카카오 주가 ADF 테스트

23.07.02 최유찬

1. ADF 테스트란?

ADF (Augmented Dickey-Fuller) 테스트는 시계열 데이터의 정상성 여부를 평가하는 데 사용되는 통계적인 방법입니다. 정상성은 시계열 데이터가 시간에 따라 평균과 분산이 일정하며, 시간에 따라 변하지 않는 특성을 가지는 것을 의미합니다. ADF 테스트는 주어진 데이터가 정상성을 가지는지 여부를 판단하는데 도움을 줍니다.

2. 귀무가설과 대립가설

-귀무가설 Null hypothesis: 원본 주식데이터(종가)는 데이터는 비정상성(non-stationarity)을 가진다.

기각에 실패하면 시계열이 비정상성(Non-stationary)임을 의미한다.

-대립가설 Alternative hypothesis : 1차분 주식데이터는 데이터는 정상성(stationarity)을 가진다. 증명하고자 하는 가설. 귀무가설이 기각되고 시계열이 정상성(Stationary)임을 의미한다.

ADF 테스트의 p-value를 통해 귀무가설을 기각하거나 채택한다

보통 p-value가 0.05보다 작을 때, 즉 유의수준 5%에서 귀무가설을 기각하면 데이터는 정상성을 가진다고 판단한다

3. ADF TEST 결과 용어 설명

**ADF 통계량**

이 값은 ADF 테스트에서 사용되는 통계량으로, 정상성을 검정하는 데 도움을 준다.

ADF 통계량이 음수로 나타날수록 데이터가 정상성을 가질 가능성이 높다고 해석할 수 있습니다.

**p-value**

가설이 참일 때 표본에서 얻은 통계량보다 극단적인 값이 관찰될 확률을 나타낸다.

num of lags

ADF 테스트에서 사용되는 시차(lag)의 수를 나타내며, 데이터 간의 자기상관성을 고려하는 데 사용(큰 의미 X)

num of observations

테스트에 사용된 관측치의 수

**Critical values**

테스트 통계량이 이 임계값보다 작으면 데이터는 정성적이라고 볼 수 있다

4. 원본 데이터 ADF TEST 결과:

ADF Statistics: -1.345605

p-value: 0.608058

num of lags: 12.000000

num of observations: 1458.000000

Critical values:

1%: -3.435

5%: -2.864

10%: -2.568

5. 차분 데이터 ADF TEST 결과:

ADF Statistics: -9.479198

p-value: 0.000000

num of lags: 11.000000

num of observations: 1458.000000

Critical values:

1%: -3.435

5%: -2.864

10%: -2.568

결과를 보면, 차분한 데이터의 ADF 테스트 결과는 원본 데이터보다 훨씬 강력한 정상성을 가지는 것으로 나타났습니다. 따라서 차분 데이터를 사용하여 시계열 분석이나 모델링을 수행하는 것이 적절할 수 있습니다.

용어 설명

-정상성과 비정상성

**정상성은 시계열 데이터가 시간에 따라 일정한 패턴과 특성을 유지하는 것을 의미합니다.** 이는 시계열 데이터의 평균과 분산이 일정하게 유지되고, 시간에 상관없이 시계열의 자기상관(autocorrelation)이 존재하지 않는 것을 의미합니다. **정상성을 가진 데이터는 예측이나 추론에 용이하며, 많은 통계적 모델링 기법들은 정상성 가정을 기반으로 합니다.**

반면에 비정상성은 시계열 데이터가 시간에 따라 추세나 주기를 갖거나, 평균과 분산이 시간에 따라 변하는 경우를 의미합니다**. 비정상성을 가진 데이터는 예측이나 추론이 어렵고, 시계열 분석 기법을 적용하기 전에 데이터를 변환하거나 차분 등의 전처리 과정을 거쳐야 합니다.**

-대립가설과 귀무가설

**귀무가설은 일반적으로 기존의 믿어온 가설이나 기존의 상태를 나타내며, 어떤 변화나 차이가 없음을 주장하는 가설입니다.** 즉, 귀무가설은 효과나 연관성이 없다는 가정을 나타냅니다. 통계적으로 귀무가설을 검정하여 통계적 증거를 찾으면, 귀무가설을 기각하고 대립가설을 받아들일 수 있습니다.

**대립가설은 귀무가설과 반대되는 가설로, 연구자가 입증하고자 하는 가설입니다.** 대립가설은 보통 귀무가설의 부정이거나 대립적인 주장을 나타냅니다. 통계적 검정 결과에 따라 귀무가설을 기각하면 대립가설을 채택할 수 있습니다.

통계적 가설 검정은 주어진 데이터를 사용하여 귀무가설의 타당성을 평가하고, 그 결과에 따라 귀무가설을 기각하거나 채택하는 것을 목적으로 합니다.