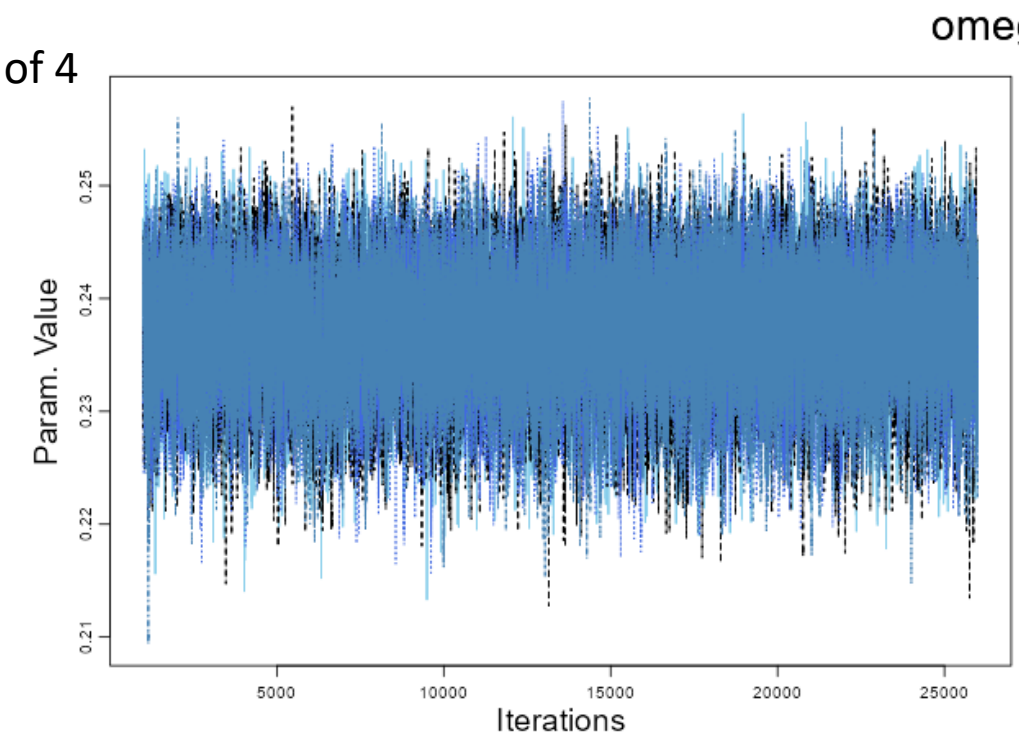
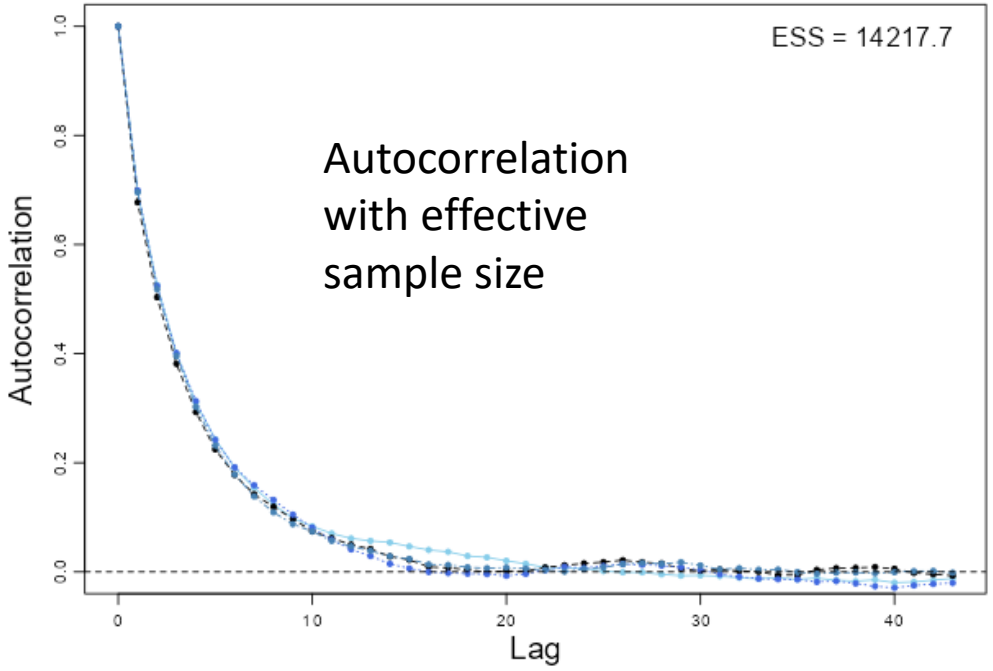
Examples from the Bayesian tab

