이동평균(Moving Average) ← '시계열 데이터'의 평균

평균이란 개념은 초등학교 때부터 들어와서 누구나 쉽게 인지하고 있습니다.

가령, 5명의 평균 나이를 계산하라고 하면 5명의 나이 합을 5로 나누어서 구합니다.

← 시점.

이동평균이라는 것은 5명 대신 "5일"로 바뀌었을 뿐 산출하는 방법은 동일합니다.

따라서 평균과 이동평균을 그냥 동일 개념으로 생각하고 넘어가기도 합니다.

> 그럼 그냥 평균이라고 쓰면 되는데 왜 굳이 이동평균이라고 사용할까요?

평균과 이동평균의 가장 큰 차이점은 **시간**이라는 개념입니다.

앞절에서 리스크를 이야기할 때도 **시간**이라는 개념을 언급했습니다.

학자들을 위험선호이기 때문에 기대값이 음수임에도 불구하고 로또를 구매한다고 이야기 합니다.

하지만 저의 생각은 다릅니다. 로또를 통해 얻을 수 있는 금전적 기대값은 비록 음수이지만, 시간적 기대값은 양수입니다.

투자에 있어서 의외로 시간의 개념은 너무 과소 평가되고 있습니다.

다시 평균과 이동평균에 대해서 이야기 하겠습니다.

← 핵심!!!

평균은 동일시점에서 산출되는 것이 흔한 반면, 이동평균은 동일 대상이지만 시점이 서로 상이해서 발생합니다.

본체는 동일하나 시점이 다릅니다.

> 어제의 나와 오늘의 나는 동일한가?

비록 같은 나이지만, 어제의 나와 오늘의 나는 서로 동일하다고는 할 수 없습니다.

2002년 16강 월드컵 때 패널티 킥을 실축한 시점에서 안정환과 연장전에서 역전 결승골을 넣은 안정환은 동일 인물이지만, 불과 몇시간만에 사람들에게서 평가가 전혀 달라집니다.

고작 2시간 남짓밖에 안되는 시간이지만, 이 때 안정환에 대한 팬들의 기대치를 평가하면 매우 큰 변화가 나타날 것입니다.

```
1. 패널티 킥 진전의 안정환

2. 패널티 킥 실축 후 안정환

3. 역전 골 후 안정환
```

3시간도 안되는 상황에서 선수에 대한 기대치는 급변하는 것을 느낄 수 있습니다.

8강 경기가 시작될 때 안정환 선수에 대한 팬들의 기대치는 어느 정도 될까요?

이를 추정하려고 할 때,

각 시점의 기대치를 단순 합해서 나누는 방법을 사용할 수도 있습니다.

가장 마지막 기대치를 통해 측정할 수도 있습니다.

가장 최근 기대치에 가장 많은 가중치를 주어서 측정할 수도 있습니다.

평균은 동일 시점에서 발생하기 때문에 단순 평균이 많이 사용됩니다.

하지만 시간시점이 다를 경우 최근의 정보가 더 많은 영향력을 미칠 수 있기 때문에

최근에 단순 평균보다 최근 정보에 가중치를 더 주는 방법을 사용할 수도 있습니다.

그 결과 이동평균을 구하는 방법엔

① 단순이동평균(Simple Moving Average)

② 가중이동평균(Weighted Moving Average)

③ 지수이동평균(Exponential Moving Average)

방법이 존재합니다.

① 단순이동평균(Simple Moving Average)

일반적인 평균을 구하는 방법입니다.

n1 , n2, n3 가 존재한다고 가정하면, (n1 + n2 + n3) / 3 입니다.

② 가중이동평균(Weighted Moving Average)

n1 , n2, n3 가 존재한다고 가정하고 각각의 가중치를 w1, w2, w3라고 하면

(w1n1 + w2n2 + w3*n3) / (w1 + w2 + w3) 입니다. `가중치의 합'으로 나눈다.

일반적으로 최근일에 높은 가중치를 줍니다.

상황에 따라서 이벤특 발생한 특정일에 높은 가중치를 주어 사용할 수도 있습니다.

최근일이 높은 가중치를 무조건 줄 필요는 없다.

③ 지수이동평균(Exponential Moving Average)

최근에 높은 가중치를 주지만, 오래된 과거도 비록 낮은 영향력이지만 가중치를 두여하도록 고려한 방법입니다.

EMV(t) = (1-w) * EMV(t-1) + w * Price(t)

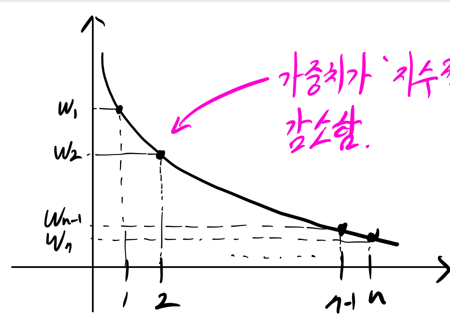
- $0 < w_i < 1$
- $\sum w_i = 1$

오늘의 종가에 w의 가중치를 주고 (1-w)를 어제의 이동평균에 주는 방식입니다.

EMV(t) = (1-w) * ((1-w) * EMV(t-2) + w * Price(t-1)) + w * Price(t)

```
 = (1-w)^2 * EMV(t-2) + (1-w )*w* Price(t-1)) + w * Price(t-1))
  ...
 = (1-w)^n * EMV(0) + (1-w)^(n-1)*w*Price(1) + ... + w * Price(t-1)
```

결국 과거의 값도 지속적으로 영향력이 잔존하게 됩니다.



가중치가 `지수적으로' 감소함.

w = 2 / (n+1)

지수이동평균 2일인 경우

w = 2/3가 됩니다.

지수이동평균 EMA

$EMA = P_{tod} \times K + EMA_{yest} \times (1-K)$

↳ P_{tod} = 당일 주가

$K = \frac{2}{N+1}$

EMA_{yest} = 전일 EMA

N = EMA 산출기간

날짜	4/16	4/17	4/18	4/19	4/20
주가	1000	1010	1000	1020	1300

4/16 ⇒ 1000 ※ 당연히 하루치의 평균값이 당일값

4/17 ⇒ $(1010 \times 0.66) + (1000 \times 0.33) \fallingdotseq 990$

4/18 ⇒ $(1000 \times 0.5) + (990 \times 0.5) \fallingdotseq 995$

4/19 ⇒ $(1020 \times 0.4) + (995 \times 0.6) \fallingdotseq 1005$

4/20 ⇒ $(1300 \times 0.33) + (1005 \times 0.66) \fallingdotseq 1092.3$

⇒ 각 당일 EMA

4/17 $k = \frac{2}{2+1} \fallingdotseq 0.66$ $(1-k) \fallingdotseq 0.33$

4/18 $k = \frac{2}{3+1} = 0.5$ $(1-k) = 0.5$

4/19 $k = \frac{2}{4+1} = 0.4$ $(1-k) = 0.6$

4/20 $k = \frac{2}{5+1} \fallingdotseq 0.33$ $(1-k) \fallingdotseq 0.66$