

Inlämningsuppgift 1

TIDSSERIEANALYS, VOLATILITETSESTIMERING OCH COPULAS

MÅL

Målet med denna uppgift är att skapa förståelse för egenskaper hos finansiella tidsserier, volatilitetsestimering och copulas samt att ge praktisk träning i att implementera finansiella modeller.

ORGANISERING

Inlämningsuppgiften består av två delar. Del 1, som ska lämnas in individuellt, handlar om att på en given datamängd slutföra ett test som finns tillgängligt via Lisam. För godkänt på denna del ska samtliga frågor vara korrekt besvarade. I del 2 ska ni arbeta med tidsserierna som finns bifogade under Inlämningar på Lisam och följa instruktionerna nedan. Del 2 genomförs i grupper om två studenter utifrån anmälan på Lisam.

Underliggande beräkningar ska implementeras i MATLAB (potentiellt med stöd av Excel) och efterfrågade resultat ska erhållas utifrån *ett* huvudskript per deluppgift (totalt två). Notera att ni ska använda funktionen `printResults()` som finns tillgänglig på Lisam för utskrift av numeriska resultat.

Inlämningen ska bestå av data och MATLAB-filer med underliggande beräkningar, *txt*-filen som genereras av `printResults()` och en *pdf*-fil innehållande svar till frågorna i Problem 1-3 nedan. Se Lisam (eller Kursinformation) för deadline och schema för när redovisning av inlämningsuppgiften sker.

Betyg: Godkänt/Uderkänt

STEG-FÖR-STEG

1. Anmäl dig till en grupp (om 2 studenter) på Lisam för genomförande av inlämningsuppgifterna.
2. Slutför testet på Lisam (samarbeta gärna men notera att varje student ska slutföra testet inloggad på sin student-id). Spara samtliga filer som genererar efterfrågade resultat, packa ihop (ex .zip) och namnge med ditt student-id. **Obs!** Ni lämnar in under Inlämningar på Lisam gemensamt i gruppen och först när både del 1 och 2 är slutförd.
3. Följ instruktionerna nedan och analysera tidsserierna som finns under Inlämningar på Lisam. Kom ihåg att spara all nödvändig kod för att erhålla efterfrågade resultat. Spara också *txt*-filen som genereras av funktionen `printResults()` och sammanställ svar på frågorna nedan i ett *pdf*-dokument.
4. Ladda upp kod från steg 2 och 3 ovan, *txt*-filen samt *pdf*-filen på Lisam. Namnge din grupp med gruppnumret enligt *Anmälan*, exempelvis Grupp 1. Notera att allt material lämnas in genom *en* inlämning, där de individuella zip-filerna från del 1 ingår.
5. Förbered er för redovisning. Vidare instruktioner kommer att publiceras på Lisam.

UPPGIFTER

Ladda dag- och veckodata (*Data.xlsx*) på OMXS30 och USD/SEK från Excel till MATLAB.

Tips: Genom uppgifterna belyser ni fördelningarnas egenskaper på många olika sätt. Använd de olika resultaten för rimlighetsbedömning.

1. Deskriptiv statistik

- a) Börja med att plotta de två tidsserierna med veckodata i samma graf (med separata y-axlar). Beräkna sedan logaritmiska avkastningar för veckodatan och plotta även dessa i en graf. Beräkna slutligen genomsnittlig avkastning och volatilitet utifrån de logaritmiska avkastningarna och bestäm 95 % (dubbelsidiga) konfidensintervall för förväntade logaritmiska avkastningar.

Numeriska resultat:¹ Genomsnittlig avkastning (2), volatilitet (2), konfidensintervall (2). *Årlig enhet och i procent!*

Grafer: Tidsserier (1), Logaritmiska avkastningar (1).

