

2019년 KBO

## 외국인 투수 스카우팅 대회

Statcast 데이터 중심의 구종 개수와  
아웃 확률에 기반한 투수 평가

안성진

# 목차

- 대회 소개
- 해결 과정
- 참가 후기

# 대회 소개

- 대회 목적: 2011년 부터 2018년까지 KBO에서 활약한 외국인 투수의 데이터를 바탕으로, 2019년 새로 들어온 외국인 투수 중 2명 스카우트

# 대회 소개

**2011~2018 KBO에서 활약한 외국인 투수 57명 : 그룹 A**



**2019 신입 외국인 투수 13명: 그룹 B**

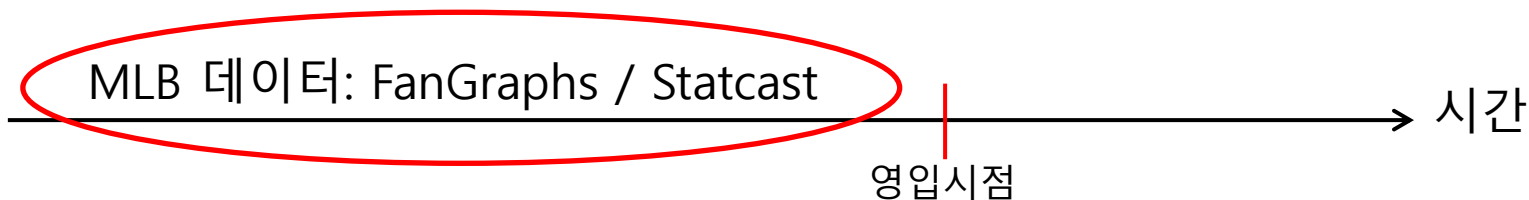


# 대회 소개

**2011~2018 KBO에서 활약한 외국인 투수 57명 : 그룹 A**



**2019 신입 외국인 투수 13명: 그룹 B**



# 해결 과정 – 스카우팅의 목적

- 스카우팅

- 국내에서 좋은 활약을 할 수 있는 선수를 찾는 것

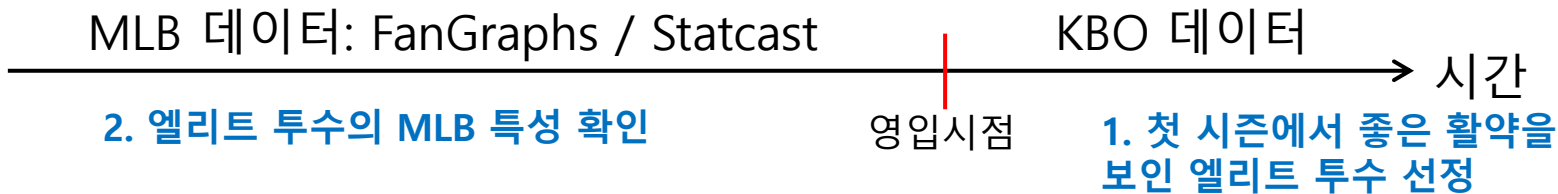
- 외국에서 잘한다고 해서, 국내에서 또한 잘할 것이라고 단정 지을 순 없다

