

시계열분석 기말 프로젝트 - 100점 만점.

학번:

이름:

자신이 사용할 자산은 12월 1일 오후 9시이후에 사캠 게시판에 공지 할 것.

보고서에 동일 자산을 사용한 학생들은 최초 공지자를 제외한 나머지 공지자 모두 감점 10.
유사한 보고서는 모두 0점처리.

평가항목: 계산의 정확성, 결과의 흥미로움, 보고서 작성의 완성도.

제출파일: 보고서, 데이터(별도제출), 코드(보고서에 포함시킴)

A. 가격과 변동성지수의 일별자료를 구할 수 있는 어떤 자산 (예: 주가, 환율, 금값, 원유가, 에너지 가격 등등) 을 하나 선택하라. 최소한 10년 이상. 가장 최근 자료 포함되어야함. 선택한 자산명을 사캠에 공지할 것.

A1. 기초통계 분석

1. 일별 가격 P_t , 로그수익률 r_t , 변동성지수 IV_t 를 구해서 시도표를 그려라, 여기서 $t = 1, 2, \dots, T$ 이며 각 t 는 특정 날짜를 나타내는 index 이다.
2. P_t , 로그수익률 r_t , 변동성지수 IV_t 의 기본 통계량 표를 작성하라. 여기에는 이들 자료의 특징 (대표값, 산포, 비대칭성, 첨도, normality 등등) 을 나타내는 통계량들이 포함되어야한다. 이들 통계량값을 통해 자료의 특징을 간단하게 기술하라.
3. P_t , r_t , IV_t 각각에 대해 표본자기상관함수를 lag 을 250 까지 그려보고 그것을 통해 각 계열의 자기상관의 특징을 간단하게 설명하라.

A2. 모형추정, 검진 및 예측

1. P_t , r_t , IV_t 각각에 대해 정상성 여부를 ADF 검정을 이용하여 판단하라.
2. P_t , r_t , IV_t 각각에 대해 ARIMA 모델을 BIC 기준으로 identify 하고 identified 된 모형을 추정하고 추정식을 써라. 추정치의 표준오차도 명시하라.
3. 위 2에서 추정한 모형의 타당성을 검진하라.
4. P_t , r_t , IV_t 각각에 대해 데이터의 마지막 시점 T 에서 일주일간 미래 $T+1, \dots, T+5$ 에서의 미래값을 예측하고 95% 예측구간을 그려라. 이그림은 데이터 시점 $T-22, T-21, \dots, T, T+1, \dots, T+5$ 만 포함되게 한다.

A3. out-of-sample 예측력 비교

r_t 에 대해 AIC order AR 모형과 BIC order AR 모형의 예측력을 비교하라

1. 전체데이터중 앞부분 85% 를 이용하여 AIC order AR 모형과 BIC order AR 모형을 추정하라.
2. 각모형의 최근 15% 데이터의 1-step, 2-step, 3-step, 4-step, 5-step 예측치의 RMSE, MAE, MAPE 비교표를 작성하라. 단 데이터가 하나 추가될때마다 모형의 추정을 갱신한다.

3. 비교 결과를 간단히 설명하라.

A5. r_t 에 대한 GARCH 모형 추정 및 P_t , r_t 의 5% VaR

1. r_t 에 대해 AR(0)+GARCH(1,1) 과 AR(0)+GJR-GARCH(1,1) 모형을 추정하고 추정식을 써라. 정규분포와 표준화 t-분포 사용하라. 총 4개의 모형을 추정하여야 한다.
2. AR(0)+GARCH(1,1) 모형으로부터 계산된 $\hat{\sigma}_t, t=1, \dots, T$ 의 시도표를 그려보고 이를 r_t , IV_t 의 시도표와 비교 설명하여라.
3. $r_{T+1}, r_{T+2}, \dots, r_{T+5}$ 의 조건부 표준편차 $\sigma_{T+1}, \sigma_{T+2}, \dots, \sigma_{T+5}$ 의 예측치를 구하시오. 또 이를 시도표 그림으로 나타내어라. 이그림은 데이터 시점 $t = T-22, T-21, \dots, T, T+1, \dots, T+5$ 에서의 σ_t 만 포함되게 한다.
4. 위 1 의 각 4개의 모형을 이용하여 r_{T+1} 의 1%, 5% VaR 구하고 표로 작성하라. 이 중 어느 모형의 VaR 더 좋은지 어떻게 평가할지를 설명하라. 또 P_{T+1} 의 1%, 5% VaR 구하고 표로 작성하라

