

UNIVERZITA PALACKÉHO V OLOMOUCI
PŘÍRODOVĚDECKÁ FAKULTA

DIPLOMOVÁ PRÁCE

Modelování závislostí více proměnných pomocí
kopul



Katedra matematické analýzy a aplikací matematiky
Vedoucí diplomové práce: **doc. RNDr. Karel Hron, Ph.D.**
Vypracovala: **Bc. Kateřina Zelinková**
Studijní program: N1103 Aplikovaná matematika
Studijní obor: Aplikace matematiky v ekonomii
Forma studia: prezenční
Rok odevzdání: 2017

BIBLIOGRAFICKÁ IDENTIFIKACE

Autor: Bc. Kateřina Zelinková

Název práce: Modelování závislostí více proměnných pomocí kopul

Typ práce: Diplomová práce

Pracoviště: Katedra matematické analýzy a aplikací matematiky

Vedoucí práce: doc. RNDr. Karel Hron, Ph.D.

Rok obhajoby práce: 2017

Abstrakt: Teorie kopul se užívá k modelování vztahů mezi dvou a více proměnnými, s četnými aplikacemi od přírodních věd až po ekonomická data. Cílem této diplomové práce je popsat teoretické základy kopul, jejich vlastnosti a typy. Dále jsou zde uvedeny metody odhadů parametrů kopul a testy vhodnosti. Následně jsou tyto poznatky využity při modelování daného datového souboru a při nastínění problematiky ohodnocování CDO.

Klíčová slova: Kopuly, závislost proměnných, CDO, eliptické kopuly, archimédovské kopuly, fundamentální kopuly.

Počet stran: 83

Počet příloh: 1

Jazyk: český

BIBLIOGRAPHICAL IDENTIFICATION

Author: Bc. Kateřina Zelinková

Title: Modeling dependences of variables using copulas

Type of thesis: Master's

Department: Department of Mathematical Analysis and Application of Mathematics

Supervisor: doc. RNDr. Karel Hron, Ph.D.

The year of presentation: 2017

Abstract: The copula theory is used to model the relationships of two and more variables with many applications to natural sciences and economical data. The thesis focuses on the basic theory of copulas, their characteristics and types. Furthermore, the work describes estimation methods of copula parameters and the goodness of fit tests. This knowledge is afterwards applied to modelling of specific dataset and to outline the process of the CDO valuation.

Key words: Copulas, dependence of variables, CDO, elliptical copulas, archimedean copulas, fundamental copulas.

Number of pages: 83

Number of appendices: 1

Language: Czech

Prohlášení

Prohlašuji, že jsem diplomovou práci zpracovala samostatně pod vedením doc. RNDr. Karla Hrona, Ph.D. a všechny použité zdroje jsem uvedla v seznamu literatury.

V Olomouci dne

.....

podpis

Obsah

Úvod	9
1 Základní pojmy	12
2 Kopuly	15
2.1 Sklarova věta	17
2.2 Fréchet-Hoeffdingovy meze	20
2.3 Vlastnosti kopul	22
2.3.1 Symetrie kopul	23
2.4 Chvosty kopul	25
2.4.1 Monotonie chvostů	25
2.4.2 Závislost chvostů	27
2.4.3 Asymetrie chvostů	28
3 Závislost	29
3.1 Pearsonův korelační koeficient	29
3.2 Konkordance	31
3.3 Kendallovo τ	31
3.4 Spearmanovo ρ_S	32
4 Typy kopul	34
4.1 Fundamentální kopuly	35
4.1.1 Součinná kopula	35
4.1.2 Horní kopula	35
4.1.3 Dolní kopula	36
4.2 Eliptické kopuly	37
4.2.1 Gaussova kopula	38
4.2.2 Studentova t-kopula	39
4.3 Archimédovské kopuly	41
4.3.1 Gumbelova kopula	42
4.3.2 Claytonova kopula	44
4.3.3 Frankova kopula	45

5	Odhady parametrů kopul	49
5.1	Metody maximální věrohodnosti	49
5.1.1	Inference marginálních distribucí	50
5.1.2	Kanonická metoda maximální věrohodnosti	51
6	Testy dobré shody	53
6.1	Cramer von Misesův test	54
6.1.1	Postup parametrického bootstrapu	55
6.2	Anderson - Darlingův test	56
7	Praktická část	57
7.1	Aplikace modelování kopul na reálných datech	57
7.2	Hodnocení CDO	63
	Závěr	72
	Dodatek	74
	Literatura	79

Seznam obrázků

4.1	Součinnová kopula	36
4.2	Horní kopula	36
4.3	Dolní kopula	37
4.4	Gaussova kopula	39
4.5	Studentova kopula	41
4.6	Gumbelova kopula	43
4.7	Claytonova kopula	45
4.8	Frankova kopula	46
7.1	Data akcií	58
7.2	Porovnání výnosové křivky Intelu	59
7.3	GOF testy pro Intel	60
7.4	Matice Spearmanových korelačních koeficientů datového souboru	60
7.5	Chvosty datového souboru.	61
7.6	Porovnání původních a nově vygenerovaných dat	63
7.7	Histogramy četností defaultů	68
7.8	Závislost výnosu tranší na ρ_S	70
7.9	Porovnání výnosové křivky Microsoftu	74
7.10	GOF testy pro Microsoft	75
7.11	Porovnání výnosové křivky GE	75
7.12	GOF testy pro GE	76
7.13	Gaussova kopula: histogramy výnosu akcií	77
7.14	Gumbelova kopula: histogramy výnosu akcií	78
7.15	Studentova kopula: histogramy výnosu akcií	78

Poděkování

Na tomto místě bych velice ráda poděkovala doc. RNDr. Karlu Hronovi, Ph.D. za jeho vstřícnost, mé rodině za jejich trpělivost, mým přátelům za podporu a Veronice Římalové za její vyčerpávající korekturu.

Úvod

„Essentially, all models are wrong, but some are useful.“

(George E. P. Box)

Metoda kopul byla v posledních letech mezi světovými ekonomy hojně vyzdvihoována pro její praktické aplikace. Bohužel, díky nesprávně pochopeným předpokladům při jejím použití se brzy stala zatracovanou teorií. Ať již se kopulám připisovaly bájně vlastnosti, které naznačovaly řešení mnoha problémů v modelování rizika, anebo byly označovány za jednu z příčin hypotéční finanční krize z let 2008–2009, je nepopiratelné, že v řízení rizika hrály a budou hrát významnou roli. Zdá se, že měl George E. P. Box pravdu, když tvrdil, že jsou všechny modely špatné, ale některé jsou užitečné.

Problematika kopul úzce souvisí s mnohorozměrnými distribučními funkcemi. Je známo, že lze z každé simultánní distribuční funkce jednoznačně určit její jednorozměrné marginální distribuční funkce dle jednotlivých proměnných. Opačný proces, tedy určení sdružené distribuční funkce spojením marginálních, je bez znalosti jejich vzájemné závislosti nemožný. Pokud je však známa závislostní struktura, byť jedno či vícerozměrná, můžeme využít metody kopul k tvorbě daného modelu.

Předchozí problém zajišťuje jejich využití. Kopuly slouží jako nástroj, který dokáže rozdělit zvláště závislost proměnných a marginální distribuční funkce, anebo naopak tyto jednotlivé složky spojit. Oproti známým mnohorozměrným distribucím navíc lze zmírnit podmínky na marginální rozdělení. Například pokud máme vektor, jež má mnohorozměrné normální rozdělení, je potřebné, aby jeho jednotlivé složky byly také normálně rozdělené. To v případě kopul není nutné,

tedy jednotlivé marginální náhodné veličiny mohou mít rozdílné rozdělení.

Alternativně lze kopuly použít k separaci mnohorozměrných distribučních funkcí na jejich jednorozměrné distribuční funkce a závislost. Vzniklá závislostní struktura je poté spojena s jinými marginálními distribučními funkcemi. Tímto způsobem se může relativně rychle dojít k modifikaci známých rozložení pravděpodobností.

Objev kopul se připisuje americkému matematikovi Abe Sklarovi, který při výzkumu pravděpodobnostních metrických prostorů pozoroval tzv. t-normy, které jsou podobně jako kopuly zobrazením z kartézského součinu n uzavřených intervalů $\langle 0, 1 \rangle^n$ do $\langle 0, 1 \rangle$, kde $n = \{2, 3, 4, \dots\}$. Shodou okolností jsou některé t-normy kopulami a naopak. Název *kopula* Sklar poprvé použil roku 1959. Inspiroval se latinským slovem *copulare*, které v překladu znamená spojovat.

Největšího rozkvětu se kopuly dočkaly až o zhruba 40 let později, kdy byla navržena jejich aplikace ve finančním sektoru. V [12] je uvedeno, že se tak stalo mimo jiné díky rozvoji metodologie kvantitativního řízení rizika a nových finančních produktů. Významným nápadem, který kopuly prosazoval, byl Liho kreditní model portfolií popsany roku 2001 v [29]. Ten se stal základem pro modelování doby do defaultu významného finančního derivátu s názvem CDO (Collateralized Debt Obligation). Defaultem se zde rozumí pojem definovaný vyhláškou ČNB č. 123/2007: „Selhání dlužníka, ke kterému dochází v okamžiku, kdy je pravděpodobné, že nesplatí své závazky řádně a včas, aniž by věřitel přistoupil k uspokojení pohledávky ze zajištění nebo alespoň jedna splátka (jejíž výše je věřitelem považována za významnou) je po splatnosti déle než 90 dnů.“ CDO, a speciálně pak špatně modelovaná doba do defaultu, je právě označována za jednu z příčin tzv. hypoteční světové krize.

S konstrukcí závislosti pomocí kopul se lze také setkat při zjišťování VaR (Value at Risk), při testování stresu (Teorie extrémních hodnot), analýze citlivosti, oceňování finančních derivátů obchodovaných na burzách apod.

Proto je hlavním úkolem této práce seznámit čtenáře se základní teorií dvou a vícerozměrných kopul a jejich základními vlastnostmi.

V první části budou zmíněny základní pojmy, bez kterých se dané tématice nelze věnovat. Bude zde uvedeno jak značení konzistentní pro celou práci, tak definice, potřebné pro oživení základních znalostí z pravděpodobnosti a matematické statistiky.

V druhé části budou již rozebrány základní klíčové pojmy. Dále zde bude uvedena Sklarova věta, jež je pilířem celé teorie kopul a udává jejich praktické využití a na konec této kapitoly budou rozebrány podstatné vlastnosti kopul jako je omezenost, symetrie a chování na svých chvostech.

Část třetí je zasvěcena mírám závislosti, určujícím celkovou míru asociace mezi náhodnými proměnnými. Je zde popsán Pearsonův korelační koeficient, Kendallovo τ a Spearmanovo ρ_S .

Následující část se věnuje popisu vzniku a vlastností různých tříd kopul. Jsou zde podrobněji uvedeny fundamentální kopuly (součinná, horní a dolní), eliptické kopuly (Gaussova a Studentova) a archimédovské kopuly (Gumbelova, Claytonova a Frankova). Na závěr je uvedena tabulka s přehledem vlastností zástupců kopul ze všech těchto rodin.

V páté kapitole jsou popsány tři metody odhadu parametrů kopul založené na maximální věrohodnosti. První část je věnována úplné metodě maximální věrohodnosti, následuje popis inference marginálních distribucí a nakonec je popsána metoda pseudověrohodnosti.

Šestá kapitola ve stručnosti zmiňuje některé způsoby, jak testovat vhodnost kopul a jejich schopnost aproximace empirického datového souboru. Jsou zde uvedeny Cramer von Misesův a Anderson-Darlingův test.

Poslední kapitola využívá teoretických poznatků popsaných v předchozích kapitolách při konkrétních výpočtech. Je rozdělena na dvě části. První praktický příklad ukazuje aplikaci modelování kopul na reálných datech a druhá část popisuje modelování hodnocení CDO za využití kopul.

V celém textu budeme vycházet především z odborné literatury [22], [25] a [32].

Kapitola 1

Základní pojmy

Pro zavedení důležitých pojmů a definic je potřebné stanovit značení základních matematických prvků a operátorů, které bude konzistentní v celé práci. Autorka se snaží o zachování v literatuře co nejrozšířenějšího značení a jeho interpretace, nicméně zavedení je pro formu stále podstatné. Definice, které jsou v této sekci uvedeny, jsou převzaty ze skript pro základní kurs teorie pravděpodobnosti [21].

V následujícím textu značí \mathbb{N} množinu celých kladných čísel ($\mathbb{N} = \{1, 2, 3, \dots\}$), \mathbb{R} otevřenou množinu reálných čísel $(-\infty, \infty)$ a symbol $\overline{\mathbb{R}}$ celou reálnou osu včetně $-\infty$ a ∞ . \mathbb{R}^n představuje kartézský součin n množin reálných čísel $\mathbb{R} \times \mathbb{R} \times \dots \times \mathbb{R}$, kde $n \in \{2, 3, 4, \dots\}$. Symbol \mathbb{I} vymezuje uzavřený interval $\langle 0, 1 \rangle$.

Z teorie pravděpodobnosti jsou hojně využívány náhodné veličiny, označované X . Symboly dané značením \mathbf{x} představují vektory $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ nebo $\overline{\mathbb{R}}^n$ a \mathbf{X} představují vektory $\mathbf{X} = (X_1, X_2, \dots, X_n)$, jejichž složkami jsou náhodné veličiny. Dále je uvažována distribuční funkce náhodné veličiny v následujícím smyslu:

Definice 1.0.1. Nechť X je náhodná veličina definovaná na pravděpodobnostním prostoru (Ω, \mathcal{A}, P) ¹. Reálná funkce F_X definovaná na $\overline{\mathbb{R}}$ předpisem

$$F_X(x) = P(X \leq x), x \in \overline{\mathbb{R}},$$

se nazývá distribuční funkce náhodné veličiny X .

¹ Dle [21] Ω znamená množinu všech možných výsledků náhodného pokusu, \mathcal{A} je jevové pole náhodného pokusu a P značí operátor pravděpodobnosti

Definice 1.0.2. Necht' F je distribuční funkce náhodné veličiny X . Potom

$$F_X^{-1}(c) = \inf\{x \in \mathbb{R} \mid F_X(x) \geq c\}, \forall c \in \mathbb{I}, \text{ kde } \inf(\emptyset) = -\infty,$$

nazýváme pseudoinverzní (kvantilovou) funkcí k distribuční funkci F .

Definice 1.0.3. Necht' je dán pravděpodobnostní prostor (Ω, \mathcal{A}, P) a náhodné jevy $A, B \in \mathcal{A}, P(B) > 0$. Podmíněnou pravděpodobností jevu A za podmínky B nazveme funkci $P(\cdot \mid B)$ definovanou na \mathcal{A} předpisem

$$P(A \mid B) = \frac{P(A) \cap P(B)}{P(B)}.$$

Jelikož ve většině textu bude potřeba i vícerozměrný případ distribuční funkce uvádíme jej v další definici.

Definice 1.0.4. Necht' $\mathbf{X} = (X_1, \dots, X_n)$ je náhodný vektor definovaný na pravděpodobnostním prostoru (Ω, \mathcal{A}, P) . Distribuční funkce náhodného vektoru \mathbf{X} je reálná funkce $F_{\mathbf{X}}$ definovaná na \mathbb{R}^n vztahem

$$F_{\mathbf{X}}(x_1, x_2, \dots, x_n) = P(X_1 \leq x_1, \dots, X_n \leq x_n) = P(\mathbf{X} \leq \mathbf{x}),$$

kde $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$.

Pro propojení znalostí získaných z těchto definic a kopuly je potřebné uvést základní vlastnosti distribuční funkce.

Věta 1.0.1. *Distribuční funkce $F_{\mathbf{X}}(x_1, x_2, \dots, x_n)$ n -rozměrného náhodného vektoru \mathbf{X} má tyto vlastnosti:*

- $\lim_{\forall i \ x_i \rightarrow \infty} F_{\mathbf{X}}(x_1, x_2, \dots, x_n) = 1, \lim_{\exists i \ x_i \rightarrow -\infty} F_{\mathbf{X}}(x_1, x_2, \dots, x_n) = 0,$
kde $i = 1, 2, \dots, n.$
- $F_{\mathbf{X}}(x_1, x_2, \dots, x_n)$ je zprava spojitá v každé proměnné (při pevných hodnotách ostatních $n - 1$ proměnných).

- Pro všechna $a_i, b_i, -\infty < a_i \leq b_i < \infty, i = 1, 2, \dots, n$, platí

$$P(a_1 < X_1 \leq b_1, \dots, a_n < X_n \leq b_n) = \sum (-1)^{\sum \epsilon_j} F_{\mathbf{X}}(c_1, \dots, c_n) \geq 0,$$

kde $\epsilon_j = 0$ nebo $\epsilon_j = 1, (j = 1, \dots, n), c_j = a_j \epsilon_j + b_j(1 - \epsilon_j)$.

- $F_{\mathbf{X}}(x_1, x_2, \dots, x_n)$ je neklesající funkcí každé své proměnné (při pevně daných hodnotách ostatních $n - 1$ proměnných).

Poslední definice, kterou uvedeme v této sekci, popisuje vztah simultánní distribuční funkce a jejích marginálních rozdělení pravděpodobností. Platí, že pokud je známo rozdělení pravděpodobností náhodného vektoru \mathbf{X} , vždy lze jednoznačně určit rozdělení pravděpodobností jeho libovolného podvektoru.

Definice 1.0.5. Náhodný vektor $(X_{i_1}, \dots, X_{i_k})$ se nazývá marginální náhodný vektor příslušný k náhodnému vektoru \mathbf{X} .

Jeho distribuční funkci $F_{X_{i_1}, \dots, X_{i_k}}(x_{i_1}, \dots, x_{i_k})$ nazýváme marginální distribuční funkcí k distribuční funkci $F_{\mathbf{X}}(x_1, x_2, \dots, x_n)$ a obdobně rozdělení pravděpodobností marginálního náhodného vektoru se nazývá marginální rozdělení pravděpodobností příslušné k rozdělení $P_{\mathbf{X}}$.

Následující vztahy umožňují určení marginálních distribučních funkcí ze sdružené distribuční funkce původního vektoru. Speciálně je zde uveden případ dvou-rozměrného rozdělení pravděpodobností, kterým se budeme v této práci zabývat především.

Věta 1.0.2. Rozdělení pravděpodobností náhodného vektoru $\mathbf{X} = (X_1, \dots, X_n)$ s distribuční funkcí $F_{\mathbf{X}}(x_1, x_2, \dots, x_n)$ jednoznačně určuje rozdělení pravděpodobností jeho libovolného podvektoru $(X_{i_1}, \dots, X_{i_k}), k = 1, \dots, n - 1, 1 \leq i_1 < \dots < i_k \leq n$. Platí

$$F_{X_{i_1}, \dots, X_{i_k}}(x_{i_1}, \dots, x_{i_k}) = \lim_{x_j \rightarrow \infty, \forall j \neq i_1, \dots, i_k} F_{\mathbf{X}}(x_1, \dots, x_n).$$

Speciálně pro náhodný vektor (X, Y) jsou určeny jednorozměrné marginální distribuční funkce takto

$$F_X(x) = \lim_{y \rightarrow \infty} F(X, Y)(x, y), \forall x \in \overline{\mathbb{R}}, \quad F_Y(y) = \lim_{x \rightarrow \infty} F(X, Y)(x, y), \forall y \in \overline{\mathbb{R}}.$$

Kapitola 2

Kopuly

Jak bylo v úvodu naznačeno, kopula je funkcí se specifickými vlastnostmi, která každému n -úhelníku ležícímu v prostoru $\langle 0, 1 \rangle^n$, kde $\langle 0, 1 \rangle^n$ značí kartézský součin n intervalů $\langle 0, 1 \rangle$, přiřadí číslo z intervalu $\langle 0, 1 \rangle$. Jejím úkolem je spojení marginálních distribučních funkcí a závislosti tak, aby vznikla odpovídající simultánní distribuční funkce. Proto se tato kapitola zaměří na bližší popis kopul. Primárními zdroji, ze kterých bylo v této kapitole čerpáno jsou [25] a [32].

Názorně může být tento typ úlohy popsán pro $n = 2$ následovně. Necht' jsou dány spojitě náhodné veličiny X a Y s distribučními funkcemi $F(x) = P(X \leq x)$ a $G(y) = P(Y \leq y)$ a spojitá sdružená distribuční funkce $H(x, y) = P(X \leq x, Y \leq y)$. Potom kopula přiřadí každému bodu $(x, y) \in \overline{\mathbb{R}}^2$ bod v prostoru \mathbb{I}^3 daný souřadnicemi $(F(x), G(y), H(x, y))$.

Obecně lze tedy pro dvourozměrný případ kopulu definovat jako:

Definice 2.0.1. Dvourozměrná kopula (2-kopula) $C : \langle 0, 1 \rangle^2 \rightarrow \langle 0, 1 \rangle$ je sdružená distribuční funkce, jejíž jednorozměrné marginální proměnné mají rovnoměrné rozdělení.

Každá kopula je spojena s náhodným vektorem $\mathbf{X} = (X, Y)$, jehož jednotlivé náhodné veličiny jsou rovnoměrně rozděleny na intervalu \mathbb{I} , značeno $X \sim Ro(0, 1)$, $Y \sim Ro(0, 1)$, a rozdělení \mathbf{X} je dáno kopulou C . Opačně lze říci, že pro všechny náhodné vektory $\mathbf{X} = (X, Y)$, kde $X \sim Ro(0, 1)$, $Y \sim Ro(0, 1)$ existuje kopula C , podle které se řídí jejich rozdělení pravděpodobností.

