프로야구 타자 연봉

독립변수: 수비점수와 세이버 메트릭스 주성분 3개, 연차 (2019년 기준), FA 유무

종속변수: 연봉 (2020년 기준)

분석대상: 2019년 KBO 프로야구 타자 중 144타석 이상을 소화한 타자

1. 수비점수

- 수비점수 계산법: 수비점수 = 수비율 \* 수비이닝 \* 포지션 가중치

- 수비율(FPCT): (자살+보살) / (자살+보살+실책)

1-1. kWAR 포지션별 가중치 반영

* 한국 프로야구 WAR 계산시 활용되는 포지션 조정점수를 활용(포지션 조정 점수는 한 시즌 모든 경기 모든 이닝에서 해당 포지션으로 수비를 하였을 경우 주어지는 점수).
* 음(-)의 값을 없애기 위해 표준화(MinMaxScaler) 사용(수비이닝에 곱할 예정인데 음의 값이면 수비 하는것 자체가 마이너스가 되기 때문에).

1-2. 자체 제작 포지션별 가중치 반영

* 센터라인(포수, 유격수, 2루수, 중견수)에 더 가중치.
* 외야수보다 내야수에 더 가중치 (1루수 제외).
* 포수, 유격수, 2루수, 3루수, 우익수, 중견수, 좌익수, 1루수 순으로 가중치 결정.

1-3. AHP 분석 포지션별 가중치 반영

* 포수>유격수>2루수=3루수>중견수>좌익수=우익수>1루수 순으로 서열 측정.
* 차이는 표준화된 kWAR Scale에 10을 곱한 값을 활용.

테이블이(가) 표시된 사진

자동 생성된 설명

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 포지션 | 포지션 조정 점수 (KBO) | | | |
| kWAR Scale | 표준화된 kWAR Scale | 자체제작 가중치 | AHP 분석 계수 |
| 포수 | 1 | 1.0 | 8 | 0.370 |
| 1루수 | -1 | 0 | 2 | 0.022 |
| 2루수 | 0.2 | 0.6 | 6 | 0.124 |
| 3루수 | 0.2 | 0.6 | 4 | 0.124 |
| 유격수 | 0.6 | 0.8 | 7 | 0.219 |
| 좌익수 | -0.6 | 0.2 | 2 | 0.036 |
| 중견수 | -0.2 | 0.4 | 3.5 | 0.067 |
| 우익수 | -0.6 | 0.2 | 2.5 | 0.036 |
| 지명타자 | -1.4 | 0 | 0 | 0 |

<그림 1> 포지션 조정 점수 출처: KBReport

2. 세이버 메트릭스 지표

- 세이버 메트릭스 지표 중 공개된 기록으로 계산 가능한 14개 지표 활용

- 출루율, 장타율, OPS, GPA, SECA, TA, RC, RC/27, XR, ISO, PSN, wOBA, BABIP, OW% 총 14개 활용

- 세이버 메트릭스 지수 계산식에 같은 타격지표가 다수 포함되어 있으므로 지표간 상관관계가 높아 다중공선성 문제 발생 -> 주성분분석으로 해결하고자 함.

3. 주성분 분석

- 세이버 메트릭스 지표 14개와 수비점수 3개를 주성분 분석

- 주성분분석 결과: 2개의 주성분 도출

= 주성분 2개의 설명력 각각 57.59%, 19.37% 총 76.96%

= 주성분 2개의 고유값 각각 9.8851, 3.3240로 모두 1을 넘기 때문에 2개 채택

- 주성분 1 (P1): 전반적인 세이버메트릭스 지표와 수비력 모두를 반영 -> 일반적인 지표

- 주성분 2 (P2): 수비지표 3개를 가장 많이 반영 -> 수비력

4. 연차

- 프로야구 지명연도 또는 프로야구 입단연도 기준

- 2019년도 기준 몇 연차인지를 반영

5. FA 유무 (더미변수)

- 2019년 기준 FA계약기간 내에 속하는 지 여부를 두고 속하면 1, 그렇지 않으면 0으로 반영

6. 연봉

- 연봉에 자연로그를 취해줌으로써 데이터 안정성 확보

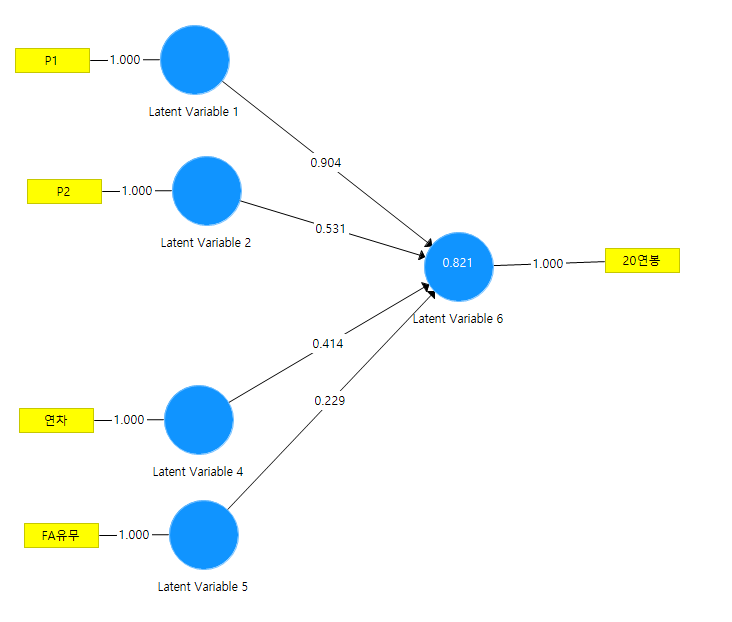
7. 다중회귀분석

- 독립변수: 주성분 3개, 연차, FA유무

- 종속변수: 20연봉(Log)

|  |  |
| --- | --- |
| R2 | 0.8097 |
| Adjusted R2 | 0.8021 |
| MSE | 0.1963 |
| RMSE | 0.4431 |

- 결과

8. PLS

테이블이(가) 표시된 사진

자동 생성된 설명