- 첫 시즌에 좋은 활약을 할 수 있는 선수를 영입

- 육성의 차원보다는 영입 직후, 팀 성적 향상에 기여할 수 있는 선수를 찾는 것

# 해결 과정 – 접근 방법

**2011~2018 KBO에서 활약한 외국인 투수 57명 : 그룹 A**



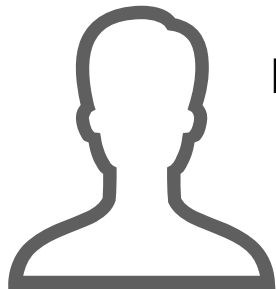
**2019 신입 외국인 투수 13명: 그룹 B**



# 해결 과정 – “좋은 활약” 정의

1. 첫 시즌에서 좋은 활약을 보인 엘리트 투수 선정

투수 A



ERA: 2.45

투수 B



ERA: 2.00



# 해결 과정 - “좋은 활약” 정의

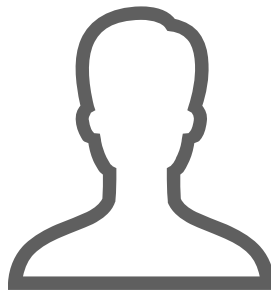
## 1. 첫 시즌에서 좋은 활약을 보인 엘리트 투수 선정

투수 A



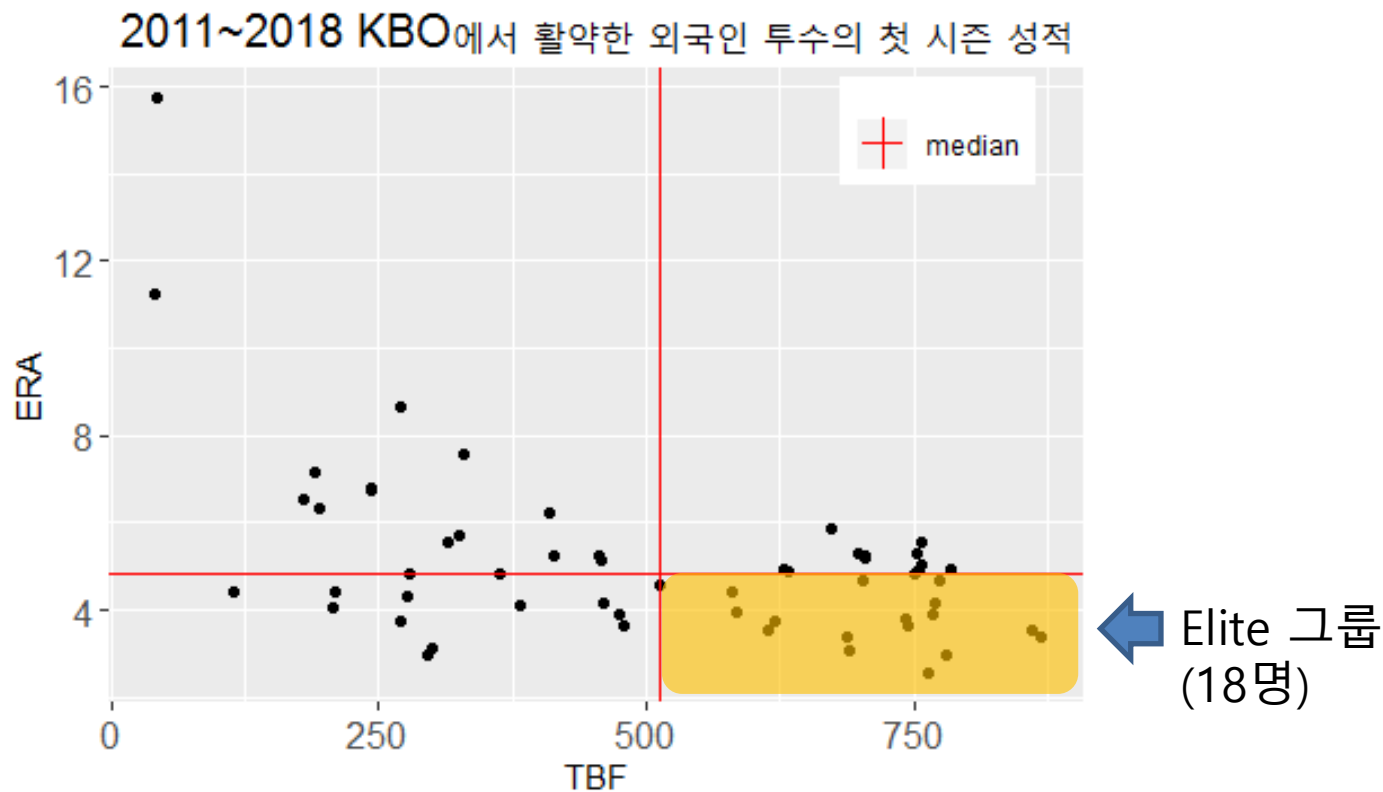
ERA: 2.45  
TBF: 100

투수 B



ERA: 2.00  
TBF: 20

# 해결 과정 - “좋은 활약” 정의



# 해결 과정 – MLB에서의 특성 확인

## 2. 엘리트 투수의 MLB 특성 확인

HR?

BB?

ERA?

H?