Z definice 2.0.1 vyplývá, že vlastnosti, které jsou od dvourozměrné kopuly vyžadovány, jsou stejné jako ty, které musí splňovat distribuční funkce dvourozměrného pravděpodobnostního rozdělení. Proto lze pro kopulu použít alternativní definici.

Definice 2.0.2. Funkce $C : \mathbb{I}^2 \rightarrow \mathbb{I}$ je dvourozměrná kopula, jestliže

- $\forall u, v \in \mathbb{I}, C(u, 0) = 0 = C(0, v)$,
- $\forall u, v \in \mathbb{I}, C(u, 1) = u$ a $C(1, v) = v$,
- $\forall u_1, u_2, v_1, v_2 \in \langle 0, 1 \rangle, u_1 \leq u_2$ a $v_1 \leq v_2$
 $C(u_2, v_2) - C(u_1, v_2) - C(u_2, v_1) + C(u_1, v_1) \geq 0$.

První vlastnost, tzv. *uzemnění*, zajišťuje přidělení nulové pravděpodobnosti celé sdružené distribuční funkci v případě, že je jedno z marginálních rozdělení pravděpodobností rovno 0.

Druhá vlastnost znamená zachování původních marginálních distribučních funkcí. Pokud je $u = 1$, resp. $v = 1$ (tedy jedna hodnota je neměnná), pak $C(u, v) = v$, resp. $C(u, v) = u$, tedy hodnota kopuly bude vždy rovna hodnotě druhé (nezafixované) proměnné ve všech jejích bodech.

Poslední vlastnost, která je v literatuře také označována jako trojúhelníková nerovnost, zajišťuje, že kumulativní distribuční funkce je neklesající. Pokud tento vztah platí pro libovolný čtyřúhelník jehož vrcholy leží v $\langle 0, 1 \rangle^2$, funkce je tzv. 2-rostoucí (kvazi-monotónní). Je tedy zajištěna nezápornost každé plochy či objemu vypočítaném na $\langle 0, 1 \rangle \times \langle 0, 1 \rangle$. To, že je funkce C neklesající v každé své proměnné, neimplikuje její kvazi-monotónnost. Aby daná implikace platila, musí být přidána ještě podmínka na uzemnění.

Z definice 2.0.2 vyplývá, že kopuly jsou isotonické, tedy jestliže $\mathbf{x} \geq \mathbf{y}$, tzn. $x_i \geq y_i$ pro $i = 1, \dots, n$ a $\mathbf{x} = (x_1, \dots, x_n), \mathbf{y} = (y_1, \dots, y_n)$, kde $\mathbf{x}, \mathbf{y} \in \mathbb{I}^2$, pak platí $C(\mathbf{x}) \geq C(\mathbf{y})$.

Další vlastností, kterou 2-rostoucí funkce splňují, je Lipschitzova podmínka, jenž definuje stejnoměrnou spojitost kopul,

$$\forall w, x, y, z \in \mathbb{I} \text{ platí } |C(x, z) - C(w, y)| \geq |x - w| + |z - y|.$$

V reálném životě se však nelze omezit pouze na dvourozměrný případ. Proto je zde uvedena zobecněná definice pro n -rozměrný případ, kde $n \in \mathbb{N}, n \geq 2$.

Definice 2.0.3. n -rozměrná kopula (n -kopula) $C : \langle 0, 1 \rangle^n \rightarrow \langle 0, 1 \rangle$ je simultánní distribuční funkce, jejíž jednorozměrné marginální proměnné mají rovnoměrné rozdělení.

Podobně jako pro dvourozměrný případ, i zde lze popsat vlastnosti n -kopul vyplývající z definice 2.0.3. Je uvažován vektor $\mathbf{u} = (u_1, u_2, \dots, u_n)$, jehož jednotlivé složky mají rovnoměrné rozdělení pravděpodobností na intervalu \mathbb{I} .

- $C(\mathbf{u}) = (u_1, u_2, \dots, u_n)$ je neklesající v každé proměnné $u_i, i = 1, 2, \dots, n$.
- $C(u_1, u_2, \dots, u_n) = 0$, jestliže existuje alespoň jedno $i \in \{1, 2, \dots, n\}$ takové, že $u_i = 0$.
- $C(1, 1, \dots, 1, u_j, 1, \dots, 1) = u_j$, kde $u_k = 1, \forall k \neq j$
- $\forall (a_1, a_2, \dots, a_n), (b_1, b_2, \dots, b_n), a_i < b_i, i = 1, \dots, n :$

$$\sum_{i_1=1}^2 \dots \sum_{i_n=1}^2 (-1)^{\sum_{k=1}^n i_k} C(u_{1i_1}, \dots, u_{ni_n}) \geq 0, u_{j1} = a_j, u_{j2} = b_j.$$

V [37] je poukázáno na to, že každá funkce, která splňuje podmínky z definice 2.0.3, je kopula. Jestliže je dána n -rozměrná kopula, potom $C(1, u_1, \dots, u_{n-1})$ je opět kopula a to samé platí pro všechny k -rozměrné marginální distribuční funkce ($2 \leq k < n$). Z praktických důvodů bude ovšem v teoretické části kladen důraz především na dvourozměrné kopuly.

2.1. Sklarova věta

Pro praktické aplikace kopul je potřebné zavést postup, který dokáže jednoduše vyjádřit všechny mnohorozměrné distribuční funkce pomocí kopul a obdobně

je dokáže konstruovat z kopul a příslušných (různě rozdělených) marginálních proměnných. Tento vztah, bez kterého by byly kopuly pouhými funkcemi bez valného významu pro praxi, je uveden v následující větě.

Věta 2.1.1. (Sklarova) *Nechť F je dvourozměrná distribuční funkce s marginálními distribučními funkcemi H a G . Potom existuje kopula C taková, že*

$$\forall x, y \in \overline{\mathbb{R}}, F(x, y) = C(H(x), G(y)). \quad (2.1)$$

Jestliže jsou H a G spojité, potom je kopula C jednoznačně určena, v opačném případě je C jednoznačně určena pouze na kartézském součinu $D_C = D_G \times D_H$, kde D_G , resp. D_H , značí obor hodnot distribuční funkce náhodné veličiny G , resp. H .

Pokud H a G jsou jednorozměrné distribuční funkce a C je kopula, potom F ze vztahu (2.1) je dvourozměrná distribuční funkce s marginálními distribučními funkcemi H a G . Konstrukce vícerozměrného rozdělení je tedy rozdělena na určení jednorozměrných marginálních rozdělení pravděpodobnosti a na nalezení vhodné kopuly.

Je zajímavé, že ačkoliv byla Sklarova věta poprvé publikována roku 1959, její důkaz pro dvourozměrné rozdělení si na zveřejnění musel počkat až do r. 1983. Pokud čtenáře tento důkaz zajímá, je uveden v [32].

Ze Sklarovy věty také vyplývá následující vztah.

Věta 2.1.2. *Nechť H, G jsou jednorozměrné distribuční funkce a C je 2-kopula. Potom funkce $F : \overline{\mathbb{R}}^2 \rightarrow \langle 0, 1 \rangle$ daná Sklarovou větou je dvourozměrná distribuční funkce s marginálními distribučními funkcemi H a G .*

Snadno lze poznat, že tato věta je potřebná při konstrukci statistických modelů, při kterých je zadáno více jednorozměrných vlastností či chování (pro nás dané jako marginální distribuční funkce) subjektů (dané náhodnými vektory) a závislost těchto komponent, modelovaná známou kopulou.

Ke Sklarově větě lze navíc dodat, že pokud jsou H a G spojité distribuční funkce, pak kopulu C lze získat následujícím vztahem:

$$C(u_1, u_2) = F(H^{-1}(u_1), G^{-1}(u_2)).$$

Kopuly jsou tedy v podstatě způsob, jak transformovat původní náhodný vektor (X, Y) na takový $(U_1, U_2) = (H(X), G(Y))$, jehož marginální proměnné budou rovnoměrně rozděleny na uzavřeném intervalu $\langle 0, 1 \rangle$ a zároveň tato transformace zachová původní závislost mezi komponentami.

Díky této transformaci se ukazuje, že kopuly při tvorbě sdružených distribučních funkcí mohou kombinovat i jednorozměrné distribuční funkce náhodných veličin, které nejsou rovnoměrně rozděleny na $\langle 0, 1 \rangle$. Navíc lze modelovat i situace, kdy mají marginální proměnné vzájemně odlišné rozdělení pravděpodobností. V této charakteristice tkví jejich přínos pro statistiku, neboť jsou vhodnou alternativou k sestrojení mnohorozměrných rozdělení pravděpodobnosti (např. Gaussovo, Paretovo apod.). Nicméně, pro tuto standardizaci neexistuje žádný validní matematický důvod, ačkoliv je ze statistického hlediska užitečná.

V praxi je hojně využívána tzv. kopula přežití. Jedná se o kopulu, jež spojuje náhodné veličiny, reprezentující např. životnost objektů nebo délku života osob v nějaké populaci. Pravděpodobnost, že jedinec přežije do daného věku (času) x , je dána funkcí přežití \bar{H} danou vztahem:

$$\bar{H}(x) = P(X > x) = 1 - H(x), \quad (2.2)$$

kde H je distribuční funkcí náhodné veličiny X . Uvažovaný obor hodnot této funkce je zobecněn na celou rozšířenou reálnou osu (ačkoliv je obvykle pouze uvažován interval $\langle 0, \infty \rangle$).

Pro náhodný vektor (X, Y) se sdruženou distribuční funkcí F je sdružená funkce přežití dána vztahem

$$\bar{F}(x, y) = P(X > x, Y > y).$$

Jednorozměrné marginální distribuční funkce jsou dány vztahy

$$\bar{G}(y) = \bar{F}(-\infty, y) \text{ a } \bar{H}(x) = \bar{F}(x, -\infty).$$

Pro tuto situaci lze analogicky odvodit tzv. kopulu přežití. Nechť je dána kopula C náhodných veličin X a Y . Potom

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

Je-li dále definována funkce $\bar{C} : \langle 0, 1 \rangle^2 \rightarrow \langle 0, 1 \rangle$ předpisem

$$\bar{C}(u, v) = u + v - 1 + C(1 - u, 1 - v), \forall u, v \in \langle 0, 1 \rangle,$$

platí $\bar{H}(x, y) = \bar{C}(\bar{F}(x), \bar{G}(y))$. Kopula \bar{C} je tedy kopulou přežití náhodných veličin X a Y a spojuje sdruženou funkci přežití s jejími jednotlivými jednorozměrnými distribučními funkcemi danými také funkcemi přežití.

2.2. Fréchet-Hoeffdingovy meze

Ukazuje se, že prvním, kdo se k teorii kopul přiblížil, byl francouzský matematik Hoeffding (v 40. letech 19. století), který studoval různé míry mnohorozměrných distribučních funkcí, ovšem na intervalu $\langle -\frac{1}{2}, \frac{1}{2} \rangle$. I když se jeho výzkum netýkal přímo kopul, jeho poznatky jsou zde využívány. Dokonce jsou po něm pojmenovány tzv. Fréchet-Hoeffdingovy meze, kterým bude věnována následující část práce.

Věta 2.2.1. *Nechť jsou dány náhodné veličiny X a Y a kopula C , jež udává jejich závislost. Nechť $T_i : \mathbb{R} \rightarrow \mathbb{R}, i = 1, 2$ jsou ryze rostoucí funkce. Potom je struktura závislosti daná náhodnými veličinami*

$$T_1(X), T_2(Y)$$

také dána kopulou C .

To znamená, že ryze rostoucí transformace nemění danou závislost mezi náhodnými proměnnými.

Bylo dokázáno, že hodnoty kopul vždy leží v určitých mezích, tedy že jsou ohraničeny případy extrémních závislostí.

Nechť jsou například dány dvě rovnoměrně rozdělené náhodné veličiny X a Y . Pokud bude platit, že se jejich hodnoty vždy rovnají, tedy $X = Y$, pak jsou na sobě maximálně závislé. V tomto případě bude platit:

$$C(x, y) = P(X \leq x, Y \leq y) = \min(x, y).$$

Pokud je tento vztah dán do souvislosti s transformací náhodných veličin T danou větou 2.2.1, lze říci, že tato kopula vždy platí, když $Y = T(X)$. Naopak pokud jsou dány dvě nezávislé náhodné veličiny, platí vztah

$$C(x, y) = P(X \leq x, Y \leq y) = x \cdot y.$$

Nicméně, tato kopula, ačkoliv pro pozdější teorii fundamentální, netvoří druhou mez. Ta je dána maximální zápornou závislostí, tedy pokud platí $Y = 1 - X$.

Pro $1 - y < x$ platí

$$C(x, y) = P(X \leq x, 1 - X \leq y) = P(X \leq x, 1 - y \leq X) = x + y - 1$$

a 0 jinak.

Věta 2.2.2. Nechť C je kopula. Potom $\forall (u, v) \in D_C$, platí

$$W(u, v) = \max(u + v - 1, 0) \leq C(u, v) \leq \min(u, v) = M(u, v). \quad (2.3)$$

Důkaz. Pro dokázání vztahu (2.3) využijeme vlastností kopul uvedených v definici 2.0.2.

Nechť (u, v) je bodem patřícím do D_C . Dle druhé vlastnosti lze určit, že

$$C(u, v) \leq C(u, 1) = u \text{ a } C(u, v) \leq C(1, v) = v.$$

Z toho vyplývá, že

$$C(u, v) \leq \min(u, v).$$

Pro vyjádření druhého vztahu vycházíme z trojúhelníkové nerovnosti bodů $(u, 1)$ a $(1, v)$, které vždy patří do D_C . Dostaneme tak postupně nerovnosti

$$\begin{aligned}
C(1, 1) - C(u, 1) - C(1, v) + C(u, v) &\geq 0, \\
1 - u - v + C(u, v) &\geq 0, \\
C(u, v) &\geq u + v - 1.
\end{aligned}$$

Je-li navíc uvažováno $C(u, v) \geq 0$, snadno dokážeme, že vždy platí

$$C(u, v) \geq \max(u + v - 1, 0).$$

□

Mez $W(u, v)$, resp. $M(u, v)$ je také kopulou a je nazývána Fréchet-Hoeffdingovou dolní mezí, resp. Fréchet-Hoeffdingovou horní mezí.

V následujícím textu se budeme zabývat tvarem a vlastnostmi těchto fundamentálních kopul.

2.3. Vlastnosti kopul

Vzhledem k povaze kopul uvedených v definici 2.0.1 lze říci, že se jednotlivé typy neliší celkovým stupněm asociace. Zaměříme-li se však na průběh jejich distribučních funkcí, lze rozpoznat podstatné rozdíly v symetrii, důrazu kladeném na chvosty a míře asociace v různých místech distribuční funkce.

Tyto vlastnosti napomáhají nejen rozlišování jednotlivých kopul, navíc je lze do určité míry rozpoznat z empirického datového souboru a tím určit závislostní datovou strukturu, která je podstatná pro modelování sdružené distribuční funkce. V takovém případě již stačí pouze určit parametr daného typu kopuly. Navíc dle věty 2.2.1 charakteristiky kopul spojitých náhodných veličin přejímají její invarianci vůči striktně rostoucím transformacím.

První vlastností, kterou se budeme zabývat, je symetrie kopul. Stručně připomeneme symetrii jednorozměrné distribuční funkce, která bude poté použita při definování dvourozměrných případů a kopul.

V této kapitole budeme převážně vycházet z [25] a [32].

2.3.1. Symetrie kopul

Symetrie distribuční funkce náhodné veličiny je ve statistice velice oblíbenou charakteristikou. Díky ní lze zjistit, zda jsou hodnoty náhodné veličiny kolem daného čísla $a \in \mathbb{R}$ symetricky rozděleny, nebo zda je příslušná distribuční funkce vychýlená/šikmá, a které hodnoty zvýhodňuje.

Náhodná veličina X je symetrická kolem bodu $a \in \mathbb{R}$, jestliže $\forall x \in \mathbb{R}$ platí následující rovnost:

$$P(X - a \leq x) = P(a - X \leq x),$$

tedy distribuční funkce náhodných veličin $X - a$ a $a - X$ jsou totožné.

Pokud je navíc X spojitá náhodná veličina s distribuční funkcí F , pak lze předchozí vztah přepsat pomocí funkce přežití jako

$$F(a + x) = 1 - F(a - x) = \bar{F}(a - x).$$

Dle předchozí definice charakteristiky symetrie lze určit, zda je náhodný vektor (X, Y) , jenž je tvořen párem náhodných veličin X a Y , symetrický kolem bodu $(a, b) \in \mathbb{R}^2$. Jsou rozlišitelné různé druhy této charakteristiky.

Náhodný vektor (X, Y) je *marginálně symetrický* kolem bodu (a, b) , jsou-li symetrické obě jeho marginální distribuční funkce X , resp. Y , kolem bodu a , resp. b .

Dalším typem je *radiální symetrie*. Za stejných předpokladů lze říci, že náhodný vektor (X, Y) je radiálně symetrický kolem bodu (a, b) jestliže jsou jednorozměrné distribuční funkce náhodných veličin $X - a$, resp. $Y - b$, a $a - X$, resp. $b - X$, identické.

Pokud jsou F a G spojitě marginální distribuční funkce náhodných proměnných X a Y , lze tuto podmínku vyjádřit prostřednictvím jejich sdružené distribuční funkce H . Náhodný vektor (X, Y) je *radiálně symetrický*, právě tehdy, když

$$H(a + x, b + y) = \bar{H}(a - x, b - y), \forall (x, y) \in \mathbb{R}^2.$$

Poslední typ symetrie dvourozměrného náhodného vektoru (X, Y) nastane právě tehdy, když následující kombinace náhodných veličin $(X - a, Y - b)$, $(X -$

$a, b - Y), (a - X, Y - b), (a - X, b - Y)$ mají identickou sdruženou distribuci. Řekneme, že (X, Y) je *sdruženě symetrický* kolem bodu (a, b) .

Je snadné určit vztahy mezi jednotlivými druhy symetrie. Nejsilnější charakteristikou je sdružená symetrie, která implikuje její jak radiální, tak marginální obdobu. Podobně vyplývá z definic prvních dvou typů symetrií, že pokud je distribuční funkce radiálně symetrická podle daného bodu, pak bude symetrická i marginálně. Opačně to však neplatí, a proto marginální symetrii lze nazvat jako nejslabší.

Uvažujeme-li dvě spojitě náhodné veličiny X, Y symetrické kolem bodů a a b , jejichž sdruženou distribuční funkcí je H , jejich marginální distribuční funkce označujeme F a G a závislostní struktura, kopula, mezi těmito marginálními proměnnými je dána C , pak lze určit, kdy je náhodný vektor (X, Y) radiálně symetrický kolem bodu (a, b) . Stane se tak v případě, že platí jedna z následujících podmínek:

- $H(a + x, b + y) = \overline{H}(a - x, b - y)$,
- $C = \overline{C}$, kde \overline{C} je kopula přežití,
- $C(u, v) = u + v - 1 + C(1 - u, 1 - v), \forall (u, v) \in \mathbb{I}^2$.

Poslední typ symetrie uvedený v této publikaci má souvislost se symetrií samotných kopul. Mějme dány náhodné veličiny X a Y se sdruženou distribuční funkcí H . V případě, že jsou náhodné vektory (X, Y) a (Y, X) stejně rozdělené, nazýváme původní náhodné veličiny *zaměnitelné*, tedy F a G jsou stejně rozdělené. Symbolicky:

$$H(x, y) = H(y, x), \forall (x, y) \in \mathbb{R}^2.$$

Jsou-li dány spojitě náhodné veličiny X a Y s marginálními distribučními funkcemi F a G , sdruženou distribuční funkcí H a kopulou C , pak je lze nazvat zaměnitelnými právě tehdy, když $F = G$ a $C(u, v) = C(v, u), \forall (u, v) \in \mathbb{I}^2$.

To znamená, že kopulu C nazveme symetrickou tehdy, když platí $C(v, u) = C(u, v), \forall (u, v) \in \mathbb{I}^2$. Zaměnitelnost náhodných veličin a symetrie jejich kopuly tedy odpovídají téže vlastnosti.

2.4. Chvosty kopul

V reálném životě se lze setkat s náhodnými veličinami nabývajících nekonečně mnoha hodnot, jejichž realizace v krajích rozdělení nelze zanedbat např. málo pravděpodobná tržní hodnota dluhopisů, nastání přírodní katastrofy apod.

Na druhou stranu pokud je při modelování pravděpodobnost výskytu těchto jevů zanedbána, může hrozit riziko ztráty podstatné informace, zvláště pak v případě, kdy lze z daných pozorování vytvořit distribuční funkci s tzv. těžkými chvosty. Je-li navíc uvažováno více náhodných veličin, je možné, že výskyt extrémních hodnot u jedné indikuje vyšší pravděpodobnost, že další náhodná veličina také nabude extrémních hodnot. Teorie extrémních hodnot se danému příkladu věnuje velice podrobně a na jejím základě byly vyvinuty tzv. extrémní kopuly.

Chvosty distribučních funkcí tedy popisují chování svých distribucí (resp. chování příslušných náhodných veličin) v těch svých částech, které jsou vzdálené od očekávané hodnoty. Distribucemi s těžkými chvosty jsou pak označovány ty, které mají vyšší pravděpodobnost výskytu neobvyklých jevů než exponenciální rozdělení viz [1].

K jejich charakteristice je možné využít různých způsobů např. analýzy monotonie či pro vícerozměrný případ závislosti chvostů.