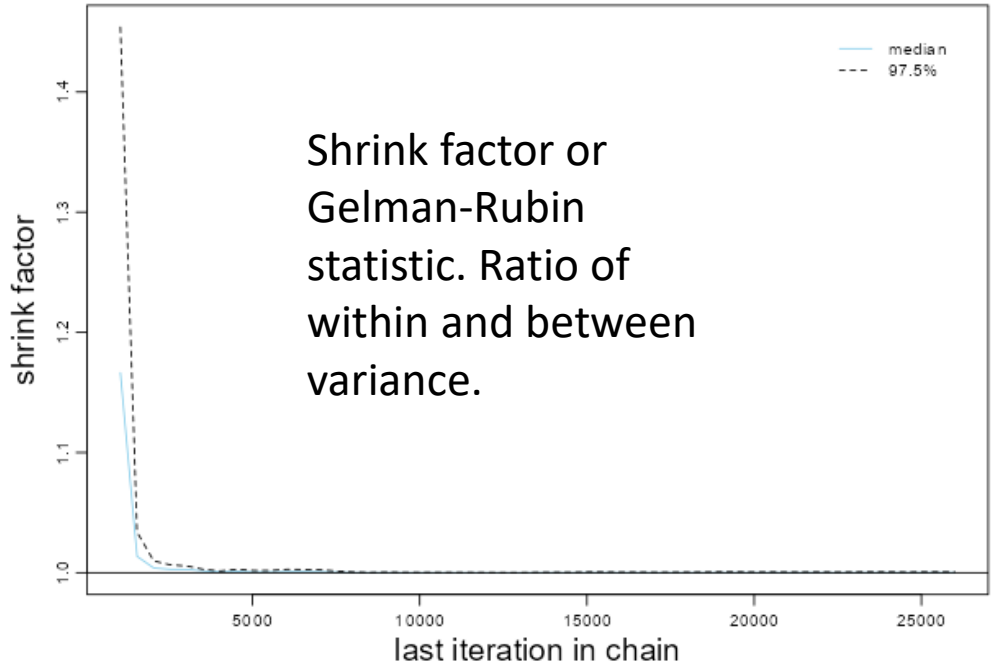
Graphs and output

Trace plots of 4 chains

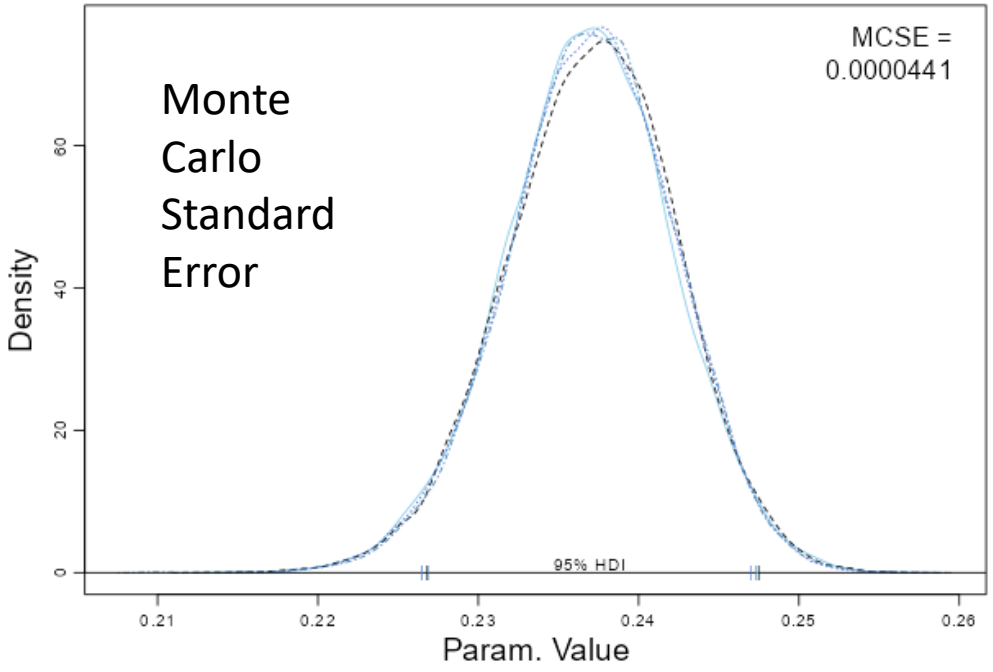




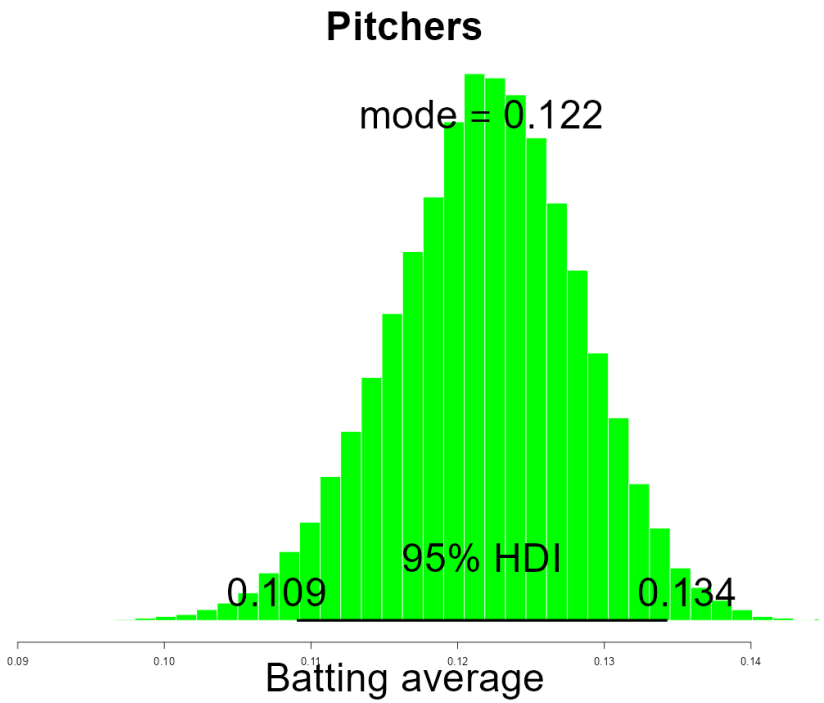
Autocorrelation with effective sample size

