Posudek

☐vedoucího ☒oponenta

☐diplomové ☒bakalářské práce

Autor/Autorka: Milan Šedivý
Název práce: Backtesting Value-at-Risk: Porovnání vybraných přístupů
Jméno vedoucího/oponenta: Jan Hurt

Matematická úroveň:

☐vynikající ☐velmi dobrá ☐průměrná ☐podprůměrná ☒nevyhovující

Grafická, jazyková a formální úroveň:

☐vynikající ☐velmi dobrá ☒průměrná ☐podprůměrná ☐nevyhovující

Výsledky:

☐originální☐původní i převzaté ☐netriviálníkompilace ☐citované z literatury ☐opsané

☒nedovedu posoudit

Použité metody:

☐nestandardní☒standardní ☐obojí

Aplikovatelnost:

☐přínos pro teorii ☐přínos pro praxi ☐přínos pro praxi i teorii ☒bez přínosu ☐nedovedu posoudit

Věcné chyby:
☐téměř žádné ☐vzhledem k rozsahu a pojednávanému tématu přiměřený počet ☐méně podstatné

četné ☒závažné

Tiskové chyby:

☐téměř žádné ☒vzhledem k rozsahu a pojednávanému tématu přiměřený počet ☐četné

Celková úroveň práce:

☐vynikající ☐velmi dobrá ☐průměrná ☐podprůměrná ☒nevyhovující

Práci

☐doporučuji ☒nedoporučuji

uznat jako diplomovou/bakalářskou. Návrh klasifikace přikládám na zvláštním papíru.

Připomínky a vyjádření vedoucího/oponenta:

(symbol s^h znamená stránku s, řádek h shora, s_d řádek d zdola, případně rozmezí)

Předně autor ani neví, co studoval: Na titulní straně uvádí jako studijní program “Stochastika”.

Četné gramatické chyby jsem částečně vyznačil v exempláři, který jsem měl k dispozici. Jejich výčet by zabral několik stránek.

1. kapitola je nesystematický výběr pojmů a tvrzení z citované Ciprovy knihy Finanční ekonometrie. Podstatné předpoklady však v práci neuvádí, jak vyplývá z některých následujících výtek.

4 uprostřed: VaR může být kladná i záporná v libovolné pozici.

4_{12} “Nejjednodušším takovým způsobem by mohlo být například to, že z nasbíraných dat spočteme všechny momenty do čtvrtého (existuje-li) a zvolíme vhodné rozdělení, které odpovídá našim napozorovaným momentům.” Řešitel by měl uvést příklad, kdy neexistuje čtvrtý moment z “nasbíraných dat”.

4_5 Kvantily $N(0, 1)$ mohou nabývat i nezáporných hodnot. Jestliže pro dlouhou pozici platí, že $u_p < 0$, pak nutně $p < 1/2$. Celá strana 4 je zmatečná.

5 Zde není jasný význam indexu t.

5_1 K dispozici nebývá nekonečný počet pozorování. Bez kritického přístupu není RiskMetrics po matematické stránce na MFF použitelný zdroj informací.

6^5 “RiskMetrics je parametrickou metodou výpočtu VaR, kde odhadujeme jediný parametr ...” je nesmysl.

6^7 “... kde $\alpha + \beta = 1$ tj. řada splňuje (viz níže):...” Nikde níže se α, β nevyskytují.

6 uprostřed: “RiskMetrics přístup dále předpokládá, že logaritmických míry zisku r_t, podmíněné veškerou dostupnou informací Ω_{t-1} do času $t - 1$ (σ-algebra) ...”. Co je dostupná informace? Proč σ –algebra?

6 uprostřed: Co je míra r_t ve frázi “... podmíněný rozptyl míry r_t ...”

7_6 Blábol: “Pokud nekorektně určíme rozdělení, pak zjištěný kvantil nebude odpovídat skutečnému.”. To nebude s. j. nikdy (i při “korektně určeném rozdělení”, i když nevím, co tímto rčením vlastně řešitel míní).

7_4 $\{\epsilon_t\}$ nejsou rezidua.

8^{7-8} Blábol: “Nevýhodou pak je to, že empirický kvantil může být špatně odhadnut, což většinou nastává pro velmi nízké kvantily.” Jak se vůbec může empirický kvantil odhadovat? Vždyť je to sám odhad!!!

8 Odstavec 1.3.4 “Vážená historická simulace”: Ani zbla o simulaci!!! Řešitel očividně netuší co simulace jsou!

8-9 Expected Shortfall (ES), česky očekávaná extrémní ztráta: Nepatří mezi metody, ale číselné charakteristiky rizika. S ohledem na zadání se řešitel vůbec nemusel touto charakteristikou zabývat. To, že tak učinil, mu jedině přitížilo, viz dále.

9 Definice 3: Uvádí se, že (L, Λ) je měřitelný prostor. Pak je Λ nějaký systém podmnožin z L. Měřitelná funkce nezobrazuje z Λ, ale z L. Co je v tomto případě L? Že by funkce měřící riziko byla funkcí množin?

9 Další blábol: "Je zřejmé, že kvantil každého rozdělení vždy existuje a je konečný, tedy i VaR vždy existuje. U ES tomu tak být nemusí a výpočet je jednoznačně komplikovanější, neboť musíme spočítat střední hodnotu nejzazšího konce možných ztrát." Máme-li již odhadnuté VaR, pak např. nejjednodušší neparametrický odhad je výběrový průměr z množiny pozorování, které spadly pod VaR (nebo překročily VaR, v závislosti na interpretaci, která je v předložené práci nejednoznačná). Nic komplikovaného.