2.4.1. Monotonie chvostů

Náhodné veličiny X a Y jsou pozitivně kvadrantově závislé, jestliže $\forall(x, y) \in \mathbb{R}^2$ platí, že

$$P(X \leq x, Y \leq y) \geq P(X \leq x) \cdot P(Y \leq y)$$

a negativně kvadrantově závislé, když

$$P(X \leq x, Y \leq y) \leq P(X \leq x) \cdot P(Y \leq y).$$

Tyto závislosti určují, zda je pravděpodobnost, že obě náhodné veličiny nabudou zároveň extrémních (vysokých, resp. nízkých) hodnot vyšší, než kdyby tak učinily nezávisle na sobě.

Mají-li tyto náhodné veličiny sdruženou funkci H , spojitě marginální distribuční funkce F a G a kopulu C , kopula i sdružená distribuční funkce tuto vlastnost přejímají. Pozitivní kvadrantovou závislost lze tedy přepsat na

$$H(x, y) \geq F(x) \cdot G(y), \forall (x, y) \in \mathbb{R}^2 \text{ a } C(u, v) = u \cdot v, \forall (u, v) \in \mathbb{R}^2.$$

Pozitivní kvadrantovou závislost lze vyjádřit i pomocí podmíněné pravděpodobnosti:

$$P(Y \leq y | X \leq x) \geq P(Y \leq y) \text{ nebo } P(Y \leq y | X \leq x) \geq P(Y \leq y | X \leq \infty).$$

Z tohoto vztahu lze odvodit monotonii chvostů. Pro náhodnou veličinu Y lze odhadnout, kdy budou hodnoty v jejím levém chvostu klesat, resp. v jejím pravém chvostu růst.

Náhodná veličina Y je klesající ve svém levém chvostu na X , jestliže je podmíněná pravděpodobnost $P(Y \leq y | X \leq x)$ nerostoucí funkcí x pro každý bod $y \in \mathbb{R}$.

Náhodná veličina Y je rostoucí ve svém pravém chvostu na X , jestliže je podmíněná pravděpodobnost $P(Y > y | X > x)$ neklesající funkcí x pro každý bod $y \in \mathbb{R}$.

Obdobně lze definovat pro náhodnou veličinu X .

Pro negativní kvadrantovou závislost existují náhodné veličiny Y rostoucí ve svém levém chvostu, resp. klesající ve svém pravém chvostu vzhledem k X , kdy je daná podmíněná pravděpodobnost nerostoucí, resp. neklesající funkcí x pro každý bod $y \in \mathbb{R}$.

Každá z podmínek monotonosti tedy implikuje pozitivní kvadrantovou závislost.

Monotonii chvostů lze analyzovat i v případě kopul. V [32] je dokázáno, že pokud jsou zadány spojitě náhodné veličiny s kopulou C , pak je náhodná veličina Y klesající ve svém levém chvostu podle X právě tehdy, když $\forall v \in \mathbb{I}$ je $\frac{C(u,v)}{u}$ nerostoucí na u , a rostoucí ve svém pravém chvostu právě tehdy, když $\forall v \in \mathbb{I}$ platí, že $\frac{v-C(u,v)}{1-u}$ je neklesající na u .

Monotonii chvostů využijeme při zjednodušování měr závislostí, viz třetí kapitola.

2.4.2. Závislost chvostů

Další vlastností chvostů je míra síly závislosti mezi extrémními výskyty náhodných veličin. Tato vlastnost je užitečná především při analýze a tvorbě závislosti extrémních hodnot.

Uvedenou vlastnost lze kvantifikovat koeficientem závislosti chvostů, který nabývá hodnot z intervalu $\langle 0, 1 \rangle$ a je invariantní vůči ryze rostoucím transformacím. Je značen symbolem λ . V následujícím textu bude primárně zmíněn dvojrozměrný případ, který je však snadno zobecnitelný pro n -rozměrnou kopulu.

Nechť existuje dvojrozměrná kopula přežití \overline{C} a koeficient

$$\lambda_U = \lim_{u \rightarrow 1^-} \frac{\overline{C}(u, u)}{(1 - u)}$$

nechť nabývá hodnot z intervalu $(0, 1)$, pak C je závislá na svém horním chvostu. V opačném případě (tedy pro $\lambda_U = 0$) je na daném chvostu nezávislá nebo nemonotonní.

Obdobně lze definovat závislost na dolním chvostu kopuly. Jestliže existuje kopula C taková, že koeficient

$$\lambda_L = \lim_{u \rightarrow 0^+} \frac{C(u, u)}{(u)} \in (0, 1),$$

pak C je závislá ve své dolním chvostu.

Předchozí definice lze odvodit prostřednictvím podmíněných pravděpodobností. Nechť jsou dány náhodné veličiny X a Y spojené kopulou C . Pak

$$\lambda_U = \lim_{u \rightarrow 1^-} P(X > u | Y > u) = \lim_{u \rightarrow 1^-} P(Y > u | X > u).$$

Vzhledem k tomu, že v definicích figurují limity, lze říci, že empirická obdoba závislosti chvostů pro data neexistuje. Tento problém lze vyřešit metodami popsanými v [9], [10] a [15], anebo lze analyzovat jednotlivé páry náhodných veličin vizualizací.

2.4.3. Asymetrie chvostů

Pokud se horní a dolní chvost kopuly chovají odlišně, např. vysokým kladným hodnotám pozorování je přisuzována vyšší pravděpodobnost, tedy $\lambda_U \leq \lambda_L$, pak jsou chvosty asymetrické. Opět se jedná o podstatnou charakteristiku, která určuje typ závislostní struktury mezi náhodnými veličinami. V [25] je definována centrální symetrie/asymetrie kopul následovně:

Definice 2.4.1. Necht' jsou dány dvě rovnoměrně rozdělené náhodné veličiny X , Y a jejich kopula C . Kopulu \widehat{C} , která je dána rovnoměrně rozdělenými náhodnými veličinami $1 - X$ a $1 - Y$ nazveme zrcadlovou kopulou k C . Pak $C = \widehat{C}$ jsou radiálně (centrálně) symetrické podle bodu $(0, 0)$.

Pokud je kopula radiálně symetrická, pak jsou symetrické i její chvosty.

Jestliže platí

$$C(u, v) \leq \widehat{C}(u, v), \forall u, v \in (0, z_0), \text{ kde } z_0 \in (0, \frac{1}{2}),$$

pak má kopula C těžší dolní chvost, což znamená vysokou pravděpodobnost společného výskytu malých hodnot, a lze říci, že je reflektivně asymetrická se šikmostí v dolním chvostu. V opačném případě, tedy

$$C(u, v) \geq \widehat{C}(u, v), \forall u, v \in (0, z_0), \text{ kde } z_0 \in (0, \frac{1}{2}),$$

jsou pravděpodobnější vysoké hodnoty a kopula je reflektivně asymetrická se šikmostí v horním chvostu.

Kapitola 3

Závislost

Kopuly jsou objekty zachycující celkové rozdělení závislosti mezi dvěma a více náhodnými proměnnými, které je nezávislé na rozdělení náhodných veličin jejich jednorozměrných distribučních funkcí. Pro odlišně rozdělené marginální náhodné veličiny se tedy závislost vyjádřená daným typem kopuly nezmění.

V teorii pravděpodobnosti, potažmo ve statistice, je běžně využíván k měření závislosti či nezávislosti dvou proměnných tzv. Pearsonův korelační koeficient, který udává míru lineární závislosti mezi dvěma náhodnými veličinami. Alternativně lze definovat dvě míry asociace, vyjadřující závislost dvou a více náhodných proměnných pomocí konkordance. Je to Kendallovo τ a Spearmanovo ρ_S . Tyto míry jsou, na rozdíl od Pearsonova korelačního koeficientu, závislé pouze na kopule, nikoliv na marginálních distribučních funkcích, a tudíž jsou neměnné v rámci daného typu kopuly.

Vzhledem ke skalární povaze těchto měř nelze jejich prostřednictvím rozlišit, jak se mění závislost v různých částech distribučních funkce.

Cílem této kapitoly je přiblížit dané míry závislosti a ukázat, jaký vztah mají ke kopulám a mezi sebou. V celé kapitole jsme vycházeli z [22], [25], [32].

3.1. Pearsonův korelační koeficient

Obyčejně se pod pojmem korelace skrývá tzv. Pearsonův lineární koeficient, který je sice nejvyužívanější mírou závislosti, avšak měří pouze sílu lineárního

vztahu mezi náhodnými veličinami. Mají-li náhodné veličiny distribuční funkci, indukující jiný typ závislosti, tato míra nedává relevantní výsledky.

Jsou-li dány dvě náhodné veličiny X, Y s konečnými rozptyly, pak základní vztah pro korelační koeficient ρ je

$$\rho(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{\text{var}(X) \text{var}(Y)}},$$

kde $\text{cov}(X, Y) = E(XY) - E(X)E(Y)$. Hodnota ρ nabývá hodnot z intervalu $\langle -1, 1 \rangle$, přičemž 1 značí úplnou kladnou a -1 úplnou zápornou lineární závislost. Pokud jsou náhodné veličiny X a Y nezávislé, ρ je rovno 0. Opačně však toto tvrzení neplatí, neboť v případě, že je ρ rovné 0, mohou být dané proměnné závislé jiným způsobem, např. kvadraticky.

Dalším problémem, se kterým se při výpočtu korelačního koeficientu lze setkat, je podmínka konečnosti druhých momentů náhodných veličin. Závažnost porušení tohoto předpokladu je významná, zvláště v případě distribučních funkcí s těžkými chvosty (viz podkapitola 2.4). Obor hodnot distribuční funkce je totiž v takovém případě neomezený a rozptyl náhodné veličiny je nekonečný.

Poslední nevýhodou Pearsonova korelačního lineárního koeficientu, na kterou bude poukázáno, je jeho závislost na marginálních distribučních funkcích náhodných veličin. Z tohoto důvodu není invariantní při nelineárních ryze rostoucích transformacích $T : \mathbb{R} \rightarrow \mathbb{R}$:

$$\rho(X, Y) \neq \rho(T(X), T(Y)).$$

Hodnota ρ bude tudíž rozdílná v případě, že je dána stejná kopula, avšak jiné marginální náhodné proměnné. Tuto vlastnost další míry závislosti nemají.

Nechť $(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)$ jsou realizace náhodného výběru o rozsahu N z rozdělení náhodného vektoru (X, Y) . Výběrová podoba Pearsonova korelačního koeficientu je dána vztahem

$$r(X, Y) = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2 \sum_{i=1}^N (y_i - \bar{y})^2}}, \text{ kde } \bar{x} = \frac{\sum_{i=1}^N x_i}{N} \text{ a } \bar{y} = \frac{\sum_{i=1}^N y_i}{N}.$$

3.2. Konkordance

Doslovným překladem slova *concordance* je soulad, shoda. Pokud jsou dány dvě náhodné veličiny a vysoké hodnoty jedné jsou spojeny s vysokými hodnotami druhé, pak je lze nazvat konkordantní. Pokud platí pravý opak, tedy vysoké hodnoty jedné jsou spojeny s nízkými hodnotami druhé, nazýváme tento pár náhodných veličiny diskonkordantní.

Formálně lze tento vztah zapsat následovně: Nechť (X, Y) je vektor spojitých náhodných veličin. Jestliže platí $(x_i, y_i)(x_j, y_j) > 0$ pro dvojici realizací $(x_i, y_i), (x_j, y_j)$ z oboru hodnot vektoru (X, Y) , pak jsou náhodné veličiny X a Y konkordantní.

Pokud analogicky pro realizace z oboru hodnot (X, Y) platí $(x_i, y_i)(x_j, y_j) < 0$, X a Y jsou diskonkordantní. Alternativně lze spojité náhodné veličiny X a Y nazvat konkordantní, resp. diskonkordantní, jestliže pro realizace z oboru hodnot (X, Y) platí

$$\forall (x_i, y_i), (x_j, y_j) : x_i < x_j \wedge y_i < y_j, \text{ resp. } \forall (x_i, y_i), (x_j, y_j) : x_i < x_j \wedge y_i > y_j.$$

3.3. Kendalovo τ

První uvedenou mírou asociace, která využívá konkordance, je Kendalovo τ . Na rozdíl od Pearsonova korelačního koeficientu je tato míra invariantní vůči rostoucím transformacím.

Nechť $(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)$ je realizace náhodného výběru o rozsahu N z rozdělení náhodného vektoru (X, Y) , kde X, Y jsou spojité náhodné veličiny. Každý pár $(x_i, y_i), (x_j, y_j)$ z tohoto výběru je konkordantní, nebo diskonkordantní. Celkově je k porovnání $\binom{N}{2}$ párů, přičemž číslo $c \in \{0, \dots, N\}$ určuje počet konkordantních a číslo $d \in \{0, \dots, N\}$ diskonkordantních párů.

Kendalovo τ ve své výběrové podobě je dáno vztahem

$$t = (c - d)/(c + d) = (c - d)/\binom{N}{2}.$$

Výše uvedené výběrové verzi odpovídá těž teoretické Kendalovo τ . Nechť je dán náhodný vektor spojitých náhodných veličin (X, Y) se spojitou distribuční

funkcí H . Pak τ je definováno jako pravděpodobnost konkordance minus pravděpodobnost diskordance,

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

Konkordantní funkce Q je definována jako rozdíl mezi pravděpodobnostmi dvou konkordantních (diskordantních) spojitých náhodných vektorů (X_1, Y_1) a (X_2, Y_2) s příslušnými (odlišnými) sdruženými distribucemi H_1 a H_2 , ale se stejnými marginálními distribučními funkcemi F a G . Tedy platí vztahy $H_1(x, y) = C_1(F(x), G(x))$ a $H_2(x, y) = C_2(F(x), G(x))$. Konkordantní funkce závisí jen na distribucích jejich kopul,

$$Q = Q(C_1, C_2) = 4 \int_0^1 \int_0^1 C_2(u, v) dC_1(u, v) - 1.$$

Díky předchozímu vztahu lze teoretickou (populační) verzi Kendallova τ přepsat následovně. Nechť X a Y jsou spojitě náhodné veličiny a C jejich kopula. Pak platí

$$\tau_{X,Y} = \tau_C = Q(C, C) = 4 \int_0^1 \int_0^1 C(u, v) dC(u, v) - 1 = 4E(C(U, V)) - 1,$$

kde $U, V \sim Ro(0, 1)$.

3.4. Spearmanovo ρ_S

Spearmanovo ρ_S neboli pořadový korelační koeficient je mírou asociace, která je podobně jako Kendallovo τ založena na porovnávání konkordantních a diskordantních párů pozorování náhodných proměnných a na rozdíl od Pearsonova koeficientu závisí pouze na dané kopule (tedy nezávisí na rozdělení marginálních náhodných veličin). Je pro něj užíváno značení ρ_S . Pro aplikaci Spearmanova koeficientu na normálních datech budou uvedeny jeho populační varianty.

Máme dány tři nezávislé náhodné vektory $(X_1, Y_1), (X_2, Y_2), (X_3, Y_3)$ se stejnou simultánní distribuční funkcí H (jejíž marginální protějšky jsou značeny F a G) a kopulou C .

Koeficient ρ_S je proporciální k rozdílu pravděpodobností konkordance a diskordance dvou vektorů (X_1, Y_1) a (X_2, Y_3) . Jedná se o pár náhodných vektorů, které mají stejná marginální rozdělení, ale první vektor má sdruženou distribuční funkci $H = C(F(x), G(y))$, zatímco komponenty druhého jsou nezávislé $H = F(x)G(y)$. Přitom C je tzv. součinnová kopula (viz kapitola 4). Pořadový korelační koeficient je pak definován jako

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

Obdobně lze vypočítat Spearmanův koeficient pro náhodné vektory (X_1, Y_1) a (X_3, Y_2) .

Nechť X a Y jsou spojité náhodné veličiny s kopulou C . Potom lze pořadový koeficient $\rho_{X,Y}$ vyjádřit vztahy:

$$\begin{aligned} \rho_S &= 3Q(C, \Pi) = 12 \int_0^1 \int_0^1 uv \, dC(u, v) - 3 = 12 \int_0^1 \int_0^1 C(u, v) \, du \, dv - 3 = \\ &= 12 \int F(x)G(y) \, dH(x, y) - 3. \end{aligned}$$

Nechť je dán soubor N pozorování $(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)$, realizací výběru z rozdělení náhodného vektoru (X, Y) a R_i , resp. S_i , jsou $\forall i = 1, 2, \dots, N$ pořadí pozorování X , resp. Y . Platí $R_i = k$ právě tehdy, když je x_i k -tým nejmenším pozorováním mezi x_1, x_2, \dots, x_N a $S_i = l$, právě tehdy, když je y_i l -tým nejmenším pozorováním mezi y_1, y_2, \dots, y_N . Výběrový tvar Spearmanovy pořadové korelace r_S je poté dán vztahem

$$\begin{aligned} r_S &= \frac{\sum_{i=1}^N R_i S_i - N[(N+1)/2]^2}{\sum_{i=1}^N [i - (N+1)/2]^2} = \frac{\sum_{i=1}^N R_i S_i - N[(N+1)/2]^2}{\sum_{i=1}^N [N(N^2 - 1)/12]} = \\ &= 12 \frac{\sum_{i=1}^N (R_i - \bar{R})(S_i - \bar{S})}{N(N^2 - 1)} = 1 - 6 \frac{\sum_{i=1}^N (R_i - S_i)^2}{N(N^2 - 1)}, \end{aligned}$$

kde \bar{R} , resp. \bar{S} je výběrový průměr vektorů pořadí R_1, \dots, R_N , resp. S_1, \dots, S_N .

Kapitola 4

Typy kopul

Následující kapitola bude zaměřena na konkrétními rodiny kopul a jejich vybrané zástupce. Každý datový soubor má svá specifika a ať už je charakterizován symetrickým rozložením nebo výskytem těžkých chvostů, je potřebné, aby vybraná kopula vystihla jeho závislostní strukturu co nejlépe. Proto byly vyvinuty různé typy kopul.

Kopuly je zároveň možné konstruovat různými způsoby. Na tomto místě pouze uvedme, že je možné využít tzv. inverzní metodu, vyplývající ze Sklarovy věty, geometrické metody, odvozené z požadavků na tvar výsledné kopuly, algebraické metody, umělé vytvoření kopul s požadovanými vlastnostmi či spojení již existujících kopul. Vzhledem k rozsahu této práce pro podrobnosti odkazujeme na [32].

Pro uživatelsky příjemné modelování kopul je podstatné, aby splňovaly alespoň jednu z následujících podmínek: buď lze jejich výpočetní vzorec vyjádřit uzavřenou formou nebo je jejich závislostní struktura snadno odvoditelná ze sdružené distribuční funkce.

Jako příklad kopul splňujících první podmínku jsou uvedeny tzv. archimédovské kopuly, které vznikají pomocí speciálních funkcí, tzv. generátorů. Kopuly splňující druhou podmínku jsou představeny prostřednictvím tzv. eliptických kopul. Dále jsou připomenuty fundamentální kopuly, které jsou potřebné při ohraničování ostatních kopul a pro vyjádření nezávislosti náhodných veličin.

V následujícím textu je využito vztahu hustoty a sdružené distribuční funkce kopuly n -rozměrného náhodného vektoru \mathbf{U} ,

$$C(\mathbf{u}) = P(\mathbf{U} \leq \mathbf{u}) = \int_{-\infty}^{u_1} \cdots \int_{-\infty}^{u_n} c(s_1, \dots, s_n) ds_1 \dots ds_n,$$

kde $c(s_1, \dots, s_n)$ je tzv. hustota kopuly a $\mathbf{u} \in \mathbb{R}^n$. Hlavními zdroji, ze kterých bylo čerpáno jsou [14], [25], [26], [32], [35], [40], [41].

4.1. Fundamentální kopuly

S fundamentálními kopulami jsme se již setkali v definici 2.2.2. V případě horní a dolní kopuly se jedná o základní kopuly, které tvoří pomyslné hranice rozdělení všech ostatních kopul. Součinnová kopula pak představuje kopulu nezávislosti mezi náhodnými proměnnými.

V následujícím textu jsou určeny vlastnosti fundamentálních kopul a je ukázáno, že jsou součástí různých tříd kopul.

4.1.1. Součinnová kopula

Nezávislá neboli součinnová kopula udává vztah mezi dvěma či více nezávislými náhodnými veličinami. Lze ji vyjádřit vztahem

$$C_{\perp}(u_1, u_2, \dots, u_n) = \prod_{i=1}^n u_i.$$

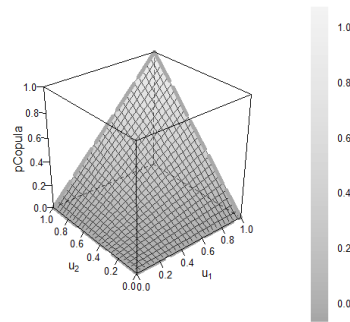
Její sdružená distribuční funkce je znázorněna na obrázku 4.1

4.1.2. Horní kopula

Horní kopula udává tzv. Fréchet-Hoeffdingovu horní mez. Určuje tedy maximální možnou závislost mezi dvěma proměnnými. Pro $n \geq 2$ je dána jako

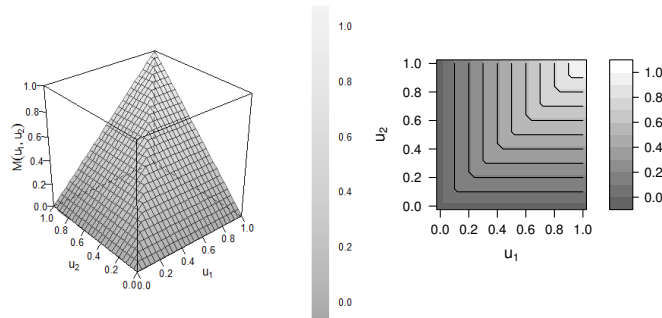
$$C(u_1, u_2, \dots, u_n) = \min(u_1, u_2, \dots, u_n).$$

Lze se také setkat s označením $M(u_1, u_2, \dots, u_n)$. Její sdružená distribuční funkce je uvedena na obrázku 4.2. Pro znázornění kopul budou využívány i vrstevnicové



Obrázek 4.1: Sdružená distribuční funkce součinné kopuly.

grafy, přičemž vrstevnice jsou dány množinami $V = \{(u, v) \in \mathbb{I}^2 \mid C(u, v) = t\}$, kde $t \in \langle 0, 1 \rangle$.