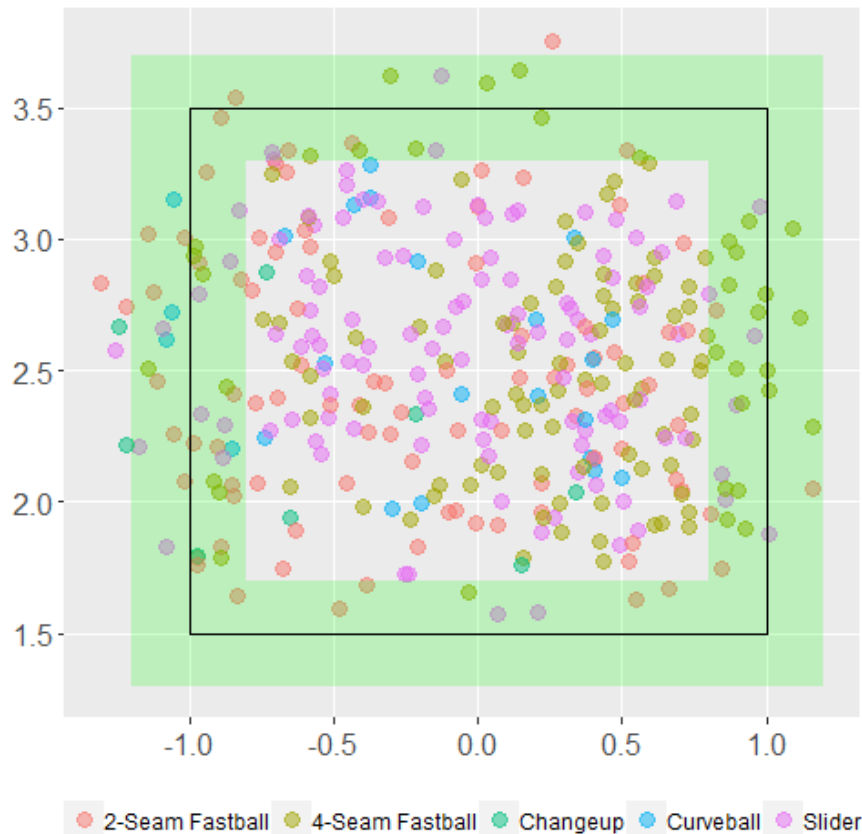
# 해결 과정 – MLB에서의 특성 확인

## 2. 엘리트 투수의 MLB 특성 확인



가설: MLB에서 많은 구종을 구사할 수록, KBO에서의 ERA가 낮을 것이다

# 해결 과정 - 구종 정의



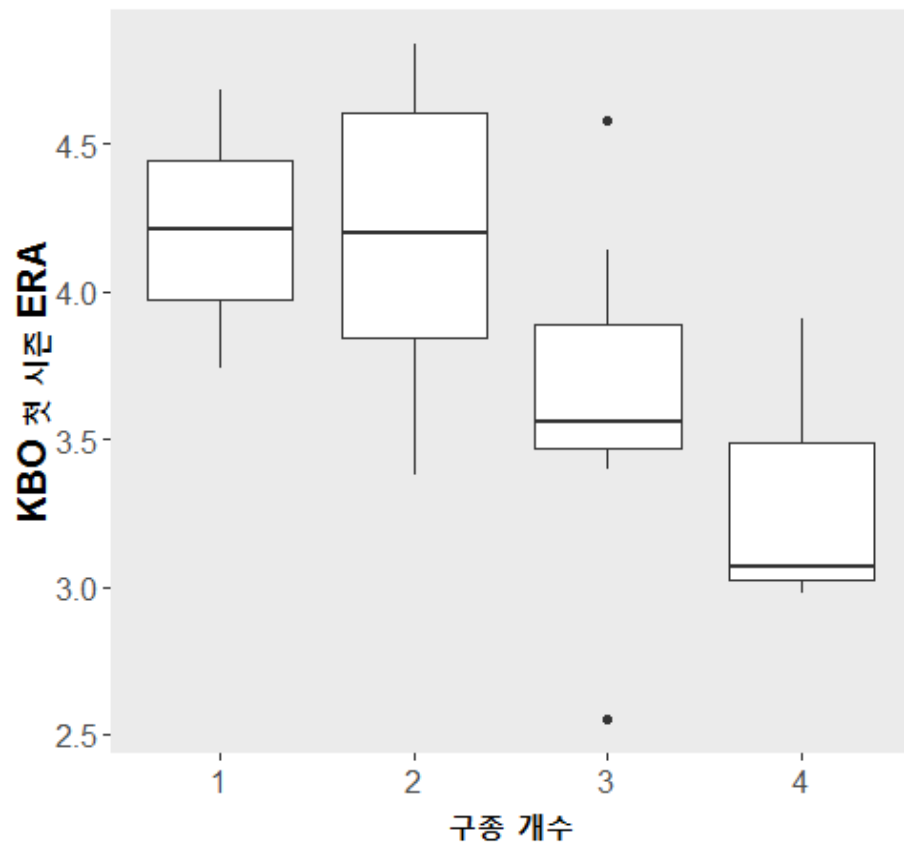
기간: 2011/06/01 ~ 2014/04/02

유형: called\_strike

구종명	구종 개수	구종 비중
투심	31	0.337
포심	35	0.380
체인지업	1	0.011
커브볼	4	0.044
슬라이더	21	0.228

**린드블럼 구종 = 3개**

## 해결 과정 - 박스플롯 / 회귀분석



Coefficients:

