1 Introduction

Lit review

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.