A6. 시간 index s 를 연속형으로 정의 하여 $\Delta s = 1$ 을 1년에 대응시키자. 그러면 1년을 250일로 가정하면 위 날을 나타내는 index t 와 년을 나타내는 index s 와는 $s = t/250$ 의 관계가 성립한다. $P_s^* = P_{\frac{t}{250}}^* = P_t, t = 1, \dots, T$ 라하자. 그러면, 예를들면 처음 1년

P_1, \dots, P_{250} 이 $P_{\frac{1}{250}}^*, \dots, P_{\frac{250}{250}}^*$ 에 각각 대응한다. P_s^* 가 다음과 같은 geometric brownian motion $dP_s^*/P_s^* = \mu ds + \sigma dw_s$ 를 따른 다고 가정 하자. 여기서 w_s 는 평균 0 분산 1 인 Wiener process 이다.

1. 위 A1-1 의 데이터를 이용하여 μ, σ 를 추정하라. (단위를 명확히 명시할 것).
2. $p_s^* = \ln P_s^*$ 라 할 때 p_s^* 가 만족하는 모형을 적어라. 모수값도 적어야함.
3. 마지막 시점 T (대응되는 s 값은 $S = T/250$)에서 조건부로 2년후의 미래값 P_{S+2}^* 의 조건부 분포를 구하고 평균 분산을 구하시오.
4. 마지막 시점 T (대응되는 s 값은 $S = T/250$)에서 조건부로 향후 2년간의 미래수익률 $r_{S+2}^* = \ln P_{S+2}^* - \ln P_S^*$ 의 조건부 분포를 구하고 평균 분산을 구하시오.
5. 이 자산의 마지막 가격 P_T 를 쓰시오. 또 만기 6개월, 행사가 P_T 인 European call option 의 공정한 현재가를 구하시오.
6. $Q_s = P_s^{*2}$ 이 Geometric Brownian Motion을 따름을 보이고 이 모형의 mean, variance 의 값을 적으시오.

A7. Simulation for GARCH and GJR-GARCH

1. 위 A5-1에서 추정된 AR(0)+GARCH(1,1) 과 AR(0)+GJR-GARCH(1,1) 모수를 이용하여 추정된 모형을 참모형으로 간주하고 $r_t, t = 1, \dots, n, n = 1000$ 을 simulate 하여라. 그 후 각 모형별로 skewness 를 계산하여라.
2. 위 1을 500 번 반복하여 skewness 의 평균을 구하여라.
3. 위 2 의 결과에 대해 간략한 토론을 하여라.

데이터 사이트 예: FRED(거시경제), worldbank (세계개발및빈곤등), EIA (에너지),
oxford-man realized library(G20 실현변동성, **종합주가지수**),
trend.google.com (구글트렌드), **CBOE 시카고 선물거래소(변동성지수)**

1) 미국 경제 데이터를 제공하는 사이트들을 모아놓은 곳

<https://library.law.yale.edu/news/75-sources-economic-data-statistics-reports-and-commentary>

2) 한국은행 경제통계 시스템

<https://ecos.bok.or.kr/EIndex.jsp>

3) 한국거래소 (**주가, 주가변동성지수**)

<http://marketdata.krx.co.kr/mdi#document=0307>

4) Bureau of Economic Analysis

<https://www.bea.gov/>

5) Yahoo finance (주식)

<https://finance.yahoo.com/>

6) 수출입무역 통계 (한국)

<https://unipass.customs.go.kr/ets/>

7) www.kaggle.com/datasets