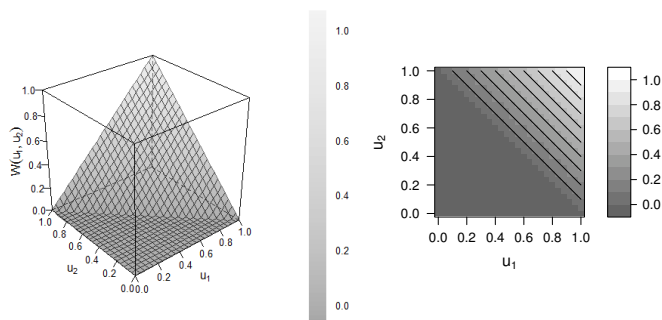
Obrázek 4.2: Sdružená distribuční funkce horní kopuly (vlevo) a její vrstevnice (vpravo)

4.1.3. Dolní kopula

Dolní kopula je tzv. Fréchet-Hoeffdingovou dolní mezí. Určuje tedy maximální možnou zápornou závislost mezi dvěma proměnnými. Pro $n \geq 2$ je dána vztahem

$$C(u_1, u_2, \dots, u_n) = \max \left\{ 1 - n + \sum_{i=1}^n u_i, 0 \right\}.$$

Dolní kopula je značena $W(u_1, u_2, \dots, u_n)$. Její sdružená distribuční funkce je graficky znázorněna na obrázku [4.3](#)



Obrázek 4.3: Sdružená distribuční funkce dolní kopuly (vlevo) a její vrstevnice (vpravo)

4.2. Eliptické kopuly

Tento typ kopul byl vytvořen přímo z třídy eliptických distribucí využitím Sklarovy věty. Eliptické neboli implicitní kopuly modelují závislostní strukturu eliptické mnohorozměrné distribuční funkce, aniž by byla zachována původní rozdělení jejích marginálních proměnných.

Výsledná kopula tedy přejímá podstatné vlastnosti dané distribuční funkce a navíc je možné flexibilně modelovat rozdělení marginálních náhodných veličin. Původní eliptická distribuční funkce má obecně parametry dané vektorem středních hodnot $\boldsymbol{\mu}$ a varianční maticí $\boldsymbol{\Sigma}$. Dále uvažujeme pouze normované eliptické distribuční funkce, a proto je $\boldsymbol{\mu}$ nulovým vektorem a matice $\boldsymbol{\Sigma}$ je shodná s korelační maticí \mathbf{R} .

Tyto kopuly nemají uzavřenou formu, nicméně jejich implicitní závislostní strukturu lze zjistit dle následujícího předpisu:

$$C(u_1, u_2, \dots, u_n, \mathbf{R}) = F\{F_1^{-1}(u_1), F_2^{-1}(u_2), \dots, F_n^{-1}(u_n), \mathbf{R}\},$$

kde $0 < u_1, u_2, \dots, u_n < 1$, a \mathbf{R} je $n \times n$ pozitivně definitní matice.

Pro implicitní kopuly spojitých simultánních distribučních funkcí F s ryze spojitými marginálními distribučními funkcemi F_1, F_2, \dots, F_n je hustota kopuly

dána vztahem

$$c(\mathbf{u}) = \frac{f(F_1^{-1}(u_1), \dots, F_n^{-1}(u_n))}{f_1(F_1^{-1}(u_1)) \cdot \dots \cdot f_n(F_n^{-1}(u_n))},$$

kde f je hustota sdružené funkce, f_1, \dots, f_d jsou marginální hustoty a $F_1^{-1}, \dots, F_n^{-1}$ jsou inverzní funkce k marginálním distribučním funkcím.

4.2.1. Gaussova kopula

První uvedenou eliptickou kopulou je kopula Gaussova, jež zachycuje závislostní strukturu mnohorozměrné normální distribuční funkce. Z daného rozdělení přijímá vlastnosti symetrie a nezávislosti chvostů. Tato rodina kopul je nejpožívanějším typem kopul. Používá se například v analýze rizika při simulaci distribuční funkce více rizikových faktorů (jednotlivé finanční toky) ovlivňujících výnos z akcií.

Nechť Φ značí distribuční funkci náhodného vektoru (X, Y) , majícího normované normální rozdělení s korelačním koeficientem ρ , a nechtě Φ_X , resp. Φ_Y , je symbol distribuční funkce jednorozměrné normované normální náhodné veličiny X , resp. Y . Potom kopula C_ρ , znázorňující závislostní strukturu distribuční funkce $\Phi_{X,Y}(x, y)$, je dána vztahem

$$\begin{aligned} C_\rho(u, v) &= \Phi(\Phi_X^{-1}(u), \Phi_Y^{-1}(v)) = \\ &= \int_{-\infty}^{\Phi_X^{-1}(u)} \int_{-\infty}^{\Phi_Y^{-1}(v)} \frac{1}{2\pi(1-\rho^2)^{\frac{1}{2}}} \exp\left\{-\frac{x^2-2\rho xy+y^2}{2(1-\rho^2)}\right\} dx dy, \end{aligned}$$

kde ρ je parametrem kopuly a Φ_X^{-1} , resp. Φ_Y^{-1} , je inverze jednorozměrné distribuční funkce normálního rozdělení.

Koeficienty závislosti dolního a horního chvostu normální kopuly lze vyjádřit pomocí

$$\lambda_L(X, Y) = \lambda_U(X, Y) = 2 \lim_{x \rightarrow -\infty} \rho \cdot \left(x \frac{\sqrt{1-\rho}}{\sqrt{1+\rho}} \right) = 0.$$

To znamená, že extrémní hodnoty proměnných X a Y se ve svých chvostech

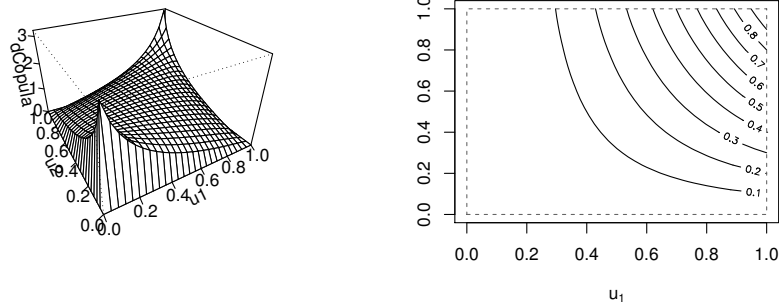
neovlivňují a nastávají nezávisle na sobě. Dále jsou uvedeny vztahy pro další míry asociace: Kendallovo τ a Spermanovo ρ_S :

$$\begin{aligned}\tau &= 2\pi^{-1} \arcsin(\rho), \\ \rho_S &= 6\pi^{-1} \arcsin\left(\frac{\rho}{2}\right).\end{aligned}$$

Vztah pro normální kopulu dvojrozměrného náhodného vektoru lze zobecnit pro n -rozměrný případ. Nechť Φ značí sdruženou distribuční funkci normovaného normálního rozdělení n -rozměrného náhodného vektoru (U_1, U_2, \dots, U_n) s korelační maticí \mathbf{R} , $\Phi_i, i = 1, \dots, n$ budiž jejich (stejně) jednorozměrná marginální distribuční funkce a $\Phi_i^{-1}, i = 1, \dots, n$ nechť jsou jejich kvantilové funkce. Pak platí

$$C_{\mathbf{R}}(u_1, u_2, \dots, u_n) = \Phi(\Phi_1^{-1}(u_1), \Phi_2^{-1}(u_2), \dots, \Phi_n^{-1}(u_n)).$$

Výsledná hustota kopuly $C_{\mathbf{R}}$ je zobrazena na obrázku 4.4. Dále jsou zde znázorněny vrstevnice její sdružené distribuční funkce. Lze si všimnout, že tato kopula je sdruženě symetrická kolem bodu $[0, 5; 0, 5]$ a je omezena Fréchet-Hoeffdingovými mezemi.



Obrázek 4.4: Hustota (vlevo) a vrstevnice sdružené distribuční funkce (vpravo) Gaussovy kopuly s parametrem $\rho = 0,5$.

4.2.2. Studentova t-kopula

Podobně jako předchozí typ kopuly je Studentova t-kopula generována danou eliptickou distribuční funkcí, v tomto případě mnohorozměrným Studentovým

rozdělením s parametrem ρ a s ν stupni volnosti a také přejímá symetrii tohoto rozdělení. Na rozdíl od Gaussovy kopuly má však těžší chvosty, a tudíž je zde větší pravděpodobnost společného výskytu extrémních hodnot náhodných veličin.

Mezi stupni volnosti ν a touto pravděpodobností panuje nepřímá úměra; pokud se stupně volnosti zvýší, pak pravděpodobnost společného výskytu extrémních hodnot poklesne. Naopak výsledkem zvýšení hodnoty parametru ρ a snížením stupně volnosti ν je větší síla závislosti mezi chvosty λ_U a λ_L .

Studentova kopula se často používá při modelování společného selhání plateb aktiv určeného portfolia.

Studentova dvourozměrná t-kopula je dána vztahem

$$C_{\rho,\nu}(u, v) = \int_{-\infty}^{t_{\nu}^{-1}(u)} \int_{-\infty}^{t_{\nu}^{-1}(v)} \frac{1}{2\pi(1-\rho^2)^{1/2}} \left\{ 1 + \frac{x^2 - 2\rho xy + y^2}{\nu(1-\rho^2)} \right\}^{-\frac{\nu+2}{2}} dx dy,$$

kde ρ a ν jsou parametry a t_{ν}^{-1} je inverzí distribuční funkce jednorozměrného Studentova t-rozdělení s ν stupni volnosti, střední hodnotou 0 a rozptylem $\frac{\nu}{\nu-2}$.

Koeficienty závislosti dolního a horního chvostu jsou vzhledem k sdružené symetrii stejné, tedy vyhovují vztahu

$$\lambda_L(X, Y) = \lambda_U(X, Y) = 2t_{\nu+1} \cdot \rho(-\sqrt{\nu+1} \cdot \sqrt{1-\rho} \cdot \sqrt{1+\rho}),$$

kde $t_{\nu+1}$ označuje distribuční funkci jednorozměrného Studentova t-rozdělení s $\nu = 1$ stupni volnosti.

Hodnoty měř asociace Studentovy kopuly lze vypočítat jako

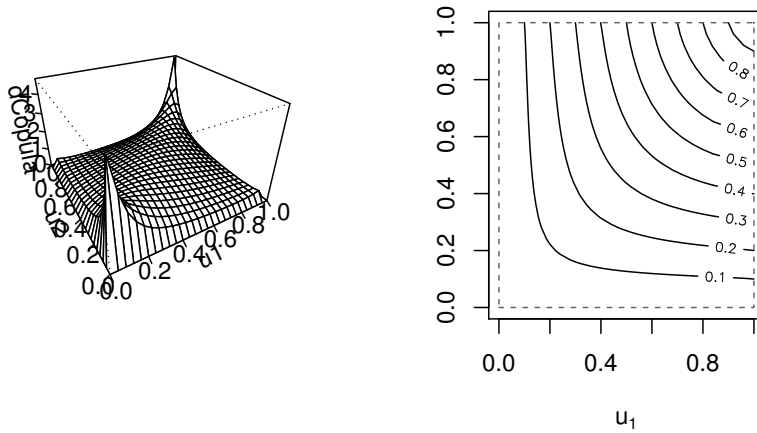
$$\begin{aligned} \tau &= 2\pi^{-1} \arcsin(\rho), \\ \rho_S &= 6\pi^{-1} \arcsin\left(\frac{\rho}{2}\right). \end{aligned}$$

Nechť je dán n -rozměrný náhodný vektor $\mathbf{U} \sim t_{\nu}^n(\boldsymbol{\mu}, \mathbf{R})$, kde $\boldsymbol{\mu}$ je nulový vektor a pro $\nu > 2$ $\boldsymbol{\Sigma} = \frac{\nu}{\nu-2}\mathbf{R}$. Kopula vektoru \mathbf{X} je Studentova t-kopula s ν stupni volnosti, která je zadána následujícím předpisem

$$C_{\nu, \mathbf{R}}^t(\mathbf{u}) = t_{\nu, \mathbf{R}}^n(t_{\nu}^{-1}(u_1), \dots, t_{\nu}^{-1}(u_n)),$$

kde $t_{\nu, \mathbf{R}}^n$ značí n -rozměrnou Studentovu distribuční funkci náhodného vektoru \mathbf{U} , $t_{\nu}(u_i), i = 1, \dots, n$ distribuční funkce jeho stejně rozdělených marginálních náhodných veličin a t_{ν}^{-1} jejich pseudoinverzní funkce.

Průběh hustoty t-kopuly s parametry $\rho = 0,5$ a $\nu = 3$ je znázorněn na obrázku 4.5.



Obrázek 4.5: Hustota (vlevo) a vrstevnice kumulativní distribuční funkce (vpravo) Studentovy kopuly s parametry $\rho = 0,5$ a $\nu = 3$.

4.3. Archimédovské kopuly

Archimédovské kopuly jsou také nazývány explicitními kvůli uzavřené formě jejich předpisů. Na rozdíl od implicitních kopul se neodvozují od známých tříd mnohorozměrných distribučních funkcí, ale jsou dány tzv. *generátory*.

Nechť φ je spojitá, ryze rostoucí funkce $\varphi : \mathbb{I} \rightarrow \langle 0, \infty \rangle$ taková, že $\varphi(1) = 0$, a necht' $\varphi^{[-1]}$ je její pseudoinverze. Ta je dána předpisem

$$\varphi^{[-1]}(t) = \begin{cases} \varphi^{-1}(t) & \text{pro } t \in \langle 0, \varphi(0) \rangle, \\ 0 & t \geq \varphi(0). \end{cases} \quad (4.1)$$

V [30] bylo zavedeno, že pokud je navíc φ konvexní, pak se $C : \mathbb{I}^2 \rightarrow \mathbb{I}$ dána předpisem

$$C(u, v) = \varphi^{[-1]} \{ \varphi(u) + \varphi(v) \}$$

nazývá archimédovská kopula.

Nechť C je dvourozměrná archimédovská kopula s generátorem φ , pak je její hustota dána předpisem

$$c_{u,v} = \frac{\varphi''(C(u, v))\varphi'(u)\varphi'(v)}{[\varphi'(C(u, v))]^3};$$

derivace na hranici $\varphi(u) + \varphi(v) = \varphi(0)$ neexistují, blíže viz [18].

V [32] je definována mnohorozměrná archimédovská kopula daná generátorem $\varphi(u)$ vztahem

$$C(u_1, \dots, u_n) = \begin{cases} \varphi^{[-1]}(\varphi(u_1) + \dots + \varphi(u_n)) & \text{pro } \sum_{i=1}^n \varphi(u_i) \leq \varphi(0), \\ 0 & \text{jinak,} \end{cases}$$

kde $\varphi(u)$ je funkce, jejíž první a druhé derivace jsou spojité, $\varphi(1) = 0$, $\varphi'(u) < 0$, $\varphi''(u) > 0$ a $0 \leq u \leq 1$.

Následně uvedeme tři specifické příklady archimédovských kopul a jejich generátorů. Přitom budou pro každou rodinu, podobně jako u eliptických kopul, popsány její základní vlastnosti: zda je symetrická, chování chvostů a z měr asociace Kendallovo τ a Spearmanovo ρ_S . Některé rodiny obsahují i fundamentální kopuly, tj. nezávislou, horní a dolní kopulu, a tudíž uvedeme i jejich parametry.

4.3.1. Gumbelova kopula

První archimédovskou kopulou, kterou uvedeme, je Gumbelova kopula. Jedná se o asymetrickou kopulu, která přisuzuje vyšší závislost hornímu chvostu (viz hustota Gumbelovy kopuly na obrázku 4.6). Gumbelova třída kopul je zajímavá především proto, že jako jediná rodina archimédovských kopul obsahuje tzv. kopulu extrémních hodnot. Navíc obsahuje i horní a nezávislou kopulu.

Generátorem Gumbelovy kopuly je funkce

$$\varphi_\delta(t) = (-\ln t)^\delta, \delta \in \langle 1, \infty \rangle, \varphi_\delta^{[-1]}(s) = \exp\{-s^{\frac{1}{\delta}}\},$$

kde $1 \leq \delta < \infty$ je parametr závislosti.

Její distribuční funkce je vyjádřena vztahem

$$C_\delta(u, v) = \exp(-[(-\ln u)^\delta + (-\ln v)^\delta]^{\frac{1}{\delta}}),$$

kde $0 \leq u \leq 1$ a $0 \leq v \leq 1$.

Pokud jde parametr závislosti Gumbelovy kopuly limitně do nekonečna to jest $\delta \rightarrow \infty$, určuje tento horní kopulu. Jeho opačná mez, tedy $\delta = 1$, vede na kopulu nezávislou.

Gumbelova kopula je závislá pouze ve svém horním chvostu. Hodnota této závislosti je dána předpisem $\lambda_U(X, Y) = 2 - 2^{\frac{1}{\delta}}$ a koeficient závislosti dolního chvostu je $\lambda_L(X, Y) = 0$. Z tohoto důvodu není dolní kopula kopulou Gumbelovou.

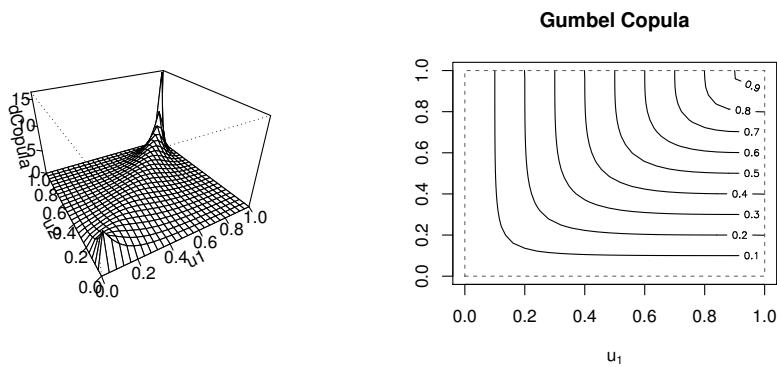
Kendalovo τ nabývá pro dvojrozměrný případ hodnoty

$$\tau(\delta) = 1 - \frac{1}{\delta}.$$

Gumbelova n -rozměrná kopula je dána vztahem

$$C_\delta(u_1, \dots, u_n) = \exp\left(-\left\{\sum_{i=1}^n (-\ln u_i)^\delta\right\}^{\frac{1}{\delta}}\right),$$

kde $u_i \in \langle 0, 1 \rangle$, pro $i = 1, \dots, n$.



Obrázek 4.6: Hustota a vrstevnice distribuční funkce Gumbelovy kopuly s $\delta = 2$.

4.3.2. Claytonova kopula

Další archimédovskou třídou kopul je tzv. Claytonova. Jedná se o rodinu kopul, které jsou asymetrické a kladou větší důraz na závislost dolního chvostu (viz obrázek 4.7). Obsahuje všechny fundamentální kopuly.

V praxi se pomocí kopul z této třídy modeluje například tzv. syndrom zlomeného srdce (jedná se o závislost času úmrtí vdovy či vdovce po smrti svého druha).

Generátor Claytonovy kopuly a jeho pseudoinverze jsou dány vztahy

$$\varphi(s, \delta) = \max\{0, (1 + \delta s)^{-\delta^{-1}}\}, -1 \leq \delta < 0, \varphi^{[-1]}(t, \delta) = t^{-\delta} - 1, \text{ kde } 0 \leq t \leq 1.$$

Rozdělení Claytonovy dvourozměrné kopuly je pak dáno jako:

$$C_\delta(u, v) = (u^{-\delta} + v^{-\delta} - 1)^{-\delta^{-1}}, u, v \in \langle 0, 1 \rangle,$$

kde $\delta \in \langle -1, 0 \rangle \cup (0, \infty)$ je parametr závislosti.

Platí, že pokud se parametr závislosti limitně blíží k nekonečnu ($\delta \rightarrow \infty$), pak určuje distribuční funkci horní kopuly a $\delta \rightarrow 0^+$, resp. $\delta = -1$, určuje distribuci součinnové, resp. dolní kopuly.

Hodnota Kendallova τ je rovna

$$\tau(\delta) = \frac{\delta}{\delta + 2}.$$

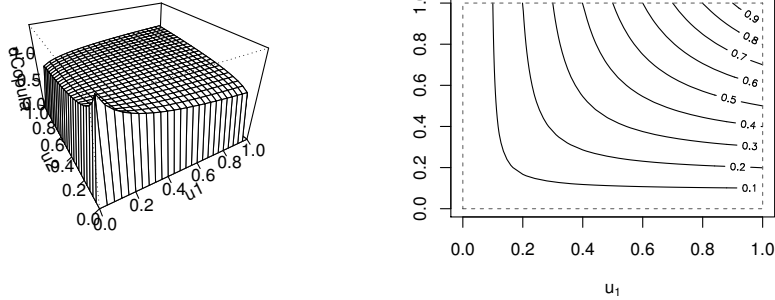
Claytonova kopula je závislá pouze na svém dolním chvostu. Tedy pro koeficient závislosti horního chvostu platí $\lambda_U(X, Y) = 0$ a koeficient závislosti dolního chvostu $\lambda_L(X, Y) = 2^{-\frac{1}{\delta}}$.