	Estimate	Std. Error	t value	Pr(> t )	
(Intercept)	4.7731	0.4135	11.543	3.6e-09	***
numberOfPitches	-0.3629	0.1499	-2.421	0.0277	*

# 해결 과정 – 19년 외국인 투수 스카우트

## 3. 인사이트 적용

구종 개수

4개

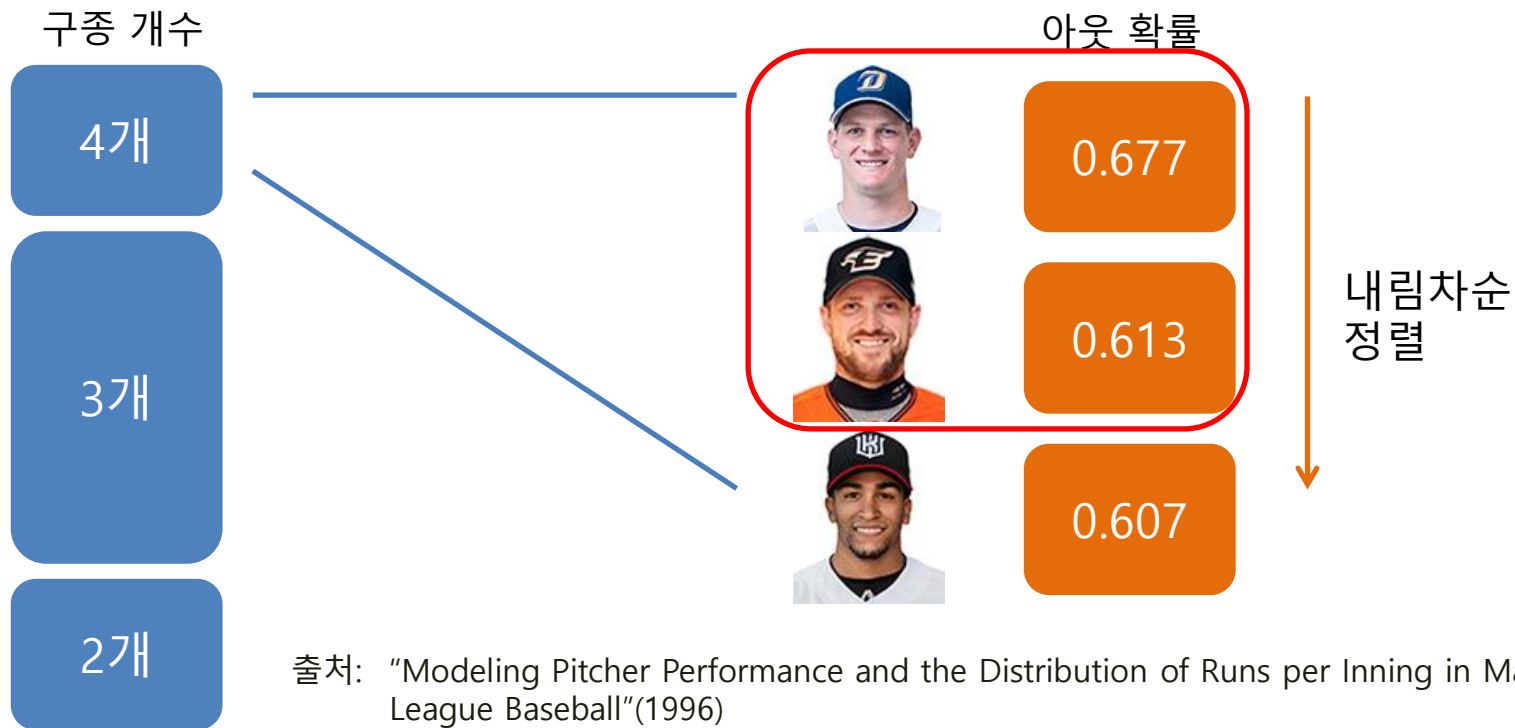
3개

2개



# 해결 과정 - 19년 외국인 투수 스카우트

## 3. 인사이트 적용





# 아웃 확률이란

이항분포

$$X \sim B(n, p)$$

동전 던지기에서

3번 동전을 던졌을 때 나온 앞면의 수

$$= X \sim B(3, 0.5) \quad (0 \leq X \leq 3)$$

n: 시행 횟수 (고정)

