin\theta}{\sqrt{3}\cos^2\theta}, \qquad \gamma = \frac{\beta}{\sqrt{\alpha}}$$

$$f(t_1, t_2) = \frac{e^{-\frac{1}{2} sec^2(\theta) t_2^2}\left(\gamma^2 + e^{\frac{\gamma^2}{2}}\sqrt{2\pi}(\gamma^2 + 3)\Phi(\gamma)\gamma + 2\right)}{\cos(\theta)\sqrt{6}\pi^{3/2}\alpha^2}$$

- a = 4:

$$\alpha = 1 + \frac{t_1^2}{4\cos^2\theta}, \qquad \beta = \frac{t_1 t_2 \sin\theta}{\sqrt{4}\cos^2\theta}, \qquad \gamma = \frac{\beta}{\sqrt{\alpha}}$$

$$f(t_1, t_2) = \frac{e^{-\frac{1}{2} sec^2(\theta) t_2^2}\left(\gamma(\gamma^2 + 5) + e^{\frac{\gamma^2}{2}}\sqrt{2\pi}(\gamma^4 + 6\gamma^2 + 3)\Phi(\gamma)\right)}{8\cos(\theta)\pi\alpha^{5/2}}$$

5. Aplikacije v financah

Dana aplikacija je črpana iz vira [15]. Ena izmed najmočnejših aplikacij kopul se tiče upravljanja s tveganji. Pri tem lahko obravnavamo štiri probleme: skupna izguba, programi testiranja izjemnih situacij, privzeto modeliranje in merjenje operativnega tveganja. Sedaj bomo pokazali, kako lahko uporabimo kopulo za združitev porazdelitev izgub in izračun VaR-a (Value-at-Risk). VaR je mera tveganja, ki ocenjuje tveganje izgube določenega portfelja finančnih sredstev.

Omejili se bomo le na zvezen primer. Računski vidik je eden izmed glavnih tem iz vidika industrijske perspektive. Tiče se problema ocenjevanja parametrov kopul in potem je tu še vprašanje simulacije. Ta dva problema ne moremo z lahkoto rešiti pri velikih dimenzijah za dano kopulo, kajti kopula ni prilagodljiva za izračun VaR-a. Simulacije, ki temeljijo na Gaussovi kopuli so enostavnejše, zato je Gaussova kopula pogosto uporabljena. Ravno tako so dovolj enostavne simulacije na Studentovi kopuli, zato je tudi slednja dober kandidat. Ker ju lahko kombiniramo s poljubnimi marginalnimi porazdelitvami, tako dobimo veliko večrazsežnih porazdelitev, ki jih lahko uporabimo za izračun tveganja portfelja.

Zapišimo še definicijo VaR-a. Naj bo $P(t)$ vektor cen sredstev ob času t, in naj bo a portfelj. Eno-obdobni VaR s stopnjo zaupanja α je definiran kot $VaR = F^{-1}(1-\alpha)$, kjer je F porazdelitev slučajne spremenljivke $a^T(P(t+1) - P(t))$.

Poglejmo si primer iz donosnosti sredstev, kjer so podatki vzeti iz baze London Metal Exchange iz januarja 1988, kjer je AL = aluminijeva zlitina, CU = baker, NI = nikelj, PB = svinec in AL-15 = 15-mesečna terminska cena aluminijeve zlitine. V naslednji tabeli so predstavljeni portfelji (P_1, P_2, P_3), kjer negativna številka pomeni kratko pozicijo.

	AL	AL-15	CU	NI	PB
P_1	1	1	1	1	1
P_2	-1	-1	-1	-1	-1
P_3	2	1	-3	4	5

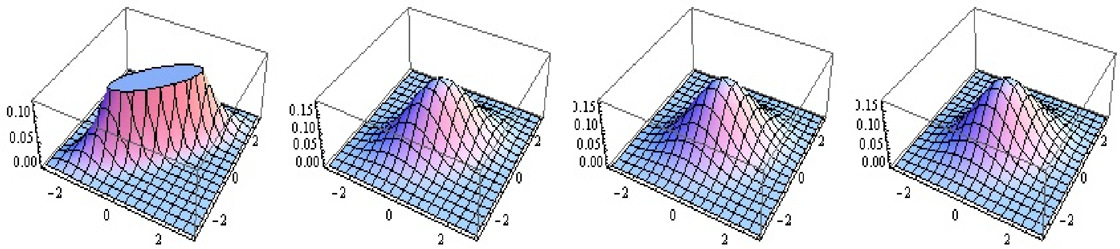
Oglejmo si korelacijsko matriko, ki je izračunana s kanonično metodo največjega verjetja za Gaussovo kopulo (torej pri privzetku, da so marginalne porazdelitve Gaussove).

