

## 1 Introduction

?: Estimates home run rates for 12 players. Modeled the aging curves as a quadratic function of the number of seasons that a player had played.

?: Baseball abstract. We should get a copy of this.

?: Bridging Different Eras in Sports. Use a nonparametric method for estimating aging curves.

?: This is a comment about ?, where Albert again mentions the quadratic aging curve. I think Albert like the interpreting of a quadratic better than the nonparametric curve.

?: Looks at effects of aging in swimming, running, and chess. Mentions Moore1975 (let's try to get this.)

?: JQAS article about aging in baseball.

## 2 Methods

- Describe what we are doing. - Trying to estimate an aging curve using imputation methods to impute the missing years of players' careers due to retirement, drop out, etc.

### 2.1 discrete

Ideas: - Raw aging curve (OPS as measure of performance) imputed with 2L.norm (and covariates) - Delta method of the aging curve (OPS as measure of performance) imputed with 2L.norm (and covariates) - Aging curves for different aspects of performance (i.e. Power hitting, OBP, stolen bases?)

- Think about HOW we are actually doing these imputations. We might need to write our own code to do something more complex than MICE??? - MICE MD? `mice.impute.2l.2stage.pmm`? `mice.impute.2l.2stage.norm`?

### 2.2 continuous

- A different idea is to take the discrete performance by age and try to fit a continuous curve (Fourier approx or loess or splines or etc.) to that data and then impute that function. - Do the "delta method" to the continuous fitted curve.

- clustering based on career types.

## 3 Results

### 3.1 Simulation results

simulate a "real" curve. We simulate dropout and then check how different types of dropout

### 3.2 Real data example

Baseball: Different positions?

Other sports? Tennis Chess Golf Running Swimming Softball

CLustering based on career types?

## 4 Conclusions

The old way of doing aging curves did not account for dropout. We believe our estimate is better. And here are the ways that it is different. He et al. (2011)

## References

- Albert, J. (1992). A bayesian analysis of a poisson random effects model for home run hitters. *The American Statistician* 46(4), 246–253.
- He, Y., R. Yucel, and T. Raghunathan (2011). A functional multiple imputation approach to incomplete longitudinal data. *Statistics in Medicine* 30(10), 1137–1156.