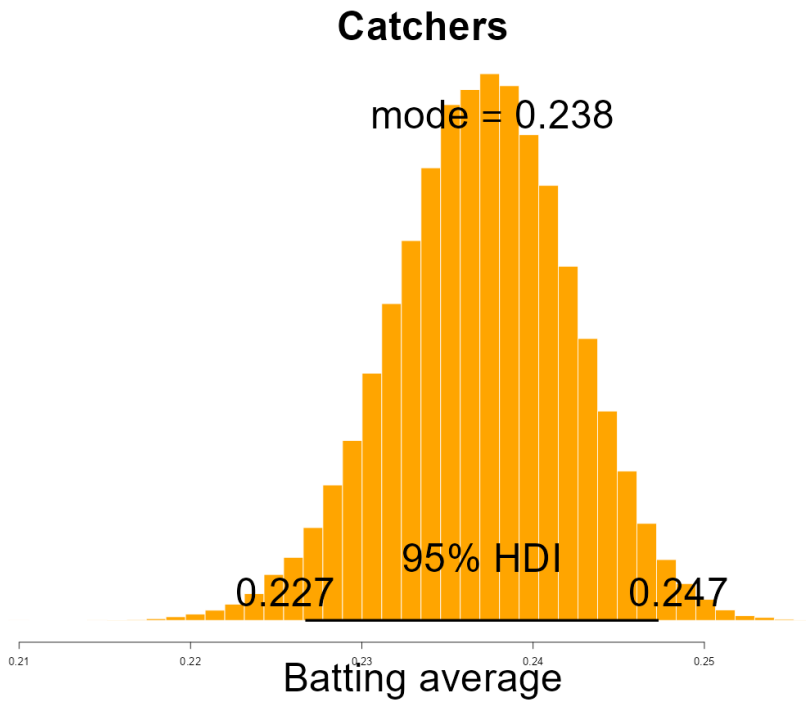


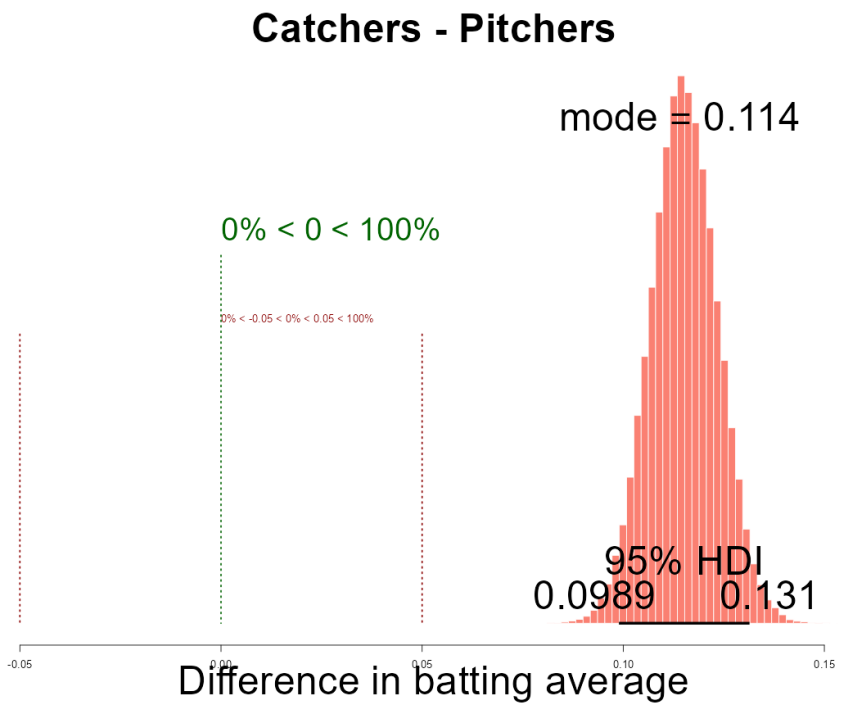
Shrink factor or Gelman-Rubin statistic. Ratio of within and between variance.