Dále uvedeme generátor pro n -rozměrnou Claytonovu kopulu, $n \geq 3$:

$$\varphi(t) = \frac{t^{-\delta}}{\delta} - 1, \delta \in \langle -1, \infty \rangle.$$

Výsledná distribuční funkce Claytonovy kopuly je tak ve tvaru

$$C(u_1, \dots, u_n, \delta) = \left[\sum_{i=1}^n u_i^{-\delta} - (n-1) \right]^{-\frac{1}{\delta}}, u_i \in \langle 0, 1 \rangle, i = \{1, \dots, n\}.$$



Obrázek 4.7: Hustota (vlevo) a vrstevnice (vpravo) Claytonovy kopuly.

4.3.3. Frankova kopula

Poslední archimédovskou třídou kopul, kterou zde uvádíme, je tzv. Frankova třída. Tato třída je speciální zejména v dvourozměrném případě, neboť je jedinou třídou archimédovských kopul, která je radiálně symetrická, tedy platí $C(u, v) = C(v, u)$ (viz obrázek 4.8). Je hojně využívána ve finančním sektoru.

Frankova kopula také dovoluje vyjádřit negativní závislost mezi marginálními proměnnými. Ačkoliv jsou její chvosty symetrické, příkládá jim pouze slabou závislost. Tato třída kopul obsahuje všechny fundamentální kopuly.

Generátorem třídy Frankových kopul nazýváme funkci

$$\varphi(s, \delta) = -\delta^{-1} \ln[1 - (1 - e^{-\delta})e^{-s}],$$

$$\varphi^{[-1]}(t, \delta) = -\ln \left[\frac{e^{-\delta t} - 1}{e^{-\delta} - 1} \right], \delta \in (-\infty, \infty) \text{ a } 0 \leq t \leq 1.$$

Distribuční funkce Frankovy kopuly je následně dána vztahem

$$C_{\delta}(u, v) = -\frac{1}{\delta} \ln \left\{ 1 + \frac{(e^{-\delta u} - 1)(e^{-\delta v} - 1)}{e^{-\delta} - 1} \right\}, \delta \neq 0, u, v \in \langle 0, 1 \rangle.$$

Fundamentální kopuly vzniknou limitním přechodem parametru δ ,

$$\lim_{\delta \rightarrow -\infty} C_{\delta} = W,$$

$$\lim_{\delta \rightarrow \infty} C_{\delta} = M,$$

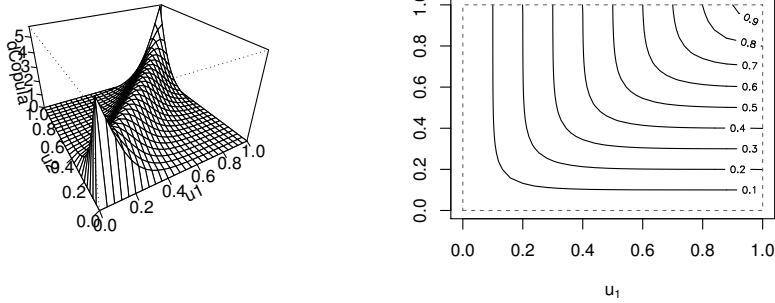
$$\lim_{\delta \rightarrow 0} C_\delta = \Pi,$$

kde W je horní, M je dolní a Π je součinnová kopula.

Hodnota Kendallova τ je dána:

$$\tau(\delta) = 1 - \frac{4}{\delta} + \frac{4}{\delta^2} \int_0^\delta \frac{t}{(e^t - 1)} dt.$$

Jestliže je parametr $\delta < 0$, pak τ nabývá záporných hodnot.



Obrázek 4.8: Hustota (vpravo) a vrstevnice (vlevo) Frankovy kopuly.

Mnohorozměrná distribuční funkce Frankovy kopuly je dána následujícím vztahem

$$C_\delta(\mathbf{u}) = \varphi \left\{ \sum_{j=1}^n \varphi^{[-1]}(u_j) \right\} = -\frac{1}{\delta} \ln \left\{ 1 - \frac{\prod_j (1 - e^{-\delta u_j})}{(1 - e^{-\delta})^{n-1}} \right\}, \mathbf{u} \in \langle 0, 1 \rangle^n,$$

kde $\langle 0, 1 \rangle^n$ je kartézský součin n intervalů $\langle 0, 1 \rangle$.

Zobecněním však přicházíme o vlastnost symetrie: n -rozměrná třída Frankových kopul ($n \leq 3$) je radiálně asymetrická se zašikmením k hornímu rohu.

Na závěr této kapitoly uvádíme tabulku s empirickými výsledky základních vlastností a mírami závislosti uvedených typů kopul. Pro každý typ byly vybrány tři rozdílné parametry a na jejich základě namodelovány teoretické distribuční funkce kopul.

Dle postupu popsaného v [27] z nich bylo následně metodou Monte Carlo náhodně vygenerováno 10000 pozorování, na základě kterých byly odhadnuty koeficienty závislosti chvostů a míry asociace dle již uvedených vztahů viz kapitoly 2 a 3.

Samotný proces generování kopul probíhá následovně:

- Generování nezávislých pozorování $(x_1, y_1), \dots, (x_N, y_N)$ náhodných veličin $X, Y \sim Ro(0, 1)$.
- Pro $i = 1, \dots, N$ je vypočtena hodnota $u_i = c_{x_i}^{-1}(y_i)$, kde $c_{x_i}(y_i) = P(X \leq x_i | Y = y_i) = \frac{\partial C(x, y)}{\partial y}$ a $c_{x_i}^{-1}$ představuje pseudoinverzní funkci k c_{x_i} .
- Výsledný pár pozorování (x_i, u_i) .

Radiální symetrie kopuly byla testována prostřednictvím bootstrapu podrobně popsaného v [28], jenž porovnává pseudopozorování $\hat{U}_i, i = 1, \dots, N$ vygenerovaná z empirické kopuly C_N a pseudopozorování $-\hat{U}_i, i = 1, \dots, N$ vygenerované z její příslušné kopuly přežití \bar{C}_N . Pokud je kopula radiálně symetrická, měl by platit vztah $C = \bar{C}$. Hodnota příslušné testovací statistiky Q_N je tedy dána vztahem $Q_N = \sum_{i=1}^N \{C_N(\hat{U}_i) - \bar{C}_N(\hat{U}_i)\}^2$, přičemž za radiální lze považovat takové kopuly, pro něž je Q_N přibližně rovno 0. Příslušné p-hodnoty vyplývající z testovací statistiky Q_N je možné nalézt v [28].

Výsledky jsou uvedeny v tabulce 4.1. Je z nich patrné, že Frankovy, Gaussovy a Studentovy kopuly jsou radiálně symetrické.

Kopula		Radiální symetrie		Chvosty		Míry asociace	
Typ	Parametry	Q_N	p-hodnota	λ_L	λ_U	t	r_S
C_n	$\rho = 0,1$	0,0182	0,6449	0	0	0,0638	0,0955
	$\rho = 0,5$	0,0457	0,2493	0	0	0,3333	0,4826
	$\rho = 0,9$	0,0245	0,3691	0	0	0,7129	0,8914
C_t	$\rho = 0,1; \nu = 2$	0,0484	0,3452	0,2152	0,2152	0,0638	0,0955
	$\rho = 0,5; \nu = 4$	0,0227	0,8377	0,2532	0,2532	0,333	0,4825
	$\rho = 0,9; \nu = 6$	0,0183	0,6439	0,5630	0,5630	0,7129	0,8915
C_{Cl}	$\delta = 0,1$	0,2152	0	0,0005	0,0010	0,0476	0,0071
	$\delta = 1$	5,0982	0,0005	0,5	0	0,3333	0,4785
	$\delta = 10$	4,5276	0,0005	0,9330	0	0,8333	0,9582
C_{Fr}	$\delta = -10$	0,0351	0,2443	0	0	-0,6658	-0,8602
	$\delta = -0,1$	0,0437	0,3382	0	0	-0,0111	-0,0167
	$\delta = 20$	0,0171	0,4810	0	0	0,8164	0,9579
C_{Gu}	$\delta = 1$	0,0512	0,2333	0	0	0	0
	$\delta = 2$	2,4584	0,0005	0	0,5858	0,5	0,6829
	$\delta = 3$	1,6176	0,0005	0	0,7401	0,6667	0,8482

Tabulka 4.1: Empirické výsledky základních vlastností a měr závislosti Gaussovy (C_n), Studentovy (C_t), Gumbelovy (C_{Cl}), Claytonovy (C_{Fr}) a Frankovy kopuly (C_{Gu}), získané náhodným generováním bodů z příslušného rozdělení.

Kapitola 5

Odhady parametrů kopul

V praxi je často zadán soubor pozorování popisující naměřené hodnoty náhodných proměnných. Tento soubor dat obsahuje informace nejen o chování jednotlivých veličin, ale také o jejich vzájemné interakci.

Analýzu a popis tohoto vzájemného působení lze provést prostřednictvím teorie kopul. Ta daný problém dělí na dvě základní části: výběr vhodné třídy kopul, jež obsahuje základní informace o průběhu interakce, a následný odhad jejího parametru.

Tato kapitola je zaměřena na druhou část. Budou zde popsány základní metody pro odhad parametrů kopul, založené na metodě maximální věrohodnosti (ML, CML, IFM).

V celém následujícím textu jsou pro jednoduchost uvažovány spojitě jednorozměrné marginální distribuční funkce. Základní citované materiály jsou [9], [16], [17], [18], [22], [25].

5.1. Metody maximální věrohodnosti

Nechť je dán výběr $\mathbf{X}_1, \dots, \mathbf{X}_N$ o rozsahu N z rozdělení náhodného vektoru $\mathbf{X} = (X_1, X_2, \dots, X_n)$. Distribuční funkce, resp. hustoty jednorozměrných marginálních proměnných jsou značeny F_j , resp. $f_j, j = 1, \dots, n$ a jejich interakce je dána hustotou kopuly c .

Hlavním úkolem je tedy odhadnout hodnoty vektoru $\boldsymbol{\theta} = (\beta_1, \dots, \beta_n, \boldsymbol{\alpha})$,

kde $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_n)$ určují (obecně vícerozměrné) parametry marginálních distribučních funkcí a $\boldsymbol{\alpha}$ vektor parametrů kopuly.

Nejznámější a nejpřesnější metodou odhadu parametrů je **metoda úplné maximální věrohodnosti**, která vychází z logaritmické věrohodnostní funkce.

$$l(\boldsymbol{\theta}) = \sum_{i=1}^N \ln\{c[F_1(x_{i1}, \boldsymbol{\beta}_1), \dots, F_n(x_{in}, \boldsymbol{\beta}_n), \boldsymbol{\alpha}]\} + \sum_{i=1}^N \sum_{j=1}^n \ln\{f_i(x_{ij}, \boldsymbol{\beta}_j)\}.$$

Maximálně věrohodný odhad (ML odhad) vektoru parametrů $\boldsymbol{\theta}$ je tedy dán vztahem

$$\hat{\boldsymbol{\theta}}_{MLE} = \operatorname{argmax}_{\boldsymbol{\theta} \in \Theta} l(\boldsymbol{\theta}),$$

kde Θ je parametrický prostor.

V [22] je ukázáno, že se za splnění podmínek regularity jedná o konzistentní a asymptoticky eficientní odhad. Mimo to zmíněná citace uvádí jeho další vlastnosti, kterými se zde nebudeme zabývat.

Tato metoda odhaduje všechny parametry najednou, což však s rostoucím n přináší i vyšší výpočetní náročnost. Proto byly navrženy další metody odhadů.

5.1.1. Inference marginálních distribucí

Inference marginálních distribucí neboli IFM byla navržena v [25] v roce 1997. Odhady parametrů $\boldsymbol{\theta}$ vychází ze stejného modelu jako ML, avšak celý proces dělí na dvě fáze, čímž snižuje výpočetní obtížnost předchozí metody.

V první fázi odhaduje parametry vektoru $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_n)$. Pokud je každá marginální distribuční funkce $F_j, j = 1, \dots, n$ určena parametry $\boldsymbol{\beta}_j$, pak $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_n)$ je odhadnut z příslušných hustot jako

$$\hat{\boldsymbol{\beta}}_{IFM} = \operatorname{argmax}_{\boldsymbol{\beta}} \sum_{i=1}^N \sum_{j=1}^n \ln f_i(x_{ij}, \boldsymbol{\beta}_j).$$

V druhé fázi je opět využita metoda maximální věrohodnosti pro odhad parametrů. Celý proces se však zjednoduší, neboť parametry marginálních roz-

dělení jsou zafixovány na hodnotách svých odhadů. Obdržíme tedy

$$\hat{\boldsymbol{\alpha}}_{IFM} = \operatorname{argmax}_{\boldsymbol{\alpha}} \sum_{i=1}^N \ln c\{F_1(x_{i1}, \hat{\boldsymbol{\beta}}_{IFM}), \dots, F_n(x_{in}, \hat{\boldsymbol{\beta}}_{IFM}), \boldsymbol{\alpha}\}.$$

Vzhledem k rozdělení výpočtu parametrů na dvě fáze je jasné, že se zvyšuje pravděpodobnost vyššího odchýlení od původního rozdělení. Jsou-li snad typy a parametry marginálních distribučních funkcí nevhodně či špatně odhadnuty, modelované parametry kopuly budou značně vzdáleny od jejich reálné hodnoty. Riziko tohoto nebezpečí se zvyšuje, pokud je původní datový soubor asymetrický či pokud přisuzuje velkou váhu svým chvostům. Více je možné se o dané metodě dozvědět z [25] a o jejích asymptotických vlastnostech z [24].

5.1.2. Kanonická metoda maximální věrohodnosti

Poslední zmíněná metoda věrohodnosti je tzv. kanonická metoda maximální věrohodnosti (CML), neboli metoda odhadu maximální pseudověrohodnosti. Tato metoda se používá především v případech, kdy je potřeba konzistentní odhad parametru $\boldsymbol{\alpha}$.

Hlavním rozdílem v její konstrukci je nspecifikace marginálních distribučních funkcí původního datového souboru. Používá empirické distribuční funkce každé marginální proměnné pro transformaci i -tého pozorování, tedy realizace vektoru (X_{i1}, \dots, X_{in}) , na pseudonáhodné pozorování s rovnoměrnými marginálními náhodnými proměnnými $\mathbf{U}_i = (U_{i1}, \dots, U_{in})$.

Pseudonáhodný výběr kopuly je zkonstruován tvorbou vektorů $\hat{\mathbf{U}}_1, \hat{\mathbf{U}}_2, \dots, \hat{\mathbf{U}}_N$. Přitom

$$\hat{\mathbf{U}}_i = (U_{i1}, U_{i2}, \dots, U_{in}) = (\hat{F}_1(X_{i1}), \hat{F}_2(X_{i2}), \dots, \hat{F}_n(X_{in})),$$

kde vztah

$$\hat{F}_j(x) = \frac{1}{N+1} \sum_{i=1}^N \mathbf{1}(x_{ij} \leq x_j)$$

označuje odhad j -té marginální distribuční funkce, kde $j = 1, \dots, n$ a $\mathbf{1}(x_{ij} \leq x_j) = 1$, právě tehdy, když $x_{ij} \leq x_j$ a $\mathbf{1}(x_{ij} \leq x_j) = 0$, když $x_{ij} > x_j$.

Transformací dochází ke změně závislostí mezi náhodnými vektory, neboť $\hat{F}_j(x), j = 1, \dots, N$ jsou konstruovány ze všech datových vektorů.

Odhad parametru $\boldsymbol{\alpha}$ je tedy dán vztahem

$$\hat{\boldsymbol{\alpha}}_{CML} = \operatorname{argmax}_{\boldsymbol{\alpha}} \sum_{i=1}^N \ln\{c(U_{i1}, \dots, U_{in}, \boldsymbol{\alpha})\}.$$

Tato metoda je detailněji popsána v [9] a [16].

Kapitola 6

Testy dobré shody

Předchozí kapitola byla věnována výběru vhodné třídy kopuly a metodám následného odhadu jejích parametrů. Pro zhodnocení toho, zda je výběr vhodný či naopak, je potřeba výsledek testovat. Hypotézou je tedy tvrzení, že závislostní struktura dat je modelována určenou kopulou a parametrem (skalárním či vektorovým).

Ověření lze dosáhnout porovnáváním hodnot empirické kopuly a naším parametrickým odhadem kopuly, za předpokladu, že je daná hypotéza správná.

V následujícím textu budou uvedeny dva nejčastěji používané testy dobré shody a jejich úpravy, které mohou být vypočítány pomocí parametrického bootstrapu anebo tzv. multiplikátoru centrální limitní věty. V celé kapitole se omezíme pouze na první variantu. Dále bylo čerpáno z [2], [25], [28], [32].

V celé sekci je uvažována následující situace: Nechť je dán náhodný výběr tvořený spojitými náhodnými veličinami X_1, X_2, \dots, X_n , kde $n \geq 2$. Dále nechť F_1, F_2, \dots, F_n jsou jejich jednorozměrné marginální distribuční funkce a kopula C popisuje jejich závislostní strukturu; představuje tedy sdruženou distribuční funkci náhodného vektoru $\mathbf{U} = (U_1, U_2, \dots, U_n)$, kde $U_j = F_j(X_j), \forall j = 1, 2, \dots, n$.

Testujeme hypotézu,

$$H_0 : C \in \vartheta_0; \vartheta_0 = \{C_{\boldsymbol{\theta}} : \boldsymbol{\theta} \in \Theta\} \text{ vs. } H_1 : C \notin \vartheta_0, \vartheta_0 = \{C_{\boldsymbol{\theta}} : \boldsymbol{\theta} \in \Theta\},$$

kde $\Theta \subset \mathbb{R}^p$, $p \geq 1$ je parametrický prostor.

6.1. Cramer von Misesův test

Hlavním konceptem Cramer von Misesova testu je porovnání vzdálenosti mezi \widehat{C}_N a $C_{\boldsymbol{\theta}_N}$, tedy neparametrickým odhadem kopuly C , a parametrickým odhadem vypočítaným pomocí $\boldsymbol{\theta}_N$ odhadu parametru $\boldsymbol{\theta}$. Protože často není známo rozdělení marginálních proměnných, lze se vyhnout jejich odhadu vytvořením tzv. pseudopozorování závislých na pořadí jednotlivých pozorování.

Nechť je dáno $N \geq 2$ nezávislých kopií vektoru \mathbf{X} , to znamená, že náhodné vektory $\mathbf{X}_1 = (X_{11}, \dots, X_{1n}), \dots, \mathbf{X}_N = (X_{N1}, \dots, X_{Nn})$ jsou nezávislé a stejně rozdělené, a F_{N1}, \dots, F_{Nn} jsou empirické distribuční funkce svých teoretických protějšků F_1, F_2, \dots, F_n . Nejprve je potřeba určit soubor pseudopozorování, který vznikne transformací užitím empirických distribučních funkcí, tedy

$$\widehat{\mathbf{U}}_i = (\widehat{U}_{i1}, \dots, \widehat{U}_{in}) = \left(\frac{N}{N+1} F_{i1}(X_{i1}), \dots, \frac{N}{N+1} F_{in}(X_{in}) \right), \forall i \in 1, \dots, N,$$

kde náhodné veličiny $(N+1)\widehat{U}_{ij}, j \in \{1, \dots, n\}$ značí pořadí pozorování X_{ij} v souboru $\{X_{1j}, \dots, X_{Nj}\}$.

Přeskálování $\frac{N}{N+1}$ je použito proto, aby nedocházelo k numerickým obtížím v případě, že při výpočtu parametrů hustoty kopuly je použito pseudopozorování ležící na hranici $\langle 0, 1 \rangle^n = \langle 0, 1 \rangle \times \dots \times \langle 0, 1 \rangle$. Vzhledem k invarianci kopuly v případě rostoucích transformací lze jistě tvrdit, že nebude mít vliv na limitní chování empirické kopuly.

Nově vzniklá pseudopozorování již nejsou nezávislá, neboť byla generována pozorováními ze všech vektorů obsažených ve výběru, a jsou přibližně rovnoměrně rozdělena na intervalu $(0, 1)$. Empirickou kopulu lze vyjádřit neparametrickým odhadem jako

$$\widehat{C}_N(\mathbf{u}) = \frac{1}{N} \sum_{i=1}^N \mathbf{1}(\widehat{\mathbf{U}}_i \leq \mathbf{u}), \mathbf{u} = (u_1, \dots, u_n) \in \langle 0, 1 \rangle^n,$$

kde $j \in \{1, \dots, n\}$ a $\mathbf{1}(u_{ij} \leq u_j) = 1$, právě tehdy, když $u_{ij} \leq u_j$ a $\mathbf{1}(u_{ij} \leq u_j) = 0$, když $u_{ij} > u_j$.