Řešitel (jako student MFF) mohl tušit, že příbuzná problematika se někde v syllabech vyskytuje. Odkazuji na svou přednášku Finanční management, pojednávající mj. o mírách rizika a tam také uvedené odkazy na moje přednášky o mírách rizika na konferencích Wolfram Technology Conference:

http://www.karlin.mff.cuni.cz/~hurt/FM2018LS.nb

9_{5-4} "I přes zmíněné zápory je podstatnou výhodou Value-at-Risk fakt, že je velmi užívaným nástrojem pro měření rizika, a tedy je i řádně regulován (Basel II, atp.)" Není mi jasné, co je řádně regulováno.

10^{7} "... hodnota v riziku není v klasickém slova smyslu pozorovatelnou náhodnou veličinou, jedná se o čistě abstraktní údaj ... " Blábol

10^{11} "Napočítanou hodnotu nelze pozorovat." Z čeho se asi "napočítaná hodnota" počítá?

11^{6} "... a označme napočítanou hodnotu rizika ..." Co je "hodnota rizika"? Z čeho se napočítá?

11^{11} Co se myslí frází "... předpověď hodnoty rizika je platná ..."?

11_{14} S ohledem na to, že N je náhodná veličina, tak se asi jedná o podmíněné rozdělení.

11 vzorec (2.1): V jakém smyslu se myslí konvergence $N \to \infty$? Je vůbec možné, aby $N \to \infty$?

11_{2-1} "... kde u každého portfolia jsme provedli měření rizika. U portfolia A došlo k překročení stanovené hranice šestkrát za stanovené období ..." Jak se provádělo "měření rizika"? Co se myslí "stanovenou hranicí"? Uvedený příklad je v každém případě nesrozumitelný.

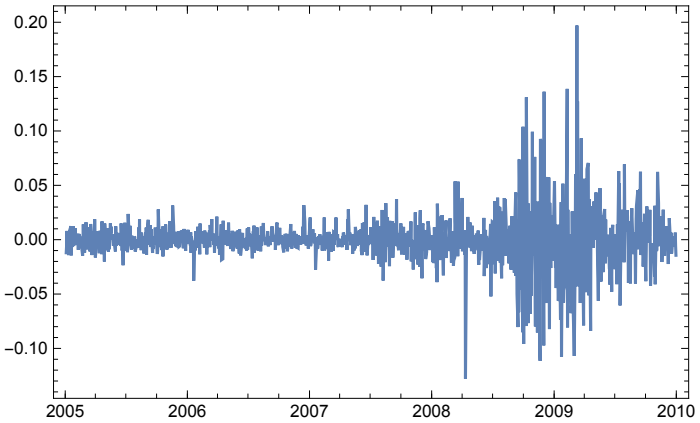
12 "2.5 Multivariate testy": Exisuje český termín mnohorozměrný či vícerozměrný.

16^{5-6} "Testujeme, zda-li distribuce rozdělení dob výskytu má geometrické rozdělení." Co je "distribuce rozdělení"?

17^{15} "... zvlášť předpokládat investice v předem nespecifikované výši (kde za měnu považujeme USD), ...". To je irelevantní, pracuje se s mírami výnosnosti (poměrovými ukazateli).

19 a dále: Na grafech nejsou logaritmické výnosy (jak tvrdí řešitel), ale **minus** logaritmické výnosy, tj. ztráty. Totéž pro tabulky. V tabulkách nejsou výběrové **rozptyly**, ale výběrové **směrodatné odchylky**. Uvádím správné hodnoty pro General Electric spolu s grafem:

Mean	−0.000701984
Variance	0.000553963
S. D.	0.0235364
Skewness	0.041789
Kurtosis	12.8588
5%−Percentile	−0.0370915



Jednoduchá kontrola: Ceny GE při prvním (velkém) propadu jsou

Out[32]=

Mon	7	Apr	2008	37.26
Tue	8	Apr	2008	36.95
Wed	9	Apr	2008	36.44
Thu	10	Apr	2008	36.75
Fri	11	Apr	2008	32.05
Mon	14	Apr	2008	31.75
Tue	15	Apr	2008	31.98
Wed	16	Apr	2008	32.23

19 a dále: Není uvedeno, zda ze serveru Yahoo používá kotace Open, High, Low, Close, Adjusted Close. Z mé kontroly vyplývá, že nejpravděpodobnější je uzavírací cena. S ohledem na to, že i výpočty elementárních charakteristik jsou špatně, mám pochybnosti o správnosti výsledků týkajících se složitějších charakteristik.

22_6 “Obecný postup určení parametru modelu ARMA by spočíval ve využití tzv. kritéria AIC (Akaike information criterion).” Proč zrovna AIC?

23 a dále: horní graf: “Pocatek simulace”: O jaké simulace se jedná? V popisu grafu: “Graf zachycující výpočet hodnot VaR aktiv AAPL ...”: Graf zachycuje výpočet? Aktiva Apple? To bychom asi museli jít do účetní rozvahy.

Přestože téma práce je ve své podstatě statistické, tak v celé práci řešitel prokazuje neznalost základních statistických pojmů jako je náhodný výběr, odhad, neznámé parametry, rozdíl mezi parametrem a jeho odhadem, naprosté nepochopení pojmu “simulace”, atd. , apod.

Místo, datum, podpis vedoucího/oponenta: Praha, 5. 9. 2018, Jan Hurt