Monte Carlo Standard Error

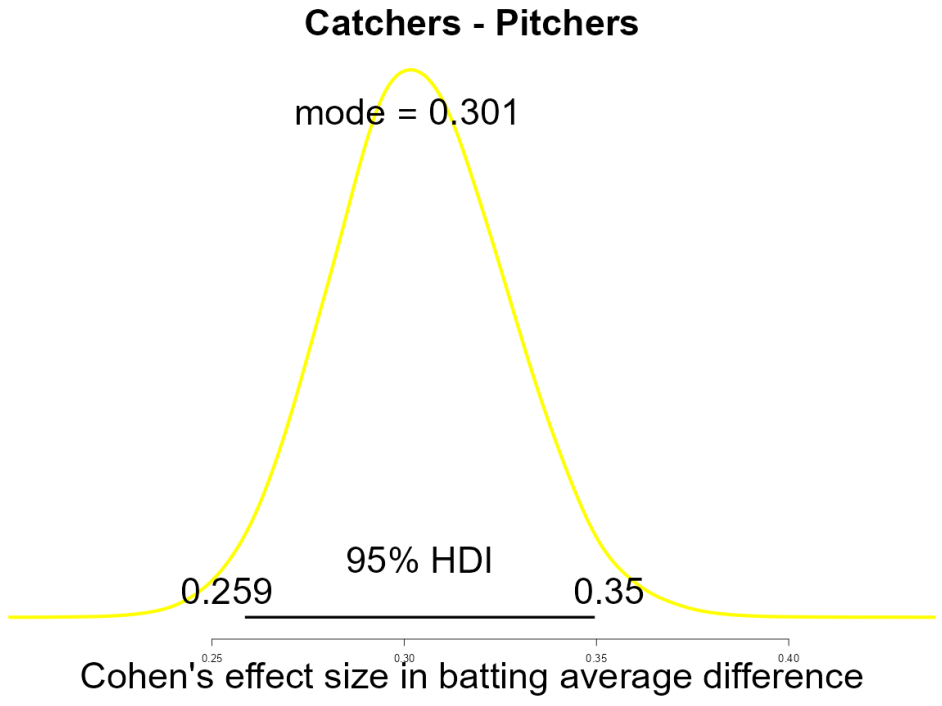


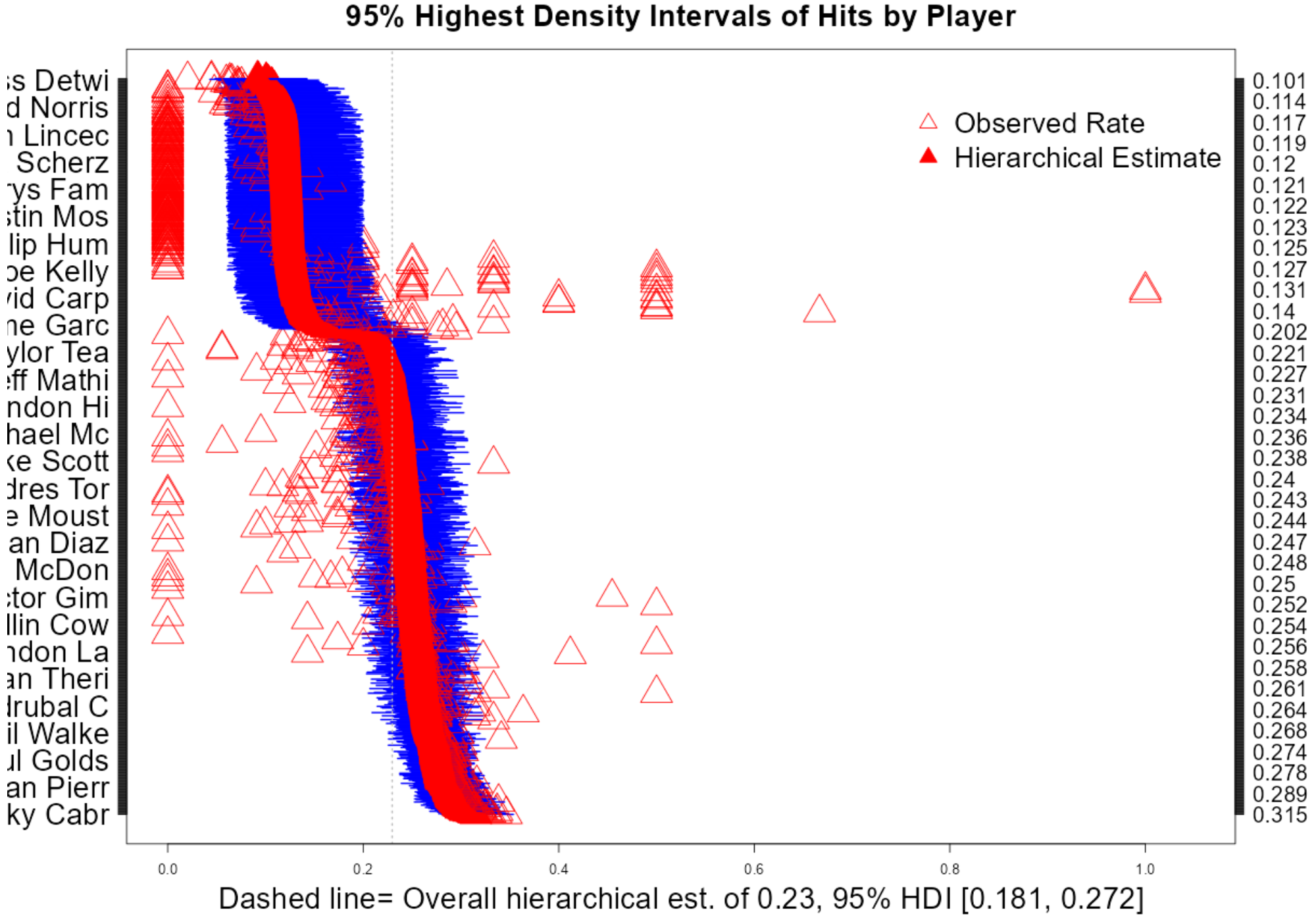


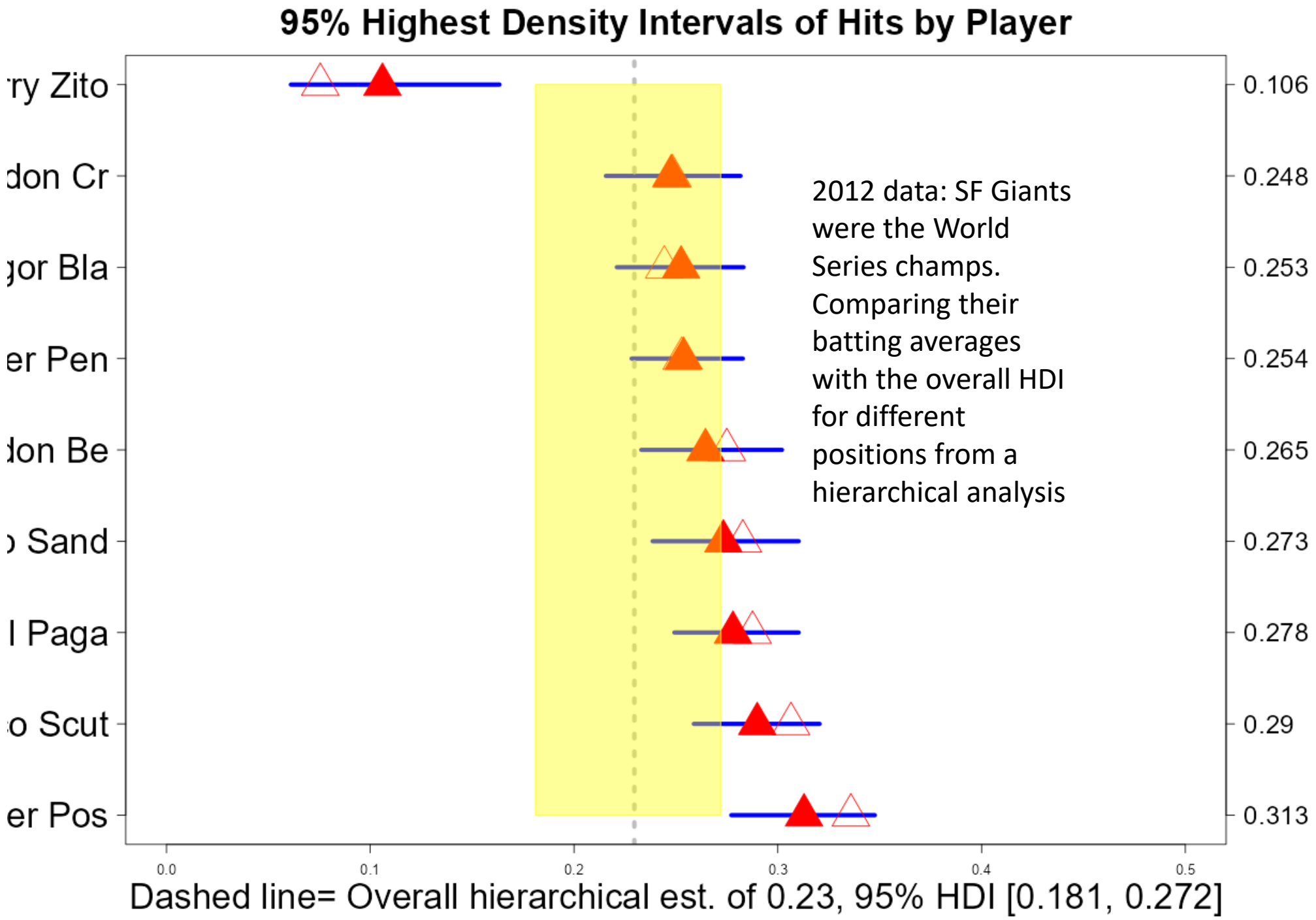


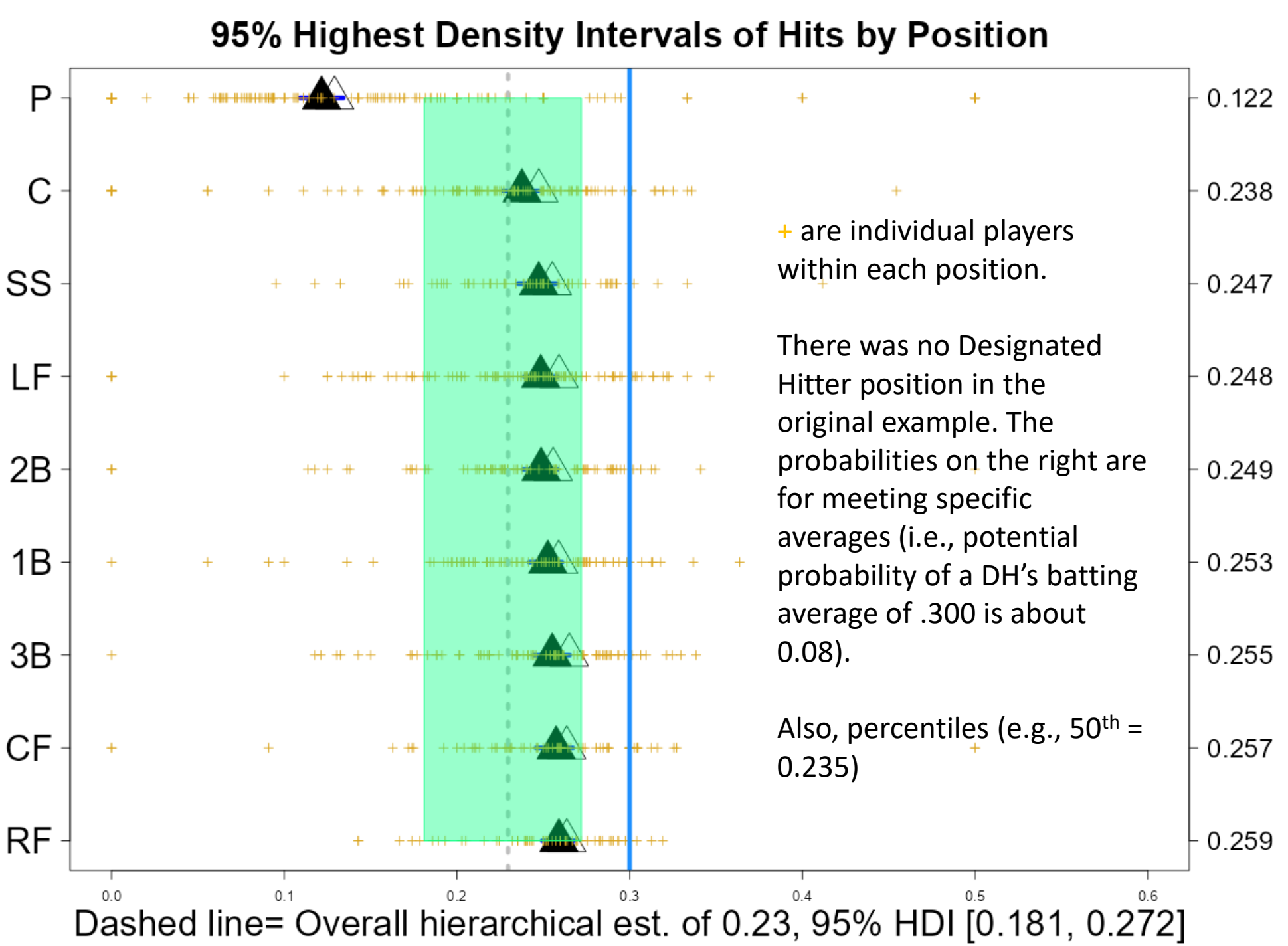
Difference in batting averages excludes 0 as a credible value and the ROPE excludes the 95% HDI.

Cohen's effect size for th difference




95% Highest Density Intervals of Hits by Player
s Detwi
d Norris
n Lincec
Scherz
ys Fam
stin Mos
lip Hum
oe Kelly
vid Carp
ne Garc
ylor Tea
eff Mathi
ndon Hi
hael Mc
ke Scott
dres Tor
e Moust
an Diaz
McDon
tor Gim
llin Cow
ndon La
an Theri
Irubal C
il Walke
ul Golds
an Pierr
ky Cabr
Observed Rate
Hierarchical Estimate
0.101
0.114
0.117
0.119
0.12
0.121
0.122
0.123
0.125
0.127
0.131
0.14
0.202
0.221
0.227
0.231
0.234
0.236
0.238
0.24
0.243
0.244
0.247
0.248
0.25
0.252
0.254
0.256
0.258
0.261
0.264
0.268
0.274
0.278
0.289
0.315
0.0
0.2
0.4
0.6
0.8
1.0
Dashed line= Overall hierarchical est. of 0.23, 95% HDI [0.181, 0.272]