Za platnosti hypotézy H_0 lze vypočítat odhady parametrů kopuly z vytvořeného pseudopozorování pomocí metod maximální věrohodnosti, viz kapitola 5. Testovací statistika hypotézy H_0 je založená na zkoumání vzdálenosti (vyjádření pomocí L^2 normy) mezi \widehat{C}_N a C_{θ_N} . Tuto normu však lze také nahradit jinými normami či jinou simulační metodou např. tzv. Rosenblattovou transformací. Toto zobrazení poskytuje jednoduchou cestu, jak lze náhodný vektor s danou distribuční funkcí, rozložit na vzájemně nezávislé a rovnoměrně rozdělené komponenty. Více informací o této metodě viz [2] a[3].

Předpis testovací statistiky Cramer von Misesova testu je dán

$$\begin{aligned} S_N &= N \int_{(0,1)^n} \{\widehat{C}_N(u_1, \dots, u_n) - C_{\theta_N}(u_1, \dots, u_n)\}^2 d\widehat{C}_N(u_1, \dots, u_n) = \\ &= \sum_{i=1}^N \{\widehat{C}_N(\widehat{U}_{i1}, \dots, \widehat{U}_{in}) - C_{\theta_N}(\widehat{U}_{i1}, \dots, \widehat{U}_{in})\}^2. \end{aligned}$$

Odchytky mezi všemi pseudopozorováními \widehat{C}_N a C_{θ_N} mají stejné váhy, a tudíž neberou v potaz požadavek na upřednostňované chování závislostní struktury v určitých částech distribuce. V případě, že je potřebné zohledňovat závislosti ve chvostech kopuly, bude tedy tento test nerelevantní.

Celý postup Cramer von Misesova testu prostřednictvím parametrického bootstrapu naznačíme v další podkapitole.

6.1.1. Postup parametrického bootstrapu

Pro velkou hodnotu N^* a $\forall k \in \{1, \dots, N^*\}$ opakujte následující kroky:

- Pro $\theta_N = T_N(\mathbf{X}_1, \dots, \mathbf{X}_n)$ generujte N nezávislých pozorování $\widehat{\mathbf{U}}_{1,k}, \dots, \widehat{\mathbf{U}}_{N,k}$ z distribuce C_{θ_N} .
- Vypočítejte $\widehat{\theta}_{N,k} = T_N(\widehat{\mathbf{U}}_{1,k}, \dots, \widehat{\mathbf{U}}_{n,k})$, a nechte $\forall \mathbf{u} \in \mathbb{R}^n$

$$\widehat{C}_{N,k}(\mathbf{u}) = \frac{1}{N} \sum_{i=1}^N \mathbf{1}(\widehat{\mathbf{U}}_{i,k} \leq \mathbf{u}).$$

- $S_{N,k} = \Phi(G_{N,k}^F)$, kde $G_{N,k}^F = \sqrt{N}(\widehat{C}_{N,k} - C_{\theta_{N,k}})$.

Předpokládáme, že pro velké hodnoty S_N zamítáme hypotézu H_0 . Za splnění předpokladů regularity je p-hodnota pro test rovna

$$p = \frac{1}{N^*} \sum_{k=1}^{N^*} \mathbf{1}(S_{N,k} > S_N).$$

U této metody je potřeba si uvědomit, že s růstem datového souboru se velmi navyšuje početní náročnost, neboť v jednotlivých iteracích potřebujeme generovat hodnoty z nafitované kopuly, ale také odhadnout parametr kopuly.

6.2. Anderson - Darlingův test

Alternativním testem dobré shody, který však zohledňuje váhy chvostů, je Andersen-Darlingův test. Jeho testovací statistika je dána vztahem

$$R_N = N \int_{(0,1)^n} \left(\frac{\widehat{C}_N(u_1, \dots, u_n) - C_{\theta_N}(u_1, \dots, u_n)}{(C_{\theta_N}(u_1, \dots, u_n)\{1 - C_{\theta_N}(u_1, \dots, u_n)\} + \zeta_m)^m} \right)^2 d\widehat{C}_N(u_1, \dots, u_n),$$

obsahujícím odhad θ_N vektoru parametrů θ a ladící parametry $m \geq 0$ a $\zeta_m \geq 0$. Přitom \widehat{C}_N je z podstaty empirického (skokového) tvaru kopuly diskrétní míra.

Jestliže je ladící parametr $m = 0$, Andersen-Darlingova a Cramer von Misesova statistika si budou rovny, tedy $R_N = S_N$. Pro $m > 0$ lze ladit nesrovnalosti chvostů mezi \widehat{C}_N a C_{θ_N} , tedy v bodech $\widehat{\mathbf{U}}_1, \dots, \widehat{\mathbf{U}}_N$, kde $C_{\theta_N}(1 - C_{\theta_N})$ je blízká 0.

Ladící parametr $\zeta_m > 0$ zajišťuje, že jmenovatel nebude moc blízko 0, což by zapříčinilo nesnáze jak v teoretickém, tak praktickém výpočtu.

Nulová hypotéza $R_N = 0$ závisí na neznámé hodnotě parametru θ , a proto je potřebné sestavit algoritmus pro výpočet p-hodnot pro každý jeho odhad. Lze tak učinit parametrickým bootstrapem uvedeným v části 6.1.1. Více informací o vlastnostech testu je uvedeno v [11], [20], [27], [28].

Kapitola 7

Praktická část

V této kapitole se zaměříme na aplikaci popsané teorie kopul na konkrétních příkladech. První z nich ilustruje aplikaci metody kopul v případě, že je znám pouze datový soubor a je potřeba aproximovat mnohorozměrnou sdruženou distribuční funkci. Jsou zde porovnávány různé třídy kopul, metody odhadu jejich parametrů a aplikovány vybrané testy dobré shody.

V druhé kapitole je stručně nastíněna problematika hodnocení finálních derivátů CDO, jíž se metoda kopul nechvalně proslavila. Je ukázán zjednodušený postup modelování doby do defaultu a následně jsou diskutovány výsledky ohodnocení této obligace. Podrobné principy užitého modelu lze najít v [6], [22], [29] a [34].

K modelování obou příkladů bylo použito softwaru **R**, konkrétně knihovny *copula* a *fitdistrplus*. Knihovna *copula* byla speciálně vytvořena pro tvorbu kopul, jejich modelování a simulace rozdělení jejich užitími, viz [19], [27], [42].

7.1. Aplikace modelování kopul na reálných datech

Jako první praktický příklad uvedeme modelování kopuly na reálných datech. Naším hlavním úkolem je aproximace sdružené distribuční funkce pomocí nalezení vhodného typu kopuly.

Data, která byla zkoumána, jsou složena ze tří časových řad hodnot akcií

na burze Dow Jones, představující denní výnosnost akcií společnosti Intel (ozn. X_1), Microsoft (ozn. X_2) a General Electrics (ozn. X_3) v časovém horizontu 5 let (1996–2000). Dohromady se jedná o 1262 pozorování hodnoty každé akcie. Datový soubor lze nalézt v programu **R** v knihovně *copula* pod názvem *rdj*. Výběr prvních deseti pozorování je uveden na obrázku 7.1.

```
> rdj
      Date      INTC      MSFT      GE
1 1996-01-03 -0.0150378774 -0.0325598781 0.0016924128
2 1996-01-04 -0.0043403997 0.0057388835 -0.0119494265
3 1996-01-05 0.0000000000 -0.0115108920 0.0017182207
4 1996-01-08 0.0021655766 -0.0014551794 0.0068351939
5 1996-01-09 -0.0466236967 -0.0728815171 -0.0096277539
6 1996-01-10 -0.0160371340 0.0269205196 -0.0321196242
7 1996-01-11 0.0495681817 0.0503017280 0.0017761609
8 1996-01-12 -0.0044073224 -0.0101453979 -0.0035554822
9 1996-01-15 -0.0591060833 -0.0386424574 0.0000000000
10 1996-01-16 0.0435328780 0.0459023042 0.0088565039
```

Obrázek 7.1: Vzorek datového souboru složeného z hodnot výnosnosti akcií společností Intel (INTC), Microsoft (MCSF) a General Electrics (GE).

Prvním úkolem je prozkoumat chování jednotlivých výnosových křivek každé ze společností, které představují realizace náhodných veličin X_1 , X_2 a X_3 . Pro tuto analýzu se využívá identifikace vhodného rozdělení pravděpodobnosti výnosnosti každé akcie. Obdržíme (empirické) jednorozměrné marginální distribuční funkce, které následně otestujeme pomocí několika metod a vybereme nejvhodnější typ distribuční funkce příslušné náhodné veličiny. Základní číselné charakteristiky jsou uvedeny v tabulce 7.1.

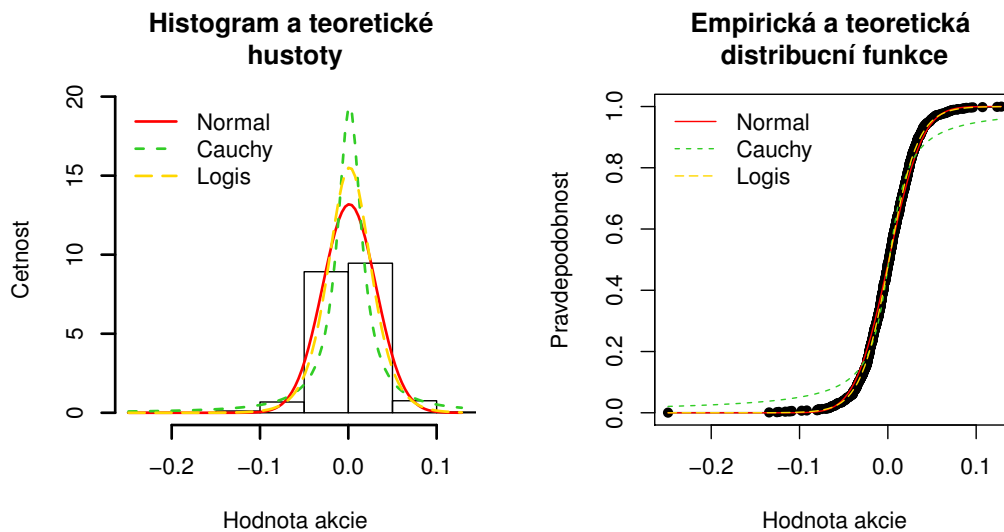
	Arit. průměr	Směr. odchylka	Šikmost	Špičatost
INTL	0,0011	0,0302	-0,5860	5,1106
MCFT	0,0011	0,0253	-0,3016	5,4941
GE	0,0011	0,0183	-0,0051	1,1800

Tabulka 7.1: Základní číselné charakteristiky (aritmetický průměr, směrodatná odchylka, šikmost a špičatost) výnosových křivek.

Vzhledem k povaze dat budeme uvažovat pouze takové spojité distribuční funkce, které umožňují hodnotám z intervalu $(-\infty, 0)$ nabývat kladných hodnot. Budeme tedy uvažovat normální, logistické a Cauchyovo rozdělení.

Pro každé z těchto rozdělení byly metodou maximální věrohodnosti odhadnuty parametry a následně jsme zkoumali, která distribuční funkce byla pro danou výnosovou křivku nejvhodnější.

První testovací metodou bylo grafické znázornění empirické a teoretických distribučních funkcí viz obrázek 7.2. Dle tohoto zobrazení se zdá, že distribucemi, které nejlépe popisují data Intelu, jsou distribuční funkce logistického rozdělení pravděpodobností s parametrem polohy 0,0015 a parametrem měřítka 0,0161 a distribuční funkce normálního rozdělení s parametry střední hodnoty 0,0011 a rozptylem 0,0302. Která z těchto distribučních funkcí je nejvhodnější, otestu-



Obrázek 7.2: Porovnávání empirické distribuční funkce výnosové křivky akcií Intelu s distribuční funkcí normálního rozdělení s parametry $(0,0011; 0,0302)$, distribuční funkcí Cauchyova rozdělení s parametry $(0,0014; 0,0163)$ a distribuční funkcí logistického rozdělení s parametry $(0,0015; 0,0161)$.

jeme prostřednictvím dalších statistických testů získaných pomocí funkce *gofstat* v **R**, viz obrázek 7.3. Jsou zde uvedeny tři statistické testy: Kolmogorovův-Smirnovův test, Cramer von Misesův test a Anderson-Darlingův test, které jsou blíže popsány v [38] a dvě statistická kritéria: Akaikeho, resp. Bayesovo informační kritérium, značené AIC, resp. BIC. Pro obě tato kritéria platí, že čím nižší je jejich hodnota, tím lépe distribuční funkce popisuje data.

Na obrázku 7.3 jsou žlutě zvýrazněny ty hodnoty testovacích statistik, pro které na hladině 0,05 nezamítáme hypotézu, že náhodná veličina má dané rozdělení pravděpodobností a zeleně jsou zvýrazněny nejmenší hodnoty testovacích kritérií. Lze vidět, že nejvhodnější se zdá být distribuční funkce logistického rozdělení s parametry 0,0015 a 0,0161.

Goodness-of-fit statistics			
	Normal	Cauchy	Logis
Kolmogorov-Smirnov statistic	0.04686213	0.06598831	0.01487581
Cramer-von Mises statistic	0.66741015	1.25374827	0.04725753
Anderson-Darling statistic	4.94784104	14.45666626	0.47534907
Goodness-of-fit criteria			
	Normal	Cauchy	Logis
Akaike's Information Criterion	-5245.379	-5032.045	-5353.870
Bayesian Information Criterion	-5235.098	-5021.764	-5343.589

Obrázek 7.3: Hodnoty testových statistik a kritérií udávající, zda distribuční funkce normálního, Cauchyova či logistického rozdělení je vhodnou aproximací empirické distribuční funkce výnosu akcií Intelu.

Podobně byly zkoumány i výnosy akcií Microsoftu a General Electrics, viz obrázky 7.9, 7.10 a obrázky 7.11, 7.12 uvedené v Dodatku. Také pro modelování těchto výnosů byly vybrány distribuční funkce logistického rozdělení.

Dalším úkolem je výběr vhodné kopuly popisující závislostní strukturu mezi daty. Vhodná kopula by měla ctít dosavadní závislosti mezi jednotlivými akciemi, které jsou popsány Spearmanovými korelačními koeficienty (viz obrázek 7.4).

```
> cor(x,method='spearman')
      INTC      MSFT      GE
INTC 1.000000 0.5688595 0.3364263
MSFT 0.5688595 1.000000 0.3994784
GE    0.3364263 0.3994784 1.0000000
```

Obrázek 7.4: Matice Spearmanových korelačních koeficientů datového souboru.

Dále byl vypočítán test radiální symetrie dle postupu uvedeného v kapitole 4. Hypotéza o symetrii empirické kopuly nebyla na základě testovací statistiky $Q_n = 0,0518$ a p-hodnoty 0,5559 zamítnuta. Je tedy vhodné uvažovat, že symetrické kopuly budou lépe aproximovat danou závislostní strukturu.

Navíc koeficienty závislostí chvostů empirické kopuly, uvedené na obrázku 7.5, ukazují, že je vhodné uvažovat rodinu kopul, jenž je ve svých chvostech závislá.

\$lower				\$upper			
	[,1]	[,2]	[,3]		[,1]	[,2]	[,3]
[1,]	1.00000	0.1134200	0.0000000	[1,]	1.0000000	0.28970624	0.18434044
[2,]	0.11342	1.0000000	0.2028493	[2,]	0.2897062	1.00000000	0.06539645
[3,]	0.00000	0.2028493	1.0000000	[3,]	0.1843404	0.06539645	1.00000000

Obrázek 7.5: Matice korelačních koeficientů závislosti dolních (vlevo) a horních (vpravo) chvostů datového souboru.

Na základě interpretace uvedených charakteristik lze předběžně usoudit, že Studentova t-kopula bude dobře aproximovat empirickou kopulu datového souboru. Abychom toto tvrzení ověřili, porovnáváme míru shody původní závislostní struktury a následujících rodin kopul: Claytonovy, Gumbelovy, Frankovy, Gaussovy a již zmíněné Studentovy.

Známe-li předpis kopuly, je potřebné co nejlépe odhadnout její parametry z datového souboru. Za tímto účelem byly vybrány dvě metody: inference marginálních distribucí a metoda maximální pseudověrohodnosti. Metodu úplné maximální věrohodnosti jsme pro její výpočetní náročnost neuvažovali. Popis a použití těchto metod nalezneme v kapitole 5. V tabulce 7.2 jsou uvedeny výsledné odhady parametrů. Kopulu, která nejvíce vyhovuje datovému souboru, vybereme pro-

	IFM	CML
Claytonova	0,4999	0,5857
Frankova	2,887	2,867
Gumbelova	1,366	1,368
Gaussova	0,5646; 0,3256; 0,388	0,5781; 0,34; 0,4017
Studentova	0,5828; 0,3599 ; 0,4265	0,5877; 0,3593; 0,4225
	$\nu = 8, 29$	$\nu = 6, 502$

Tabulka 7.2: Odhady parametrů Gumbelovy, Claytonovy, Frankovy, Gaussovy a Studentovy kopuly prostřednictvím inference marginálních distribucí (IFM) a kanonické metody maximální věrohodnosti (CML).

střednictvím testů dobré shody a hodnot věrohodnostních funkcí, kdy testujeme všech pět tříd kopul a jejich příslušné odhadnuté parametry.

Testy, které za tímto účelem byly vybrány, jsou Cramer von Misesův test (s hodnotou testovací statistiky S_N) a Anderson-Darlingův test (s hodnotou statistiky R_N), který, na rozdíl od S_N , zdůrazňuje rozdíly v závislosti chvostů prostřednictvím parametru $\zeta_m = 3$. Navíc byly vypočítány hodnoty log-věrohodnostní funkce (LL) dané kopuly. Čím vyšší hodnota tohoto kritéria, tím lépe model odpovídá původním datům.

Tabulky 7.3 a 7.4 uvádí hodnoty jednotlivých testů a příslušných p-hodnot pro kopuly s odhadnutými parametry metodou maximální pseudověrohodnosti, resp. inference marginálních distribucí. Ačkoliv dávají obě metody odhadů parametrů podobné výsledky, je zřejmé, že CML je dle kritéria log-věrohodnosti vhodnější.

Kopulami, jež nebyly na hladině $\alpha = 0,05$ zamítnuty žádným testem, jsou obě Studentovy t-kopuly a Gaussova kopula s parametry 0,5781, 0,34 a 0,4017. Na základě LL kritéria je upřednostňována Studentova t-kopula s parametry $\theta = (0,5877; 0,3593; 0,4225)$ a stupněm volnosti $\nu = 6,502$, a proto ji přijmeme jako pro nás nejvhodnější.

Druh kopuly	LL	S_N		R_N	
		statistika	p-hodnota	statistika	p-hodnota
Gauss	375,7089	0,0137	0,0347	0,0137	0,0347
Gumbel	294,5982	0,3607	0,0050	0,1154	0,0050
Frank	323,3839	0,1408	0,0050	0,0447	0,0050
Student	419,2701	0,0406	0,0545	0,0145	0,0248
Clayton	273,9629	0,8812	0,0050	0,2742	0,0050

Tabulka 7.3: Testy dobré shody pro Gaussovu, Gumbelovu, Frankovu, Studentovu a Claytonovu kopulu s parametry odhadnutými metodou CML.

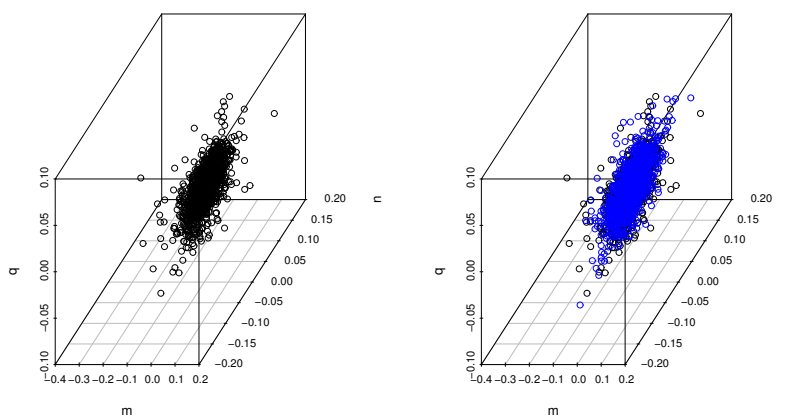
Posledním a finálním úkolem je vytvoření mnohorozměrné simultánní distribuční funkce. To provedeme spojením příslušných marginálních distribučních funkcí $X_1 \sim \text{Log}(0,0015; 0,0161)$, $X_2 \sim \text{Log}(0,0012; 0,0124)$ a $X_3 \sim \text{Log}(0,0011; 0,0101)$ se Studentovou kopulou $t_{6,502}(0,5877; 0,3593; 0,4225)$. V softwaru **R** lze pro tento proces využít funkce *mvd*.

K posouzení vhodnosti modelované sdružené distribuční funkce jsme využili

Druh kopuly	LL	S_n		R_n	
		statistika	p-hodnota	statistika	p-hodnota
Gauss	364,4283	0,0436	0,0248	0,0137	0,0545
Gumbel	291,7988	0,3607	0,0050	0,1154	0,0050
Frank	322,6404	0,1408	0,0050	0,0447	0,0050
Student	407,5654	0,0406	0,0446	0,0145	0,0347
Clayton	236,3491	0,8812	0,0050	0,2742	0,0050

Tabulka 7.4: Testy dobré shody pro Gaussovu, Gumbelovu, Frankovu, Studentovu a Claytonovu kopulu s parametry odhadnutými metodou IFM.

grafickou metodu. Z nové simultánní distribuce bylo vygenerováno 1262 pozorování a ty byly následně zobrazeny do grafu původních pozorování, viz obrázek 7.6. Vidíme, že data relativně dobře popisují převážnou část původních dat.



Obrázek 7.6: 3D graf původních dat (černé) a 1262 pozorování vygenerovaných z vytvořené simultánní distribuční funkce (modré).

