

SUMMARY

시계열 분석의 의의와 필요성

- 시계열 분석은 불확실한 미래에 대한 통계적 예측을 위해서 사용
- 시계열 분석은 여러 엄격한 가정에 기반하기 때문에 계량분석 기법 중에서도 난도가 높은 편
- 제대로 수행된다면 기업의 의사결정 과정에 강력한 과학적 무기로써의 가치 ↑

간단하면서도 다양한 활용가치를 가진 시계열 기법, 런 검정(Runs test)

- 모멘텀 : 무작위 진행과 대비되는 것으로 시계열 데이터의 변화 양상에 가속 추세가 존재한다는 것을 지칭
- 런 검정(Runs test) : 모멘텀의 존재 여부를 파악하는 통계적 기법
- 런(Run) : 한 종류의 부호(+ 혹은 -)가 시작해 끝날 때까지 런으로 정의

+++++ ----- +++ --- +++++ -----

모멘텀이 존재한다면, 런의 길이는 평균적으로 길어지게 되는 반면 런의 개수는 줄어들

런 검정은 런의 수가 충분히 적으면 양의 계열 상관이,

런의 수가 충분히 많으면 음의 계열 상관이라고 정의

사례) 투수 B의 영입과 관련된 의사결정을 위한 런검정

질문

B 투수의 경기력에 기복이 있는지를 알고 싶다. 투구에 기복이 있다면 스트라이크 혹은 안타가 경기의 흐름에 따라 무리 지어 몰려 있을 공산이 크다.

결론

B 투수의 투구에 양의 계열 상관, 즉 모멘텀이 존재한다고 결론을 내릴 수 있다. 따라서 B 선수의 경기력에 기복이 있다는 의견을 담은 보고서를 구단주에게 제출해야 하는 것이다.

런 검정을 통한 시사점

이러한 계량분석 기법은 여러 가정에 기반하고 있다는 점에서 절대로 만병통치약과 같은 '완벽한 무기'가 돼줄 수는 없다. 하지만 각각의 상황에서 요구되는 가정을 철저하게 이해하고 있는 사람에 의해 절차적 오류 없이 기업 실무에서 제대로 활용된다면 '완벽한 미래 예측'은 아니더라도 '과학적인 추정'은 가능하게 해 줄 것이다


```
#런테스트 결과
runs.test(x)

##
##  Runs Test
##
## data:  x
## Standard Normal = -0.81693, p-value = 0.414
## alternative hypothesis: two.sided

#랜덤하지 않게 동전을 50 번 던진다
#20% 80% 확률
a1<-sample(coin,10,replace=TRUE, prob =c(20,80))
#70% 30% 확률
a2<-sample(coin,10,replace=TRUE, prob =c(70,30))
#90% 10% 확률
a3<-sample(coin,10,replace=TRUE, prob =c(90,10))
#20% 80% 확률
a4<-sample(coin,10,replace=TRUE, prob =c(20,80))
#70% 30% 확률
a5<-sample(coin,10,replace=TRUE, prob =c(70,30))

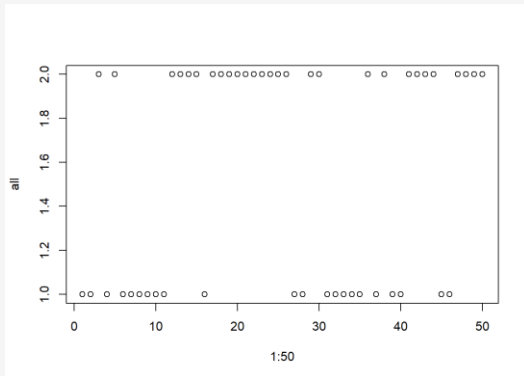
all<-c(a1,a2,a3,a4,a5)
all<-factor(all)

#최종결과
print(all)

##  [1] 뒤 뒤 앞 뒤 앞 뒤 뒤 뒤 뒤 뒤 앞 앞 앞 앞 뒤 앞 앞 앞 앞 앞 앞 앞
## [24] 앞 앞 앞 뒤 뒤 앞 앞 뒤 뒤 뒤 뒤 뒤 앞 뒤 앞 뒤 뒤 앞 앞 앞 앞 뒤 뒤
## [47] 앞 앞 앞 앞
## Levels: 뒤 앞
```

시각화

```
plot(1:50, all)
```



런테스트 결과

```
runs.test(all)
```

```
##
```

```
## Runs Test
```

```
##
```

```
## data: all
```

```
## Standard Normal = -2.2159, p-value = 0.0267
```

```
## alternative hypothesis: two.sided
```

WARD

횡단적, 종단적

횡단조사	종단조사
표본조사	현지조사
일정한 시점에서 특정 표본이 갖고 있는 특성을 파악	시간의 흐름에 따른 조사대상이나 상황의 변화를 측정하여 의미를 찾고자 함
정태적 성격	동태적 성격

패널데이터(panel data)

패널데이터(panel data)는 종단자료(longitudinal data)라고도 하며, 여러 개체들을 복수의 시간에 걸쳐서 추적하여 얻는 데이터를 말한다

- 위키백과 -

-

시계열(time series)

시계열(時系列, time series)은 일정 시간 간격으로 배치된 데이터들의 수열을 말한다.

- 위키백과 -

귀무가설(영가설 Null hypothesis)

설정된 가설이 진실할 확률이 극히 적어 처음부터 버릴 것이 예상되는 가설. 검정을 할 때 비교되는 2개의 표본집단(標本集團)의 결과차(結果差)가 확실히 조건차(條件差)가 있다고 생각되는 경우 또는 동일 모집단(母集團)에 귀속하고 있지 않다고 생각되는 경우 그 추측과는 반대의 가설을 설정하는 것이다.

- 두산백과 -

-

선물(Futures)

선물(先物)은 파생상품의 한 종류로서 품질, 수량, 규격 등이 표준화되어 있는 상품 또는 금융자산을 미리 결정된 가격으로 미래 일정시점에 인도·인수할 것을 약정한 거래를 말한다. 이러한 표준화된 상품, 자산은 정해진 시장을 통해서만 거래된다

- 두산백과 -

ISSUE

1. 런검정이 경영을 하거나 의사결정을 할 때 어떠한 상황에서 유용하게 쓰일 수 있을까요? 그리고 런검정은 어떠한 단점을 가지고 있을까요? 런검정의 장점과, 기사에서 언급되지 못한 단점까지 이야기해봅시다.
2. 런검정과 같은 통계기법들은 여러가지가 있지만, 적재적소에 적합한 기법을 쓰지 않으면 무용지물이 될 뿐만 아니라 잘못된 의사결정까지 할 수 있습니다. 런검정 뿐만 아니라 다른 통계기법까지 통틀어서 잘못된 통계기법을 적용하여, 낭패를 본 사례를 이야기해봅시다.

(수준 높은 통계기법이 아니라, 그래프와 히스토그램과 같은 단순한 통계치를 이용한 사례 또한 괜찮으니 여러 사례를 생각해 보아요~^^)