	AL	AL-15	CU	NI	PB
AL	1.000	0.850	0.492	0.386	0.354
AL-15		1.000	0.429	0.346	0.316
CU			1.000	0.409	0.359
NI				1.000	0.333
PB					1.000

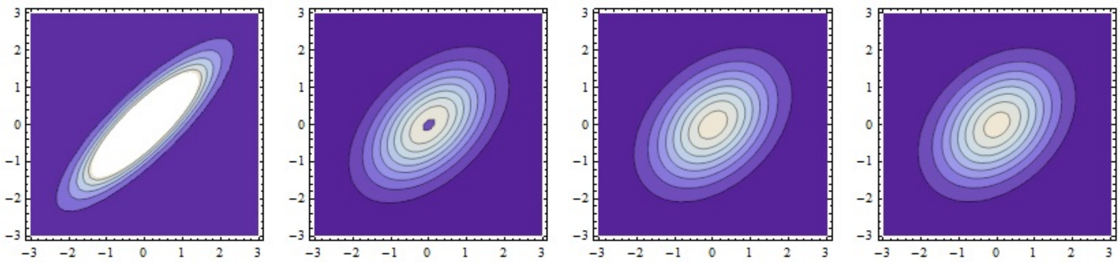
Sedaj si lahko ogledamo nekatere pripadajoče dvorazsežne Gaussove kopule z danimi korelacijami.

Na sliki 5 vidimo gostote Gaussove kopule, na sliki 6 pa njene obrise. Na prvi izmed slik je prva spremenljivka podatkov AL, druga pa podatkov AL-15 s korelacijskim koeficientom $0,85$; oziroma korelacijsko matriko $\mathbf{P} = \begin{bmatrix} 1 & 0,85 \\ 0,85 & 1 \end{bmatrix}$.

Naslednja kopula ima ravno tako prvo spremenljivko podatkov AL, drugo pa CU. Njun korelacijski koeficient je $0,492$. Sledita še kopula s podatki AL in NI s korelacijskim koeficientom $0,386$ ter kopula s podatki AL in PB s korelacijskim koeficientom $0,354$.



Slika 5. Dvorazsežne Gaussove kopule.



Slika 6. Obrisi kopul.

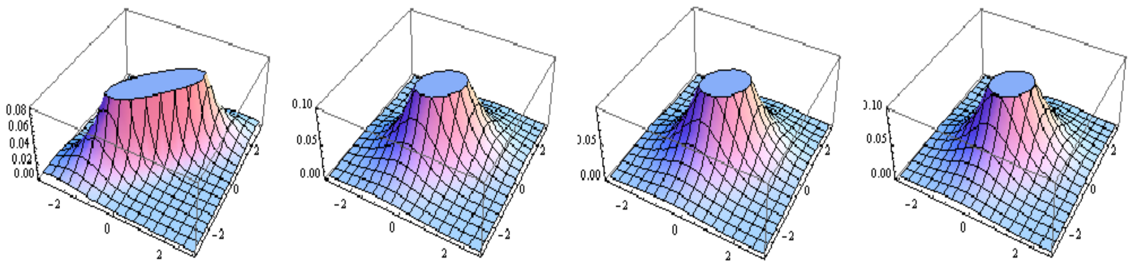
Sedaj pa si poglejmo še korelacijsko matriko, ki je izračunana s kanonično metodo največjega verjetja za Studentovo kopulo. Vidimo, da so izračuni drugačni.

$\nu = 1$	AL	AL-15	CU	NI	PB
AL	1.000	0.820	0.326	0.254	0.189
AL-15		1.000	0.271	0.223	0.164
CU			1.000	0.267	0.219
NI				1.000	0.195
PB					1.000

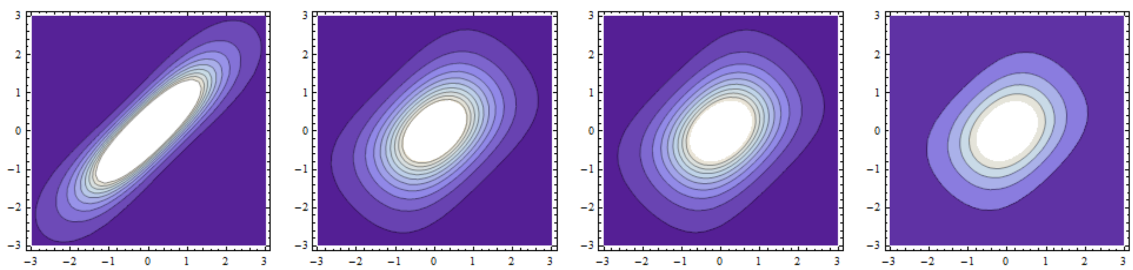
ν =5	AL	AL-15	CU	NI	PB
AL	1.000	0.893	0.493	0.410	0.350
AL-15		1.000	0.437	0.373	0.320
CU			1.000	0.423	0.362
NI				1.000	0.337
PB					1.000