7.2. Hodnocení CDO

Při popisu modelování hodnocení CDO budeme čerpat z následující literatury: [4], [5], [6], [13], [22], [23], [31] a [34].

Nezajištěné dluhové obligace (neboli CDO) jsou finanční deriváty, jejichž

hlavní podstatou je diverzifikace rizika různých finančních aktiv (úvěry, hypotéky, jiná CDO) mezi věřitele s odlišnými preferenčními stupni rizika.

Vydavatel CDO vybírá na základě daných kritérií portfolio dluhových pohledávek, jež tvoří rizikovou strukturu CDO, a dle této struktury vydává dluhopisy (na rozdíl od tradičních cenných papírů) s různou rizikovou mírou a výnosem tzv. „tranše“. Platí, že čím je tranše rizikovější, tím vyšší výnos slibuje. Jejich druhy se nejčastěji dělí na:

- ekvitní: nejrizikovější s nejvyšší mírou výnosu,
- mezaninové,
- seniorní: nejméně riziková s nejmenším úrokovým procentem.

Výplaty probíhají na základě tzv. principu vodopádu, kdy je nejprve vyplácen výnos seniorní tranše, poté mezaninové a nakonec ekvitní. Dojde-li k přerušení finančního toku některého z aktiv, prvněmu, komu nebude vyplácena výplata v celé své výši, budou majitelé ekvitní tranše. V případě, že ekvitní tranše nebude stačit na pokrytí ztráty, odebere se odpovídající část výplaty mezaninové tranše atd.

Názorný příklad modelování a ohodnocování tranší nezajištěných dluhových obligací uvádíme dále. Tento proces lze rozdělit na dvě na sebe navazující části. První část se skládá z určení sdružené distribuční funkce, jejíž marginální distribuční funkce jsou tzv. kreditní křivky, určené časem do defaultu jednotlivých aktiv z podkladového portfolia. Část druhá se poté zabývá samotným ohodnocováním jednotlivých tranší.

Předpokládejme, že máme portfolio n aktiv a nechť spojitě náhodné veličiny Z_1, \dots, Z_n představují příslušné doby do defaultu, tedy tu dobu $t \geq 0$, za kterou již nebude dlužník (emitent aktiva) schopen splácet své závazky. Pak lze dle [29] definovat pravděpodobnostní rozdělení selhání finančního toku v daném roce, tedy kreditní křivky, prostřednictvím distribučních funkcí těchto náhodných veličin.

Uvažujme dluhopis A a nechť je dána funkce

$$H(t) = P(Z_A \leq t), t \geq 0.$$

Dle vztahu (2.2) lze funkci přežití, která udává pravděpodobnost, že firma, emitující toto aktivum, nedefaultuje alespoň do času t , definovat vztahem

$$\overline{H} = 1 - H(x) = P(Z_A > t), t \geq 0.$$

Dále definujeme symboly

$${}_tq_x = P(Z_A - x \leq t \mid Z_A > x), t \geq 0,$$

$${}_tp_x = 1 - {}_tq_x = P(Z_A - x > t \mid Z_A > x), t \geq 0.$$

Nechť $t = 1$ a $m \geq 1$. Posloupnost q_0, q_1, \dots, q_m danou pravděpodobnostmi defaultu dluhopisu A v roce x za podmínky, že tak neučinil do začátku roku x , kde $x \in \{0, 1, \dots, m\}$, tj.

$$q_x = P(Z_A - x \leq 1 \mid Z_A > x),$$

označujeme jako kreditní křivku pro diskrétní modely.

Strukturu měr defaultů, potřebnou pro modelování kreditních křivek, lze dle [6] získat třemi způsoby: z historických dat defaultů firem publikovaných ratingovými agenturami, Mertonovým opčním modelem a prostřednictvím tržních cen dluhopisů.

V této práci jsme využili první možnosti, tedy dat a informací poskytnutých ratingovými agenturami, konkrétně vážených průměrných měr defaultů firem z let 1981–2014. Dané míry se liší dle ratingu svých jednotlivých akcií. Ten je dle [8] vnímán jako hodnocení schopnosti emitentů dluhových cenných papírů splácet své závazky. Ratingová stupnice je pak dána písmenovou škálou od AAA až do D, kde hodnocení AAA značí velice nízkou a D skoro jistou pravděpodobnost defaultu akcie (viz [36]). Po zvolení ratingu akcie budou jednotlivé členy posloupnosti (až na první člen, jenž je roven první míře defaultu) dány vztahem

$$q_{x+t} = \frac{{}_{t+1}q_x - {}_tq_x}{{}_tp_x},$$

kde ${}_tq_x$ jsou vážené průměrné míry defaultů firem v roce x .

Máme-li namodelovány marginální distribuční funkce, je potřebné obrátit pozornost k jejich vzájemné interakci, neboť v reálném světě lze jen málokdy

předpokládat úplnou nezávislost. Ke zjištění společné distribuční funkce je tedy potřebná metoda pro modelování závislostí, n -rozměrná kopula. Ta umožňuje sestavení sdružené distribuční funkce, ze které je možné (pomocí simulace Monte Carlo) náhodně generovat realizace dob do defaultu všech podkladových aktiv pro dané časové období T . Metoda Monte Carlo je blíže popsána v [39].

V příkladu 7.1 jsme snadno odvodili nejvhodější typ a parametr kopuly pro daná data pomocí testů dobré shody. V tomto případě to však tak lehce nepůjde, neboť pouze z marginálních distribučních funkcí nelze vyčíst korelační strukturu mezi jednotlivými objekty. Je zřejmé, že naše situace nemá jednoznačné řešení.

Z tohoto důvodu se autoři rozhodli ukázat, jaký vliv má výběr kopuly na konečný výsledek. Nejprve jsme uvažovali tři typy kopul: Gaussovu, která je obvykle využívána při modelování CDO, Gumbelovu kopulu, jež je těžká ve svém horním chvostu, a Studentovu kopulu, která je symetrická a těžká v obou svých chvostech. Jejich parametry byly následně odhadnuty prostřednictvím uměle vytvořené korelační matice.

V našem příkladu uvažujeme portfolio složené z 20 dluhopisů, ohodnocených ratingem BBB, tedy se středně nízkou pravděpodobností defaultu za dané období. Dle průměrných korporátních měr defaultu z let (1981–2014) uvedených v tabulce 7.8 jsme určili, že lze doby do defaultu jednotlivých aktiv modelovat prostřednictvím exponenciální distribuční funkce s odhadnutým parametrem 329,5772 tedy $Z_i \sim Exp(329,5772)$, $i = (1, \dots, n)$.

Dále jsme pro 10000 nezávislých portfolií (tj. $N^* = 10000$) zjistili, kolik závazků aktiv z každého nebude splaceno za období jednoho roku. Tuto náhodnou veličinu dále označujeme X . Je zřejmé, že Gaussova, Studentova a Gumbelova kopula mají významný vliv na závislostní strukturu mezi dobami selhání, a tudíž i na výsledné počty defaultů a očekávanou hodnotu CDO.

K tomu, aby výsledky byly porovnatelné, jsme dle praktických zkušeností popsanych v [34] vytvořili „umělou“ korelační matici danou Spearmanovými ρ_S . Závislosti mezi všemi aktivy jsme pevně stanovili na hodnotě 0,2 a následně jsme hledali odpovídající parametry kopul, které by nejlépe vystihovaly danou

závislostní strukturu prostřednictvím vzdálenosti mezi teoretickou a odhadnutou korelační maticí. V tabulce 7.5 uvádíme výsledné odhadnuté parametry kopul. Spojením daných marginálních funkcí a kopuly s určeným parametrem, jsme poté

	Gauss	Gumbel	Student
θ	0,2110	1,163	0,2200; $\nu = 5$

Tabulka 7.5: Odhady parametrů Gaussovy, Gumbelovy a Studentovy kopuly.

získali n -rozměrnou sdruženou distribuční funkci, udávající společné chování dob do defaultu podkladových aktiv.

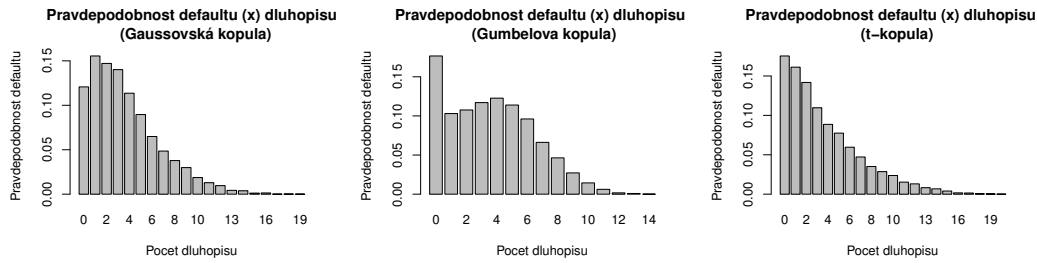
Náhodně bylo generováno N^* vzájemně nezávislých stejně rozdělených portfolií s n aktivy (dluhopisy) a byl zaznamenáván počet defaultů aktiv každého tohoto portfolia. Z výsledků byl vytvořen histogram, udávající četnosti portfolií s danými počty defaultů. Relativní četnosti počtu defaultů (představující zde odhady příslušných pravděpodobností) vzniknou vydělením četností jednotlivých pozorování jejich celkovým počtem.

Histogramy těchto simulací jsou vykresleny na obrázku 7.7. Je znatelné, že jsou při generování díky exponenciálnímu rozdělení přikládány větší váhy nižším hodnotám defaultů. To se projeví na výsledcích jejich očekávaných hodnot, které jsou rovny v případě generování z Gaussovy, Gumbelovy, resp. Studentovy kopuly hodnotám 3,6736, 3,6312, resp. 3,6391, tedy neschopnosti splácet přibližně čtyř firem za dané období, a ve velice malé šanci, že všichni dlužníci defaultují.

Zatímco Gaussova a Studentova kopula mají podobné rozložení četností počtů defaultů, histogram Gumbelovy kopuly se výrazně liší především v pravděpodobnosti výskytu více jak 14 defaultů v portfoliu ($P(X < 14) = 0$), a také v nižší špičatosti rozdělení pravděpodobností počtů defaultů.

Jedinou kopulou, jenž vykazuje jistou (malou) pravděpodobnost, že by všechny finanční toky mohly selhat, je Studentova kopula. Její těžké chvosty zajišťují, že bude brán zřetel i na extrémní případy.

Empirické rozdělení pravděpodobností náhodné veličiny X je uvedeno v Dodatku v tabulkách 7.9, 7.10 a 7.11.



Obrázek 7.7: Histogramy četností portfolií s danými dobami defaultů získané generováním z Gaussovy (vlevo), Gumbelovy (uprostřed) a Studentovy kopuly (vpravo).

Následně se zaměříme na nastínění ohodnocování tranší CDO a ukážeme, jak námi vybraný typ kopuly může ovlivnit příslušné očekávané výnosy. K odhadu očekávané hodnoty CDO je potřebné znát celkovou hodnotu portfolia a rozložení tranší. Fiktivní portfolio, které uvažujeme níže, bude mít hodnotu 20 Kč. Tranše, jimž budou přiděleny výnosy prostřednictvím principu vodopádu, jsou stanoveny procentuálními podíly z celého portfolia aktiv uvedenými v tabulce 7.6. Dále

	Ekvitní	Mezaninová	Seniorní
%	0 - 5	5 - 20	20 - 80

Tabulka 7.6: Struktura tranší fiktivního CDO.

je potřebné odhadnout míru náhrady, kterou lze vymoci v případě, že dlužník není schopen splatit své závazky, a inflační míra. Dle praktických zkušeností se obvykle uvádí míra návratnosti $M = 0,2$. Vzhledem k relativně krátkému období jednoho roku bude uvažována nulová inflační míra.

V případě, že nominální hodnota každé akcie je 1 Kč, bude očekávaná hodnota ($OH(t)$) každé tranše vypočítána vztahem

$$OH(t) = \frac{1}{N^*} \sum_{i=1}^{N^*} \sum_{j=1}^{20} M_{ij}(t) \cdot N_{ij}(t)$$

kde $i = \{1, \dots, N^*\}$, $j = \{1, \dots, 20\}$ a symbol $M_{ij}(T)$ značí míru návratnosti j -té akcie i -tého portfolia v roce T , tedy hodnotu 1 v případě, že v daném portfilu

j -tý dlužník nedefaultuje v roce T , anebo 0,2 když defaultuje. $N_{ij}(T)$ se rovná hodnotě 1 v případě, že je realizace doby do splatnosti vyšší než T a 0 jinak.

Na základě simulace jsme pro každý typ kopul obdrželi histogramy uvedené v Dodatku na obrázcích 7.13, 7.14 a 7.15, udávající výnosy jednotlivých tranší a následně jsme vypočítali jejich očekávané hodnoty, viz tabulka 7.7.

	Ekvitní	Mezaninové	Seniorní
Gauss	0,1207	1,2629	13,6508
Gumbel	0,1766	1,1710	13,5288
Student	0,1754	1,4030	13,4889

Tabulka 7.7: Očekávané hodnoty výnosu ekvitní, mezaninové a seniorní tranše modelované Gaussovou, Gumbelovou a Studentovou kopulou.

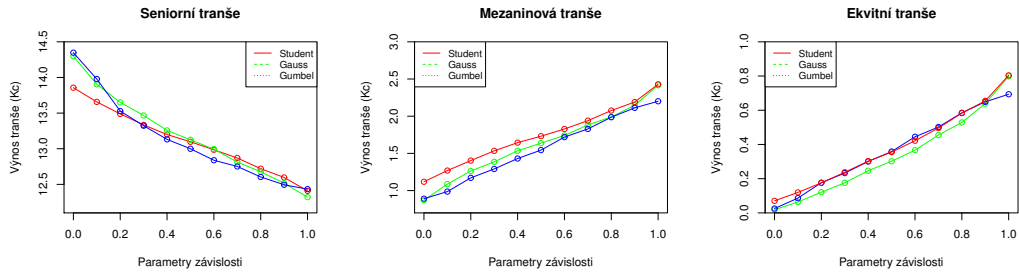
Lze vidět, že ekvitní tranše generovaná prostřednictvím Gumbelovy kopuly má nejvyšší očekávanou výnosnost, očekávaná hodnota výnosu pro mezaninovou tranši je nejvyšší u modelu generovaného Studentovou kopulou a Gaussova kopula vykazuje nejvyšší výplatu seniorní tranše. Zdá se, že závislosti v chvostech mají velký vliv na uvedené očekávané hodnoty. Připomínáme, že Gaussova kopula je ve svých chvostech nezávislá, a tudíž vykazuje pouze ojedinělé výskyty extrémních hodnot. Gumbelova kopula má těžký horní chvost, a proto není modelováno mnoho případů selhání, a Studentova kopula je souměrně těžká na obou svých chvostech, což umožní modelování výskytu jak vysokých, tak nízkých extrémních hodnot.

Na obrázcích 7.13, 7.14 a 7.15 si je také možné všimnout, že zatímco četnosti výnosů seniorní, resp. ekvitní tranše, jsou pro vyšší hodnoty rostoucí, resp. klesající, očekávaná hodnota mezaninové tranše je dána hlavně svými extrémy. Jedná se o tzv. korelační úsměv, který je popsán v [5].

K výběru vhodné kopuly je však ještě potřebné znát vhodnou závislostní strukturu defaultů mezi páry aktiv. Dříve jsme uvažovali stálou hodnotu 0,2. Nyní nás zajímá, jak by se výsledná očekávaná hodnota tranší změnila v případě, že by původní aktiva byla méně diverzifikována, tedy vzrostla by hodnota ρ_S .

Vytvořili jsme tedy další umělé korelační matice, jejichž párové korelační ko-

eficienty se v rámci této matice rovnaly (pro rozdílná aktiva) a vypočítali jsme příslušné parametry pro jednotlivé typy kopul. Zajímala nás především nezáporná závislost podkladových aktiv. Následně jsme postupně ohodnotili všechny tranše fiktivního CDO. Výsledky jsou znázorněny v grafu 7.8.



Obrázek 7.8: Očekávané výnosy seniorní, mezaninové a ekvitní tranše v závislosti na hodnotě ρ_S .

Z grafů je vidět, že hodnota ekvitní a mezaninové tranše se s nižší diverzifikací zvyšuje pro všechny tři typy kopul. Až na portfolia složené z dokonale závislých a nezávislých aktiv, jsou hodnoty očekávaného výnosu ekvitní tranše nejvyšší pro modely Gumbelovy a Studentovy kopuly, zatímco hodnota výnosu mezaninové tranše je zcela dominována modelem generovaným prostřednictvím Studentovy kopuly.

Klesající očekávaná hodnota výnosu pro seniorní tranši je však poněkud komplikovanější. Hodnota této tranše v případě Gumbelovy kopuly, ač nejvyšší v případě nezávislosti aktiv, rychle klesá na nejnižší. Model Gaussovy kopuly dominuje očekávaný výnos seniorní tranše do bodu, kdy je $\rho_S = 0,6$. Také lze sledovat, že pro více diverzifikovaná aktiva je hodnota výnosu této tranše nejvyšší pro kopulu Studentovu.

Jak již bylo zmíněno, v případě, kdy nejsou známa přesná historická data, a tudíž jednotlivé modely závislostí nelze posoudit např. testy dobré shody, řešení není jednoznačné. Nicméně, na předchozích grafech jsme chtěli znázornit, že pokud je vybírána kopula, je vhodné brát v úvahu jak chování kopul tak i závislosti mezi proměnnými.

Závěrem bychom chtěli poukázat na některé další nedostatky konkrétněji. Je zjevné, že velký počet uvedených argumentů drasticky zvyšuje riziko jeho nepřesnosti či nevhodnosti, a proto se zaměřujeme na nejkritičtější z nich. Upozorňujeme, že nejsou uvažovány chybně ohodnocené ratingy jednotlivých aktiv a záměrné zneužití modelu. Dále jsme se zde nevěnovali problému tzv. korelačního úsměvu, který je blíže popsán v [5].

První nedostatek se týká samotného výběru podkladového portfolia. V uvedeném modelu bylo složeno z 20 aktiv, jejichž pravděpodobnosti defaultu byly stejně rozdělené dle ratingu BBB. V reálném případě jsou však k modelování CDO používány různé druhy aktiv (dluhopisy, akcie, CDO, atd.) s rozdílným ratingem, dobou splatnosti a v daleko větším množství. Článek [6] uvádí, jak tyto předpoklady mohou ovlivnit následný model. Zde uvedená analýza prokázala, že je model velice citlivý na jejich změny.

Dalším důležitým předpokladem je použitá struktura závislostí defaultů jednotlivých aktiv. Ačkoliv jsme v našem modelu uvedli stejnou míru asociace pro selhání všech párů aktiv, je vhodné předpokládat, že v normální situaci budou tyto závislosti rozdílné. To zásadně ovlivní jak vhodný typ, tak vypočtené parametry kopul.

Ukázali jsme tak, že výběr druhu kopuly, modelujícího závislosti doby do defaultu, může také významně ovlivnit výsledné hodnoty modelu. Třída použité kopuly musí být tedy pečlivě zvážena na základě jejích vlastností a předpokladů.

Posledním zmíněným problémem je výpočetní náročnost daného modelu. Pokud bude uvažováno portfolio složeno z průkazného vzorku aktiv s různými ratingy (např. 100) a provedeme simulace Monto Carlo pro 10000 nezávislých pozorování, výpočetní čas se mnohonásobně zvýší.

Závěr

Hlavním záměrem této práce bylo seznámit čtenáře se základní teorií a vlastnostmi dvou a vícerozměrných kopul a jejich praktickým využitím.

Toho bylo dosaženo pomocí popisu dané problematiky na dvou příkladech, přičemž první z nich se týkal modelování kopul na reálných datech a druhý z nich znázorňoval fiktivní model ohodnocování CDO. Při řešení praktických příkladů bylo využito softwaru **R**, konkrétně knihovny *copula*.

Tematika kopul je jedna z velice diskutovaných, a dle autorky také velice zajímavých, matematických teorií. V posledních letech vzniklo mnoho (převážně anglických) článků, které se zabývají jak jejími teoretickými poznatky, tak praktickými aplikacemi. Těmi je především myšleno jejich použití při modelování CDO.

Díky tomuto „boomu“ narazila autorka během psaní této práce na několik problémů. První z nich byl, jak již to u odborných článků bývá, překladatelský. Některé názvy matematických metod, které se vyskytují v teorii kopul, ještě nebyly přeloženy do českého jazyka (nebo námi alespoň nebyly objeveny).

Dalším oříškem byl výběr dat a příkladů použití kopul. Jak již bylo zmíněno, ohodnocování a ratingu CDO se již věnovalo spoustu autorů, kteří poukazovali na přínosy, avšak i na slabá místa těchto finančních derivátů. Navíc modely CDO prošly po finanční krizi, která se jim z části kladla za vinu, dalšími změnami. Vzhledem k nejednoznačnosti modelování CDO pomocí kopul, a tudíž k nemožnosti jistého výsledku, se autorka rozhodla praktickou část rozdělit na dvě části.