- b) Nu ska vi börja studera om tidsserierna (dag- och veckodata) beskrivs väl av en log-normalfördelning, d v s om de logaritmiska avkastningarna beskrivs väl av en normalfördelning.
- Beräkna skevhet och (excess) kurtosis.
 - Konstruera histogram baserat på de historiska avkastningarna och bestäm 1, 5, 95 och 99:e percentilen. *Tips:* Använd `histfit()` för att jämföra era empiriska fördelningar med normalfördelningen.
 - Konstruera Q-Q plottar för samtliga fyra tidsserier. För att stödja er förståelse rekommenderas ni att själva implementera Q-Q plot utifrån föreläsningsmaterialet. Jämför er implementation med funktionen `qqplot()` i MATLAB.

Numeriska resultat: Skevhet (4), kurtosis (4), percentiler (4x4)

Grafer: Histogram (4), Q-Q plottar (4) (använd subplots!)

Frågor: Beskrivs tidsserierna väl av en log-normalfördelning? Hur skiljer sig resultaten för de två tillgångarna och för dag- respektive veckodata? Om ni observerat skillnader, kan ni föreslå en rimlig förklaring?

¹Numret inom parentes representerar antalet resultat som ska bestämmas, ex. 2 genomsnittliga avkastningar.

2. Volatilitets-estimering. (Använd de två tidsserierna med veckodata)

- a) Beräkna historiska volatiliteter (Equal Weighted Moving Average) rullande över hela tidsserien utifrån 30 och 90 veckors fönster.
- b) Konstruera tidsserier av volatilitetsestimat baserat på EWMA (Exponentially Weighted Moving Average) med $\lambda = 0.94$.
- c) Estimera EWMA och GARCH(1,1) modeller (med och utan variance targeting) baserat på Maximum likelihood estimering (MLE), givet antagande om normalfördelade logaritmiska avkastningar med väntevärde noll (se föreläsning 2). Använd solver i Excel och/eller `fmincon()` i MATLAB.
- d) Standardisera era tidsserier utifrån estimerade GARCH(1,1)-volatiliteter och konstruera Q-Q plottar.

Numeriska resultat: Log-likelihood värden för EWMA (2) och GARCH (4) och parameterestimater (MLE).

Grafer: EqWMA (2x2), EWMA-RiskMetrics (2), EWMA (2), GARCH (2), Q-Q plottar (2). Använd subplots och plotta (årliga) volatiliteter!

Frågor:

- I. Beskrivs era tidsserier bäst som homoskedastiska eller heteroskedastiska?
- II. Hur skiljer sig målfunktionsvärdet (log-likelihood-funktionen) för EWMA och GARCH (med och utan variance targeting)? Är det rimligt?
- III. Hur skiljer sig historisk volatilitet med den medelvärdesåtervändande nivån (jmf V_L) bestämd med MLE i GARCH-modellerna?
- IV. Jämför Q-Q plottarna från 1 b) med de som tagits fram efter standardisering med GARCH-volatiliteter i 2 d). Verkar den antagna modellen med tidsvarierande volatilitet beskriva tidsserierna bättre än en modell med normalfördelade logaritmiska avkastningar?

3. Korrelation och bivariata copulas. (Använd de två tidsserierna med veckodata)
- a. Bestäm korrelationen (Pearson) mellan tidsserierna med logaritmiska avkastningar. (MATLAB: `corr()`)
 - b. Bestäm autokorrelationen (1-5 veckors lag) i tidsserierna med logaritmiska avkastningar. (MATLAB: `corr()`)
 - c. Använd MLE (Inference for margins) givet de univariata fördelningarna från GARCH(1,1)-modellerna i uppgift 2. Bestäm den copula (Normal, Student-t, Gumbel, Clayton, Frank) som bäst beskriver er data genom att jämföra log-likelihood värden. (MATLAB: `copulafit()`, `copulapdf()`). Generera slumpstal baserat på den copula som bäst beskriver datan och jämför dessa med tidsserierna transformerade till likformiga fördelningar (inverteringsprincipen) utifrån scatter plots (MATLAB: `copularnd()`, `scatter()`).

Numeriska resultat: Korrelation (1), autokorrelationer (2x5), log-likelihood värden (5).

Grafer: Scatter plots (2)

Frågor: Hur skiljer sig era scatter plots för simulerad och verklig data? Uppvisar dessa liknande mönster?