95% Highest Density Intervals of Hits by Player
ry Zito
don Cr
gor Bla
er Pen
don Be
o Sand
l Paga
o Scut
er Pos
0.106
0.248
0.253
0.254
0.265
0.273
0.278
0.29
0.313
0.0
0.1
0.2
0.3
0.4
0.5
2012 data: SF Giants were the World Series champs. Comparing their batting averages with the overall HDI for different positions from a hierarchical analysis
Dashed line= Overall hierarchical est. of 0.23, 95% HDI [0.181, 0.272]



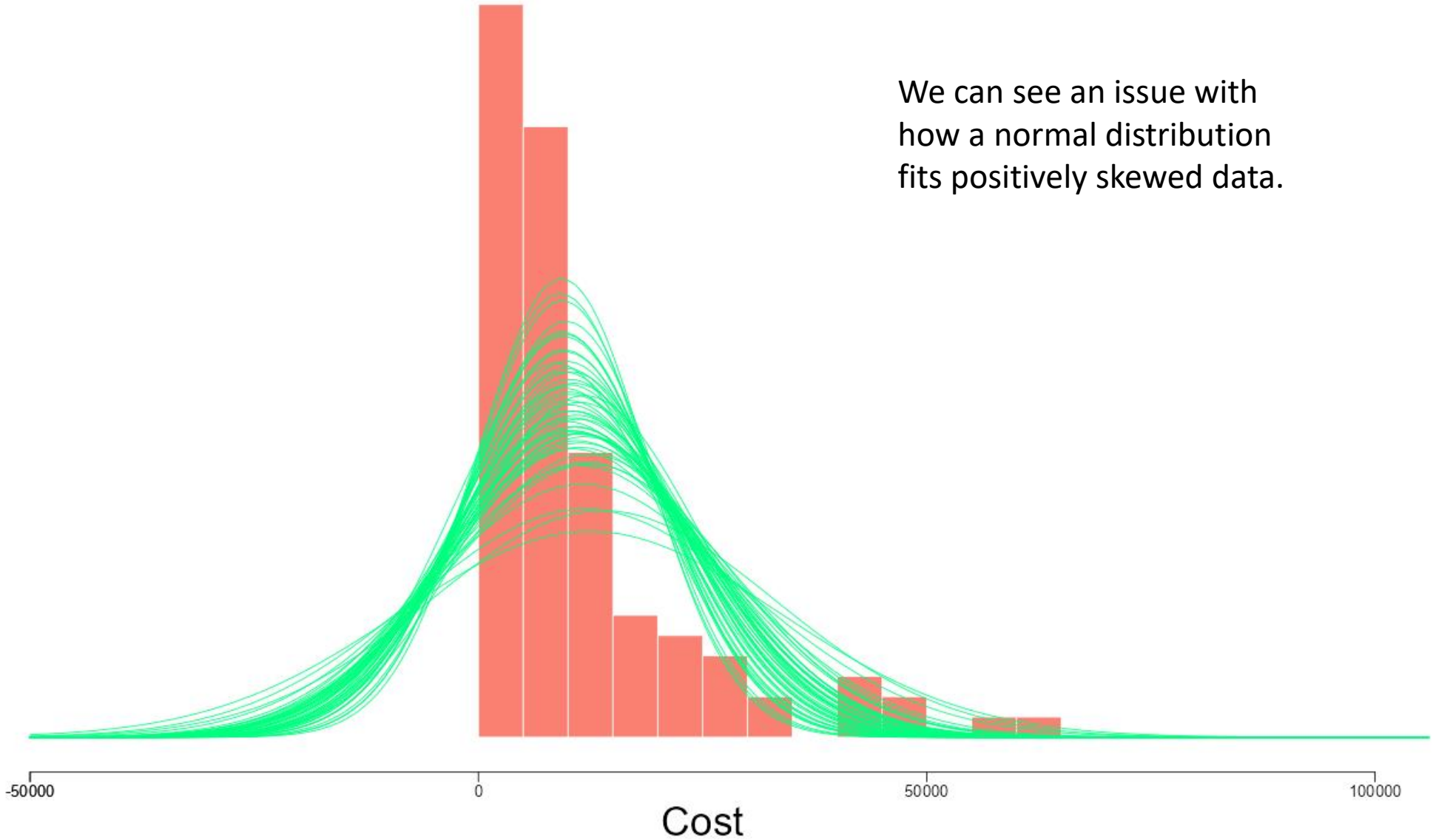
+ are individual players within each position.

There was no Designated Hitter position in the original example. The probabilities on the right are for meeting specific averages (i.e., potential probability of a DH's batting average of .300 is about 0.08).

Also, percentiles (e.g., 50th = 0.235)