První (exaktnější) část byla zaměřena na situaci, kdy je dán soubor finančních

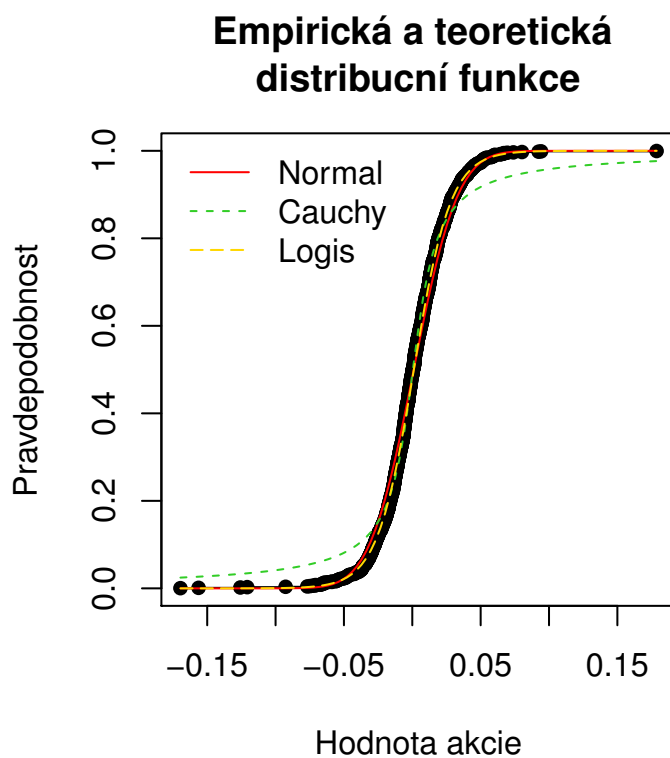
dat. Na nich byl předveden optimální postup modelování mnohorozměrné distribuční funkce pomocí kopul. Byly zde porovnány dvě metody odhadů parametrů (IFM a CML). Ukázalo se, že ačkoliv byly odhady parametrů pro obě, metody podobné, metoda maximální pseudověrohodnosti lépe aproximovala datový soubor. Rodinou, která nejlépe popisovala původní soubor, byla Studentova t-kopula.

Druhá část se již zaměřovala na modelování nezajištěné dluhové obligace. Tato část práce měla poukázat na jeho základní milníky a úskalí. Hlavním úkolem nebylo vybrání vhodné kopuly, ale spíše znázornění, jak jednotlivé typy kopul mohou ovlivnit celé ocenění tranší CDO.

Je zřejmé, že pokud uživatel nerozumí dané problematice či model špatně specifikuje, může obdržet nerelevantní výsledky. Bylo by tedy vhodné, podobně jako v prvním příkladu, nejprve otestovat jednotlivé typy kopul a zjistit, který je nejvhodnější. Bohužel to většinou není možné, neboť jsou CDO při své tvorbě nejprve modelovány fiktivními aktivy a svou opravdovou strukturu nabývají až po emitaci tranší, případně se během své životnosti ještě mění. Tvůrce modelu je tedy odkázán na své odborné znalosti a notnou dávku předpokladů.

Dodatek

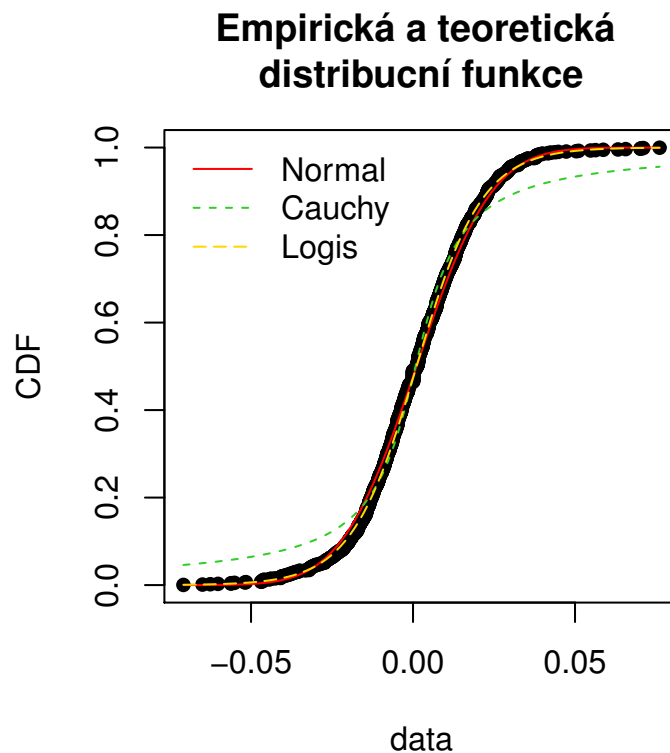
Příklad 1: Aplikace modelování kopul na reálných datech



Obrázek 7.9: Porovnávání empirické distribuční funkce výnosové křivky Microsoftu s distribučními funkcemi normálního rozdělení pravděpodobností s parametry (0,0011; 0,025), Cauchyova rozdělení pravděpodobností s parametry (0,0004; 0,0131) a logistického rozdělení pravděpodobností s parametry (0,0012; 0,0124).

Goodness-of-fit statistics			
	Normal	Cauchy	Logis
Kolmogorov-Smirnov statistic	0.0503882	0.06468534	0.02001198
Cramer-von Mises statistic	0.9450912	1.07980082	0.08634771
Anderson-Darling statistic	5.9938209	12.87105486	0.56247544
Goodness-of-fit criteria			
	Normal	Cauchy	Logis
Akaike's Information Criterion	-5691.439	-5527.796	-5817.396
Bayesian Information Criterion	-5681.158	-5517.515	-5807.113

Obrázek 7.10: Hodnoty testovacích statistik a kritérií udávající, zda distribuční funkce normálního, Cauchyova a logistického rozdělení je vhodnou aproximací empirické distribuční funkce výnosu akcií Microsoftu.



Obrázek 7.11: Porovnávání empirické distribuční funkce výnosové křivky General Electrics s distribučními funkcemi normálního rozdělení pravděpodobností s parametry (0,0011; 0,0182), Cauchyova rozdělení pravděpodobností s parametry (0,0008; 0,0104) a logistického rozdělení pravděpodobností s parametry (0,0011; 0,0101).

Goodness-of-fit statistics			
	Normal	Cauchy	Logis
Kolmogorov-Smirnov statistic	0.0340322	0.06947837	0.01907169
Cramer-von Mises statistic	0.2563692	1.37371847	0.07994654
Anderson-Darling statistic	2.0786294	15.50562877	0.54369191
Goodness-of-fit criteria			
	Normal	Cauchy	Logis
Akaike's Information Criterion	-6520.705	-6189.758	-6557.80
Bayesian Information Criterion	-6510.424	-6179.478	-6547.52

Obrázek 7.12: Hodnoty testovacích statistik a kritérií udávající, zda distribuční funkce normálního, Cauchyova a logistického rozdělení je vhodnou aproximací empirické distribuční funkce výnosu akcií General Electrics.

Příklad 2

Roky	1	2	3	4	5	6	7	8	9	10
	0,2	0,57	0,96	1,46	1,95	2,43	2,84	3,26	3,66	4,06
Roky	11	12	13	14	15					
	4,49	4,84	5,17	5,5	5,84					

Tabulka 7.8: Průměrné kumulativní míry defaultů firem (v %) dané ratingem BBB (1981–2014), viz [36].

# defaultů	0	1	2	3	4	5	6
$P(X \leq x)$	0,1207	0,1555	0,1471	0,1401	0,1136	0,0896	0,0649
# defaultů	7	8	9	10	11	12	13
$P(X \leq x)$	0,0485	0,0378	0,0299	0,0187	0,0129	0,0096	0,0043
# defaultů	14	15	16	17	18	19	20
$P(X \leq x)$	0,0038	0,0012	0,0014	0,0001	0,0002	0,0001	0,0000

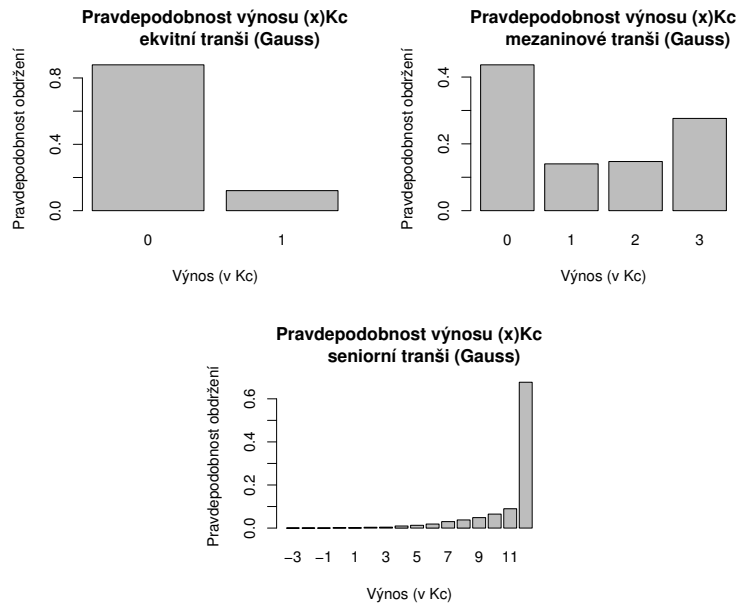
Tabulka 7.9: Rozdělení pravděpodobnosti náhodné veličiny X generované z Gaussovy kopuly.

# defaultů	0	1	2	3	4	5	6
$P(X \leq x)$	0,1766	0,1030	0,1076	0,1170	0,1227	0,1139	0,0961
# defaultů	7	8	9	10	11	12	13
$P(X \leq x)$	0,0662	0,0464	0,0272	0,0144	0,0062	0,0017	0,0009
# defaultů	14	15	16	17	18	19	20
$P(X \leq x)$	0,0001	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000

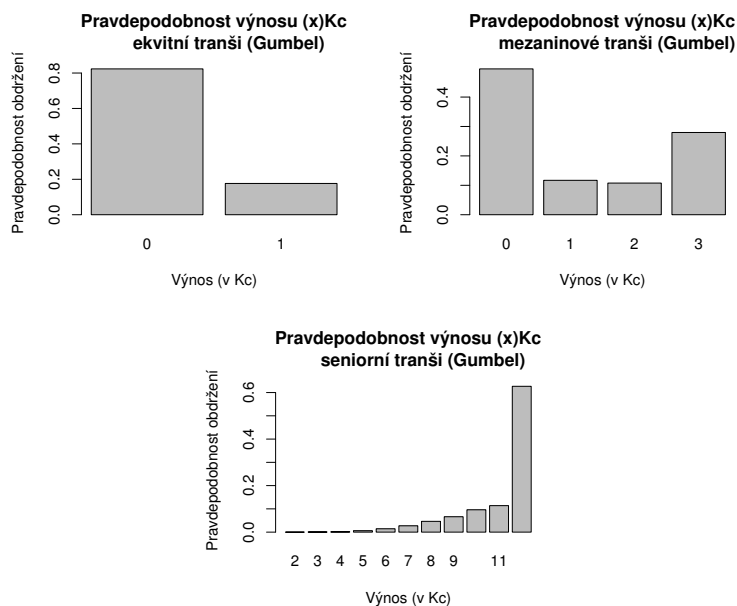
Tabulka 7.10: Rozdělení pravděpodobnosti náhodné veličiny X generované z Gumbelovy kopuly.

# defaultů	0	1	2	3	4	5	6
$P(X \leq x)$	0,1754	0,1612	0,1418	0,1096	0,0886	0,0775	0,0596
# defaultů	7	8	9	10	11	12	13
$P(X \leq x)$	0,0472	0,0351	0,0286	0,0237	0,0153	0,0131	0,0082
# defaultů	14	15	16	17	18	19	20
$P(X \leq x)$	0,0068	0,0040	0,0017	0,0015	0,0006	0,0004	0,0001

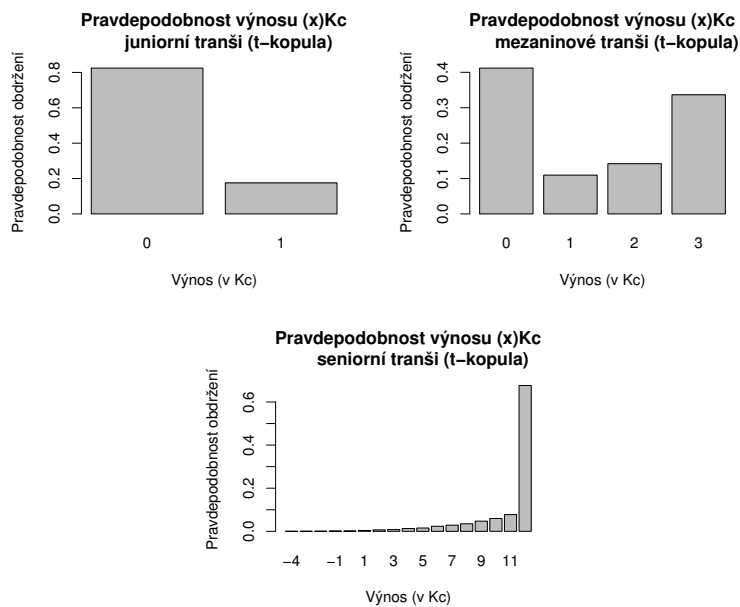
Tabulka 7.11: Rozdělení pravděpodobnosti náhodné veličiny X generované ze Studentovy kopuly.



Obrázek 7.13: Histogramy výnosu akcií každé tranše za dané období modelované prostřednictvím Gaussovy kopuly.



Obrázek 7.14: Histogramy výnosu akcií každé tranše za dané období modelované prostřednictvím Gumbelovy kopuly.



Obrázek 7.15: Histogramy výnosu akcií každé tranše za dané období modelované prostřednictvím Studentovy kopuly.

Literatura

- [1] Asmussen, S. R.: *Steady-State Properties of GI/G/1*. Stochastic Modelling and Applied Probability **51** (2003), s. 266–301.
- [2] Beaudoin, D., Genest, C., Rémillard, B.: *Goodness-of-fit tests for copulas: A review and a power study*. Insurance: Mathematics and Economics **44** (2009), s. 199–214.
- [3] Berg, D.: *Copula Goodness-of-Fit Testing: An Overview and Power Comparison*. The European Journal of Finance **9/15** (2009), s. 675–701.
- [4] Bluhm, Ch.: *CDO Modeling: Techniques, Examples and Applications*. [online]. 2003, [cit. 2017-02-02]. Dostupné z: <http://www.quantcandy.com/blog/wp-content/uploads/2008/01/44.pdf>.
- [5] Brigo, D., Pallavicini, A., Torresetti, R.: *Credit Models and the Crisis. A Journey into CDOs, Copulas, Correlations and Dynamic Models*. John Wiley & Sons, New York, 2010.
- [6] Coval, J., Jurek, J., Stafford, E.: *The Economics of Structured Finance*. Journal of Economic Perspectives **1/23** (2009), s. 3–25.
- [7] Česká národní banka: *ČNB vyhláška č. 123/2007 Sb., o pravidlech obezřetného podnikání bank, spořitelních a úvěrních družstev a obchodníků s cennými papíry*. [online]. [cit. 2017-03-13]. Dostupné z: http://www.cnb.cz/cs/legislativa/vyhlasky/vyhlaska_123_2007_uz_k_2012_07_01.pdf.

- [8] Česká národní banka – Slovník pojmů [online]. [cit. 2017-03-13]. Dostupné z: <https://www.cnb.cz/cs/obecne/slovník/index.html>.
- [9] Demarta, S., McNeil, A. J.: *The t-copula and related copulas*. International Statistical Review **73** (2005), s. 111–129.
- [10] Dobrić, J., Schmid, F.: *Nonparametric estimation of the lower tail dependence λ_l in bivariate copulas*. Journal of Applied Statistics **4/32** (2005), s. 387–407.
- [11] Dufour, J., M., Genest, C., Huang, W.: *A regularized goodness-of-fit test for copulas*. Journal de la Société française de statistique **154** (2013), s. 64–77.
- [12] Embrechts, P.: *Copulas: A personal view*. Journal of Risk Ins. **3/76** (2009), s. 639–650.
- [13] Embrechts, P., Frey, R., McNeil, A.: *Quantitative Risk Management: Concepts, Techniques, Tools*. Princeton University Press, New Jersey, 2005.
- [14] Embrechts, P., Lindskog, F., McNeil, A.: *Modelling Dependence with Copulas and Applications to Risk Management*. [online]. 2001, [cit. 2017-02-21]. Dostupné z: <https://people.math.ethz.ch/~embrecht/ftp/copchapter.pdf>.
- [15] Frahm, G., Junker, M., Schmidt, R.: *Estimating the tail-dependence coefficient: properties and pitfalls*. Insurance: Mathematics and Economics **1/37** (2005), s. 80–100.
- [16] Genest, C., Favre, A. C.: *Everything you always wanted to know about copula modeling but were afraid to ask*. Journal of Hydrologic Engineering **12** (2007), s. 347–368.
- [17] Genest, C.: *Frank's family of bivariate distributions*. Biometrika **74** (1987), s. 549–559.

- [18] Genest, C., MacKay, J.: *The Joy of Copulas: Bivariate Distributions with Uniform Marginals*. The American Statistician **4/40** (1986), s. 280–283.
- [19] Hofert, M., Kojadinovic, I., Yan, J.: *Multivariate Dependence with Copulas*. [online]. 2017, [cit. 2017-02-21]. Dostupné z: <https://cran.r-project.org/web/packages/copula/copula.pdf>.
- [20] Holmes, M., Kojadinovic, I., Yan, J.: *Fast large-sample goodness-of-fit tests for copulas*. Statistica Sinica **21** (2011), s. 841–871.
- [21] Hron, K., Kunderová, P.: *Základy pravděpodobnosti a matematické statistiky*. Vydavatelství Univerzity Palackého, Olomouc, 2013.
- [22] Cherubini, U., Luciano, E., Vecchiarelli, W.: *Copula Methods in Finance*. Wiley & Sons, New York, 2004.
- [23] Cherubini, U., Gobbi, F., Mulini, S., Romagnoli, S.: *Dynamic Copula Methods in Finance*. John Wiley & Sons, New York, 2012.
- [24] Joe, H.: *Asymptotic efficiency of the two-stage estimation method for copula-based models*. Journal of Multivariate Analysis **94** (2005), s. 675–701.
- [25] Joe, H.: *Dependence Modeling with Copulas*. Prometheus, Praha, 2015.
- [26] Kjersti, A.: *Modelling the dependence structure of financial assets: A survey of four copulas*. [online]. 2004, [cit. 2017-02-15]. Dostupné z: <https://www.nr.no/files/samba/bff/SAMBA2204b.pdf>.
- [27] Kojadinovic, I., Yan, J.: *Modeling Multivariate Distributions with Continuous Margins Using the copula R Package*. Journal of Statistical Software **9/34** (2010), s. 1–20.
- [28] Kojadinovic, I.: *Some copula inference procedures adapted to the presence of ties*. [online]. 2017, [cit. 2017-03-03]. Dostupné z: <https://arxiv.org/abs/1609.05519>.

- [29] Li, D.: *On default correlation: a copula function approach*. Journal of Fixed Income **4/9** (2001), s. 43–54.
- [30] Ling, C. M.: *Representation of associative functions*. Publ. Math. Debrecen **12** (1965), s. 189–212.
- [31] Marcantoni, E.: *Collateralized Debt Obligations: A Moment Matching Pricing Technique based on Copula Functions*. Springer Fachmedien Wiesbaden, Wiesbaden, 2014.
- [32] Nelsen, R. B.: *An Introduction to Copulas*. SNTL, Praha, 1983.
- [33] Nelsen, R. B.: *Properties and applications of copulas: A brief survey*. [online]. 2004, [cit. 2016-04-20]. Dostupné z: <http://w4.stern.nyu.edu/ioms/docs/sg/seminars/nelsen.pdf>.
- [34] Riggs, Ch.: *Estimating with the Gaussian, Clayton, and t Copulas. An Exploration of Bond Defaults and Value of a Collateralized Debt Obligation*. [online]. 2013, [cit. 2017-02-16]. Dostupné z: https://economics.nd.edu/assets/134106/riggs_estimating_with_the_gaussian_clayton_and_t_copulas_an_exploration_of_bond_defaults_and_value_of_a_collateralized_debtsubmissioncopulas.pdf.
- [35] Romano, C.: *Applying copula function to risk management*. [online]. 2002, [cit. 2017-02-21]. Dostupné z: https://www.researchgate.net/publication/228760120_Applying_copula_function_to_risk_management.
- [36] S&P Global Ratings: *Default, Transition, and Recovery: 2014 Annual Global Corporate Default Study And Rating Transitions*. [online]. 2015, [cit. 2017-02-16]. Dostupné z: <http://www.spratings.com/documents/20184/774196/2015+Annual+Global+Corporate+Default+Study+And+Rating+Transitions/6d311074-5d56-4589-9ef8-a43615a6493b>.

- [37] Schmidt, T.: *Coping with Copulas*. [online]. 2006, [cit. 2016-04-25]. Dostupné z: <http://down.cenet.org.cn/upfile/36/20061219225057121.pdf>.
- [38] Stephens, M. A.: *Goodness-of-fit techniques*. Marcel Dekker, New York, 1986.
- [39] *Studijní opory předmětu KMA/AR*. [online]. [cit. 2017-03-15]. Dostupné z: http://elearningmath.upol.cz/pluginfile.php/17316/mod_resource/content/6/AR_6_2017_moodle_opvk.pdf.
- [40] Trivedi, P. K., Zimmer, D. M.: *Copula Modeling: An Introduction for Practitioners*. Foundations and Trends in Econometrics, **1/1** (2005), s. 1–111.
- [41] Venter, G. G.: *Tails of copulas*. [online]. [cit. 2017-02-15]. Dostupné z: https://www.casact.org/library/studynotes/Venter_Tails_of_Copulas.pdf.
- [42] Yan J.: *Enjoy the Joy of Copulas*. Journal of Statistical Software **21/4** (2010), s. 1–21.