X: n번의 시행 횟수 내에서,  
해당 사건이 발생한 건수

음이항분포

$$X \sim NB(k, p)$$

동전 던지기에서

3번의 앞면이 나오기 위해서 동전을  
던진 횟수

$$= X \sim NB(3, 0.5) \quad (3 \leq X)$$

k: 성공 횟수 (고정)

X: 해당 사건이 k번 발생하기까지  
시행한 횟수

# 아웃 확률이란

3번의 아웃이 나오기  
위해서 상대한 타자의 수



음이항분포

$$X \sim \text{NB}(k, p)$$

동전 던지기에서

3번의 앞면이 나오기 위해서 동전을  
던진 횟수

$$= X \sim \text{NB}(3, 0.5) \quad (3 \leq X)$$

k: 성공 횟수 (고정)

X: 해당 사건이 k번 발생하기 까지  
시행한 횟수

# 아웃 확률이란

$$X \sim \text{NB}(3, p)$$

X: 3번의 아웃이 나오기  
위해서 상대한 타자의 수

➡ 이닝이 끝날 때 까지  
상대한 타자의 수

p: 아웃 확률

➡ 주어진 데이터를 통해 추정

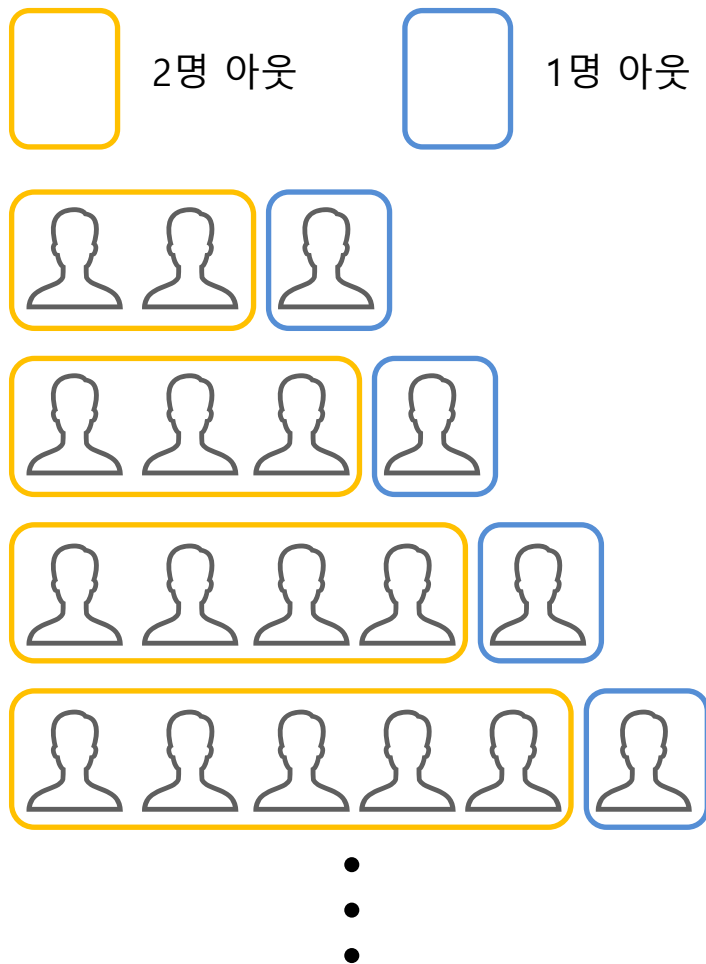
$$x = 3$$

$$x = 4$$

$$x = 5$$

$$x = 6$$

⋮



# 아웃 확률이란

$$X \sim \text{NB}(3, p)$$

X: 3번의 아웃이 나오기  
위해서 상대한 타자의 수

p: 아웃 확률



주어진 데이터를 통해 추정

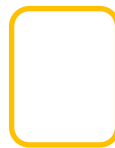
$$X = 3$$

$$X = 4$$

$$X = 5$$

$$X = 6$$

⋮



2명 아웃



1명 아웃



⋮

# 스카우팅 선수



선수명: 채드벨

MLB 구종 개수: 4개

MLB 아웃 확률: 0.613

2019 상반기 ERA: 3.97 (13명의 신입 외국인 선수 중 5위)



선수명: 루친스키

MLB 구종개수: 4개

MLB 아웃 확률: 0.677

2019 상반기 ERA: 2.58 (13명의 신입 외국인 선수 중 1위)

# 참가 후기



감사합니다