ν =100	AL	AL-15	CU	NI	PB
AL	1.000	0.861	0.496	0.394	0.358
AL-15		1.000	0.436	0.354	0.322
CU			1.000	0.415	0.363
NI				1.000	0.337
PB					1.000

Poglejmo si dvorazsežne Studentove kopule z danimi korelacijami, za primer, ko je $\nu = 5$. Na sliki 7 so torej gostote Studentove kopule s stopnjo prostosti 5, na sliki 8 pa njeni obrisi. Na vseh slikah je prva spremenljivka podatkov AL. Na prvi izmed štirih slik je dvorazsežna Studentova kopula s korelacijskim koeficientom $0,893$, torej korelacijsko matriko $\mathbf{P} = \begin{bmatrix} 1 & 0,893 \\ 0,893 & 1 \end{bmatrix}$, za podatke AL in AL-15. Naslednja ima korelacijski koeficient 0,493 za spremenljivki podatkov AL in CU. Sledi kopula s korelacijskim koeficientom 0,41 za podatke AL in NI. Na koncu pa še koeficient 0,35 za spremenljivki podatkov AL in PB.



Slika 7. Dvorazsežne Studentove kopule.



Slika 8. Obrisi Studentovih kopul.

Predvidevajmo, da merimo kapital z enodnevnim VaR-om portfelja. V praksi je 99% stopnja zaupanja za VaR ocenjena z zgodovinskim pristopom. Dejanske baze podatkov so premajhne za izračun zgodovinskega VaR-a z višjo stopnjo zaupanja. Projekti gospodarskega razporejanja kapitala pa temeljijo na višjih stopnjah zaupanja, zaradi česar si pri izračunu pomagamo s kopulami. Tu so klasične stopnje zaupanja.

ocena	BBB	A	AA	AAA
stopnja zaupanja α	99.75	99.9	99.95	99.97

VaR je izračunan glede na to, da so mejne porazdelitvene funkcije porazdeljene standardno normalno, kopuli pa sta različni. Najprej so rezultati, kjer je kopula Gaussova.

	90 %	95 %	99 %	99.5 %	99.9 %
P_1	7.265	9.297	13.15	14.59	17.54
P_2	7.263	9.309	13.17	14.59	17.65
P_2	13.92	17.92	25.24	27.98	33.40

Poglejmo si sedaj še rezultate, ki so dobljeni s Studentovo kopulo za različne prostostne stopnje.

$\nu = 1$	90 %	95 %	99 %	99.5 %	99.9 %
P_1	5.696	7.977	13.22	15.46	20.16
P_2	5.709	7.980	13.19	15.46	20.15
P_2	13.43	19.34	34.32	40.65	55.12

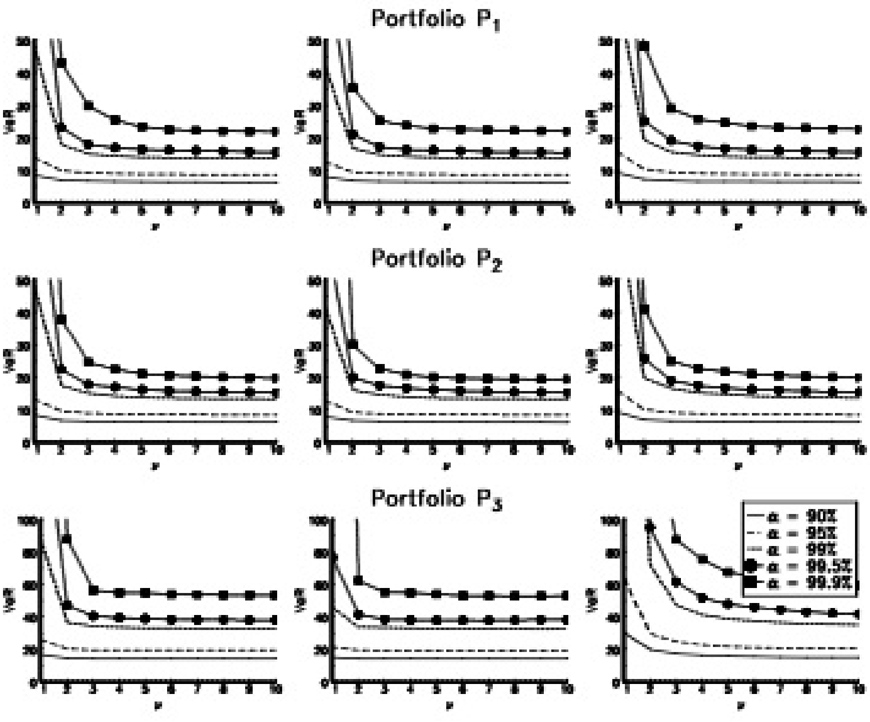
$\nu = 5$	90 %	95 %	99 %	99.5 %	99.9 %
P_1	7.081	9.283	13.88	15.68	19.58
P_2	7.109	9.351	13.94	15.72	19.57
P_2	13.31	17.63	27.18	31.18	40.78