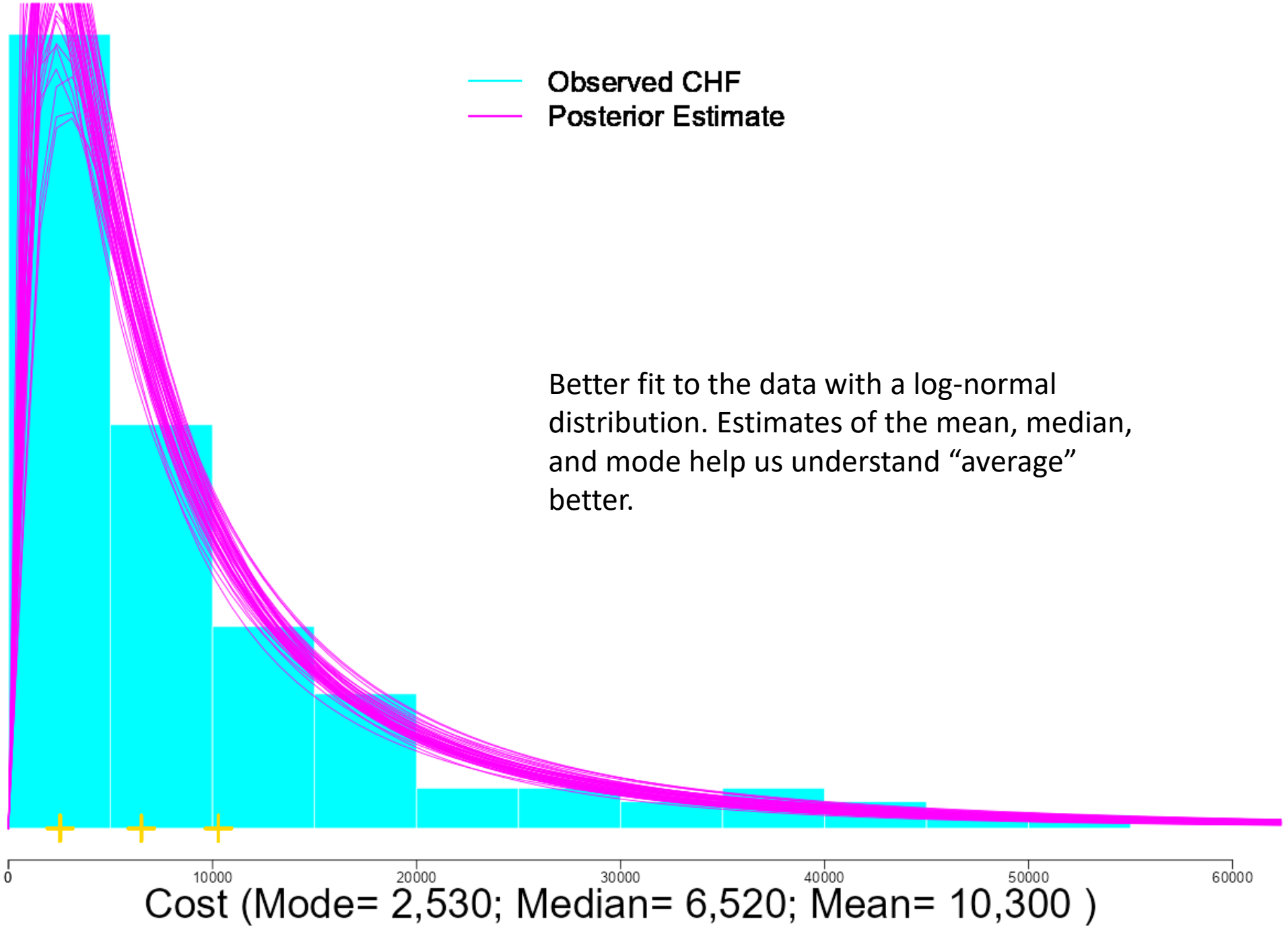
```
$Est.Prop.GT.Y
$Est.Prop.GT.Y$Y_0.23
     Mode    HDIlow   HDIhigh
0.5521178 0.2884207 0.7772578

$Est.Prop.GT.Y$Y_0.3
       Mode      HDIlow     HDIhigh
0.082626078 0.008068365 0.356581483


$Est.Effect.Size.2Y
     Mode    HDIlow   HDIhigh
0.8461453 0.4639655 1.2680568

$Est.Quantile.Y
$Est.Quantile.Y$Percentile_0.05
     Mode    HDIlow   HDIhigh
0.1559087 0.0934855 0.1953441

$Est.Quantile.Y$Percentile_0.25
     Mode    HDIlow   HDIhigh
0.1993641 0.1515007 0.2369288

$Est.Quantile.Y$Percentile_0.5
     Mode    HDIlow   HDIhigh
0.2352480 0.1932466 0.2780868

$Est.Quantile.Y$Percentile_0.75
     Mode    HDIlow   HDIhigh
0.2681554 0.2297939 0.3329683

$Est.Quantile.Y$Percentile_0.95
     Mode    HDIlow   HDIhigh
0.3238013 0.2739714 0.4327274


$Est.Mean.Beta
     Mode    HDIlow   HDIhigh
0.2368235 0.1970613 0.2808094
```

