

STAT 532
Stein's Paradox in Statistics
10/27/17

Name: _____

Read Brad Efron and Carl Morris's *Stein's Paradox in Statistics*. Note your responses to these questions may be turned in at the end of the class.

1. Import the data used in Efron and Morris's paper, it can be found on the course website at: <http://www.math.montana.edu/ahoegh/teaching/stat532/data/SteinData.csv>. Fit the James-Stein estimator on this data (using avg 45). Note: $c = 1 - \frac{(k-3)\sigma^2}{\sum(y-\bar{y})^2}$ and that the authors estimate $\sigma^2 = \frac{\hat{p}(1-\hat{p})}{45}$, where $\hat{p} = \frac{1}{18} \sum y = \bar{y}$.
 - (a) Compute the MSE between the estimated and season ending averages for the James-Stein estimator and y .

 - (b) Summarize your results.

2. Create a simulation study to mimic this scenario.
 - (a) Generate a set of 18 baseball players, each with some "true batting average" (typically between .150 and .350).

 - (b) For each batter, give them 45 at bats and record the batting average.

 - (c) Compute the MSE between the estimated and season ending averages for the James-Stein estimator and the observed average.

 - (d) Repeat this entire procedure 1000 times and record the proportion of simulations where the James-Stein estimator is better.

 - (e) Summarize your results.