$\nu = 100$	90 %	95 %	99 %	99.5 %	99.9 %
P_1	7.293	9.387	13.27	14.79	17.56
P_2	7.297	9.427	13.42	14.78	17.56
P_2	13.85	17.81	25.32	28.11	34.07

Vidimo, da je pri Gaussovi kopuli ekonomski kapital z 99.9 % intervalom zaupanja nižji kot pri Studentovi. To pomeni, da ima odvisnostna struktura (oziroma funkcija kopule) velik vpliv na izračun VaR-a.

Poglejmo si še finančno bolj realističen primer, ko imajo marginalne porazdelitve težke repe. Privzemimo, da je odvisnostna struktura podana s Studentovo kopulo z dvema prostostnima stopnjama. Prej smo privzeli, da so bile vse marginalne porazdelitve Gaussove. Sedaj privzamemo, da je en nabor finančnih donosov porazdeljen Studentovo s prostostno stopnjo ν, ostali pa Gaussovo. Levi grafi na sliki 9 pripadajo podatkom AL, srednji podatkom CU in desni podatkom PB.

Iz slike 9 je razviden velik vpliv težkih repov na VaR pri visokih vrednostih stopnje zaupanja.



SLIKA 9. Težki repi.

Literatura

[1] J. Globevnik in M. Brojan, *Analiza 2*, verzija 10. 8. 2010, [ogled 25. 2. 2013], dostopno na `http://www.fmf.uni-lj.si/~globevnik/skriptaII.pdf`.

[2] *Copula (probability theory)*, [ogled 25. 2. 2013], dostopno na `http://http://en.wikipedia.org/wiki/Copula_(probability_theory)`.

[3] *Student's t-distribution*, [ogled 25. 2. 2013], dostopno na `http://http://en.wikipedia.org/wiki/Student's_t-distribution`.

[4] W. T. Shaw, *Sampling Student's T distribution - use of the inverse cumulative distribution function*, Journal of Computational Finance, Vol. 9, No. 4 (2006)

[5] *Slučajna spremenljivka*, [ogled 15. 3. 2013], dostopno na `http://sl.wikipedia.org/wiki/Slu%C4%8Dajna_spremenljivka`.

[6] *Porazdelitev hi-kvadrat*, [ogled 15. 3. 2013], dostopno na `http://sl.wikipedia.org/wiki/Porazdelitev_hi-kvadrat`.

[7] D. B. Ntwiga, *Copulas in Finance*, African institute for Mathematical Sciences Report, Maj, 2004.

[8] R. B. Nelsen, *An introduction to Copulas*, Second Edition, Springer, New York, 2006.

[9] M. Abramowitz in I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Dover Publications, New York, 1972.

[10] J. S. Rosenthal, *A first look at rigorous probability theory*, Second Edition, World Scientific, Singapore, 2009.

[11] L. J. Slaterl, *Generalized hypergeometric functions*, Cambridge : at the University Press, 1966.

[12] W. H. Kruskal, *Ordinal measures of association*, Journal of the American Statistical Association, 53, str. 814 (1958)

[13] H.- B. Fang, K.- T. Fang, S. Kotz, *The meta-elliptical distribution with given marginals*, Journal of Multivariate Analysis, 82, 1-16 (2002)

[14] D. Anevski, *Riemann-Stieltjes integrals*, Mathematical Sciences, Lund University, (2012)

[15] E. Bouyé, V. Durrleman, A. Nikeghbali, G. Riboulet, T. Roncalli *Copulas for Finance - A Reading Guide and Some Applications*, Financial Econometrics Research Centre, City University Business School London, (2000)

[16] W. T. Shaw, K. T. A. Lee, *Copula Methods vs Canonical Multivariate Distributions: the multivariate Student T distribution with general degrees of freedom*, KCL Working paper, [ogled 25. 2. 2013], dostopno na `http://www.kcl.ac.uk/nms/depts/mathematics/research/finmath/publications/2006ShawLee.pdf`.