Posterior Predictive Check of health costs

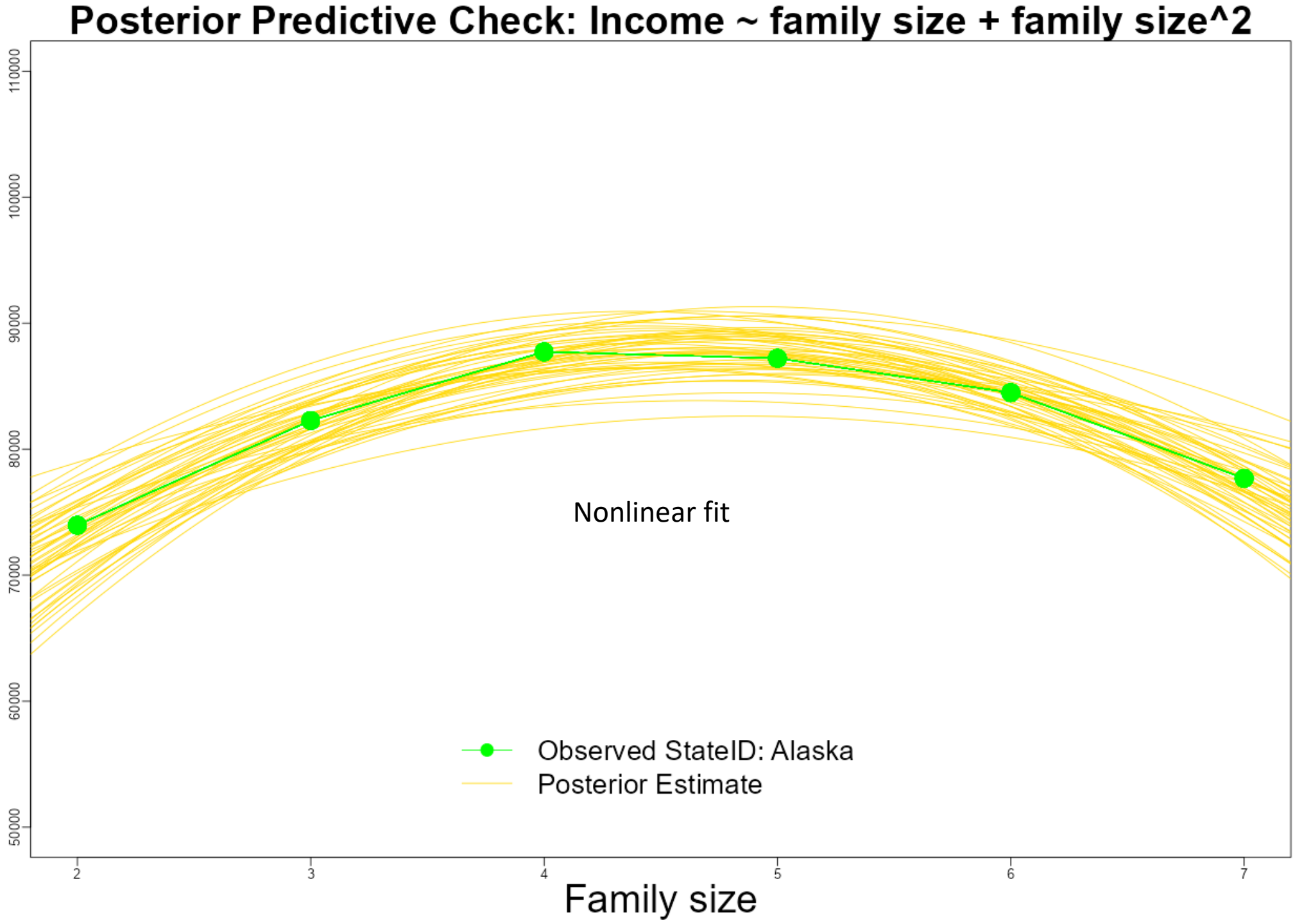


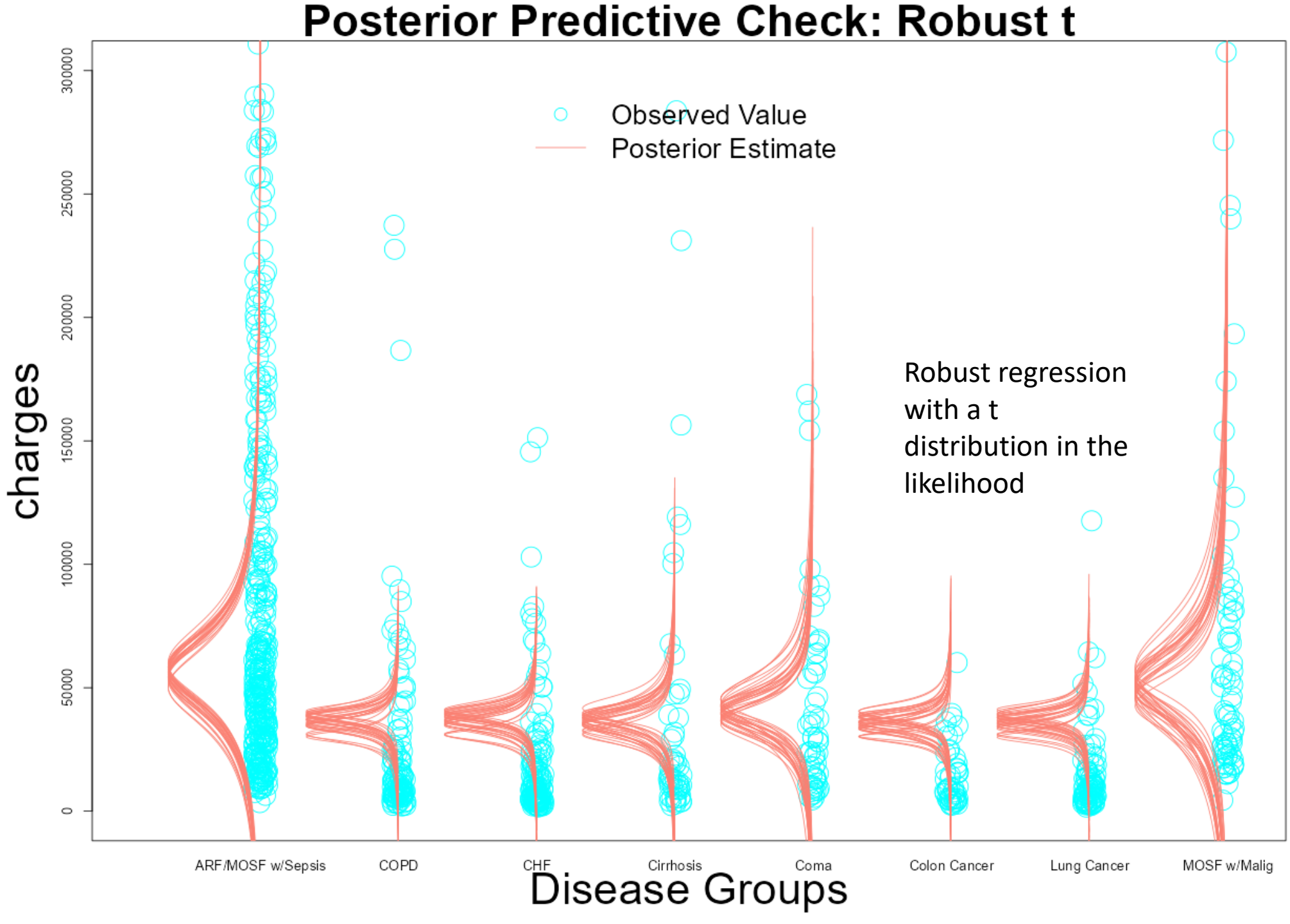
We can see an issue with how a normal distribution fits positively skewed data.

Posterior Predictive Check of health costs



Better fit to the data with a log-normal distribution. Estimates of the mean, median, and mode help us understand “average” better.


Posterior Predictive Check: Income ~ family size + family size^2
110000
100000
90000
80000
70000
60000
50000
2
3
4
5
6
7
Family size
Nonlinear fit
Observed StateID: Alaska
Posterior Estimate


Posterior Predictive Check: Robust t
Observed Value
Posterior Estimate
charges
0
50000
100000
150000
200000
250000
300000
ARF/MOSF w/Sepsis
COPD
CHF
Cirrhosis
Coma
Colon Cancer
Lung Cancer
MOSF w/Malig
Disease Groups
Robust regression with a t distribution in the likelihood