Computer Intensive Methods
Presentation 1

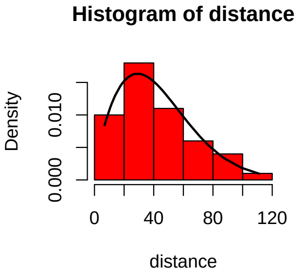
Jorriet Marc[1](), Olusoji Oluwafemi[1](1541893), [1]Orlowski Robert[1](1541889), Owokotomo Olajumoke Evangelina[1](1539654)

[1]Center for Statistics and Biostatistics, Universiteit Hasselt, Agoralaan D, 3590, Diepenbeek, Belgium.

October 14, 2016

Table of Contents

Exploration & Bootstrap Algorithm (Nonparametric)



i. Draw $B = 5000, 50000, 100000$ from the data (stopping distance) with replacement
ii. Compute statistic of interest(maximum) from each bootstrap sample(bootstrap replicates)
iii. Approximate $se(\hat{\theta})$ by the sample standard deviation of bootstrap replicates.

Results, Problem & Remedies

Results

Replicates	5000	50000	100000
SError	13.75	13.74	13.77

Table: Standard Error estimates

Problems and Remedy

- Discrete not continuous distribution observed (see Appendix graph)
- Poor approximation of the distribution of maximum (Bickel, Gotze, and van Zwet (1997)) but standard error still appropriately computed.
- Remedy includes m out of n bootstrapp (Bickel, Gotze, and van Zwet (1997)) and Semiparametric bootstrapp (Zelterman (1993)).

Bootstrap Algorithm (Parametric)

Candidate distributions based on extreme value theories include;

- i Gumbel
- ii Weibull
- iii Generalized extreme value distribution (GEV)

Question of the best distribution to assume

	Gumbel	Weibull	GEV
AIC	462.72	461.47	464.72

Table: AIC values of Candidate distributions

- i Draw $B = 5000, 50000, 100000$ from a distribution that best summarizes(weibull) the data with parameters replaced with their plugin estimates(1.7239251, 48.1481931)
- ii Compute statistic of interest(maximum) from each bootstrap sample(bootstrap replicates)
- iii Approximate $se(\hat{\theta})$ by the sample standard deviation of bootstrap replicates.

Results & Comarison

Results

Replicates	5000	50000	100000
SError	17.9	18.25	18.24

Table: Standard Error estimates

Comparison & Discussions

i Continuous and not discrete distribution observed.

ii Standard error computed larger than that of the semi-parametric bootstrapp.

iii This is partly because of the distribution used.

Bootstrap Algorithm (Non-Parametric) & Results

Algorithm

i Draw $B = 5000, 50000, 100000$ from the data (stopping distance) with replacement

ii Compute probability of interest $P(distance > 65ft) = \frac{\#(distance>65)}{n}$ from each bootstrap sample (bootstrap replicates)

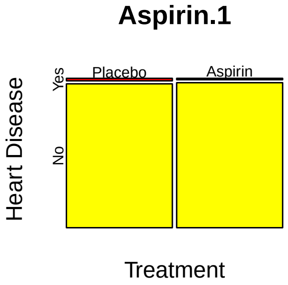
iii Approximate $P(distance > 65ft)$ by the mean of bootstrap replicates.

Results

Replicates	5000	50000	100000
Prob.	0.2	0.2	0.2

Table: Standard Error estimates

Odds Ratio and Confidence Interval



Results

i Observation unit are

ii Odds ratio is 1.83(1.44,2.33)

iii The odds of.....

Bootstrap Algorithm and Results

Bootstrap t Algorithm

i Draw B $=$ 5000 and 50000 bootstrap samples from the data and compute $Z^* = \frac{\hat{\theta^*} - \hat{\theta}}{se(\hat{\theta}^*)}$, where $\hat{\theta^*}$ is the bootstrap log-odds and $se(\hat{\theta}^*)$ is the corresponding bootstrap standard error.

ii Note that pairs are sampled

iii obtain quantiles of the replicate Z^*

iv Compute $(\hat{\theta} - \hat{t}_{1-\alpha} se(\hat{\theta}), \hat{\theta} + \hat{t}_{\alpha} se(\hat{\theta}))$, where $\hat{\theta}$ is the estimated log odds ratio and $se(\hat{\theta})$ is the estimated standard error of the odds ratio.

Results

Replicates	5000	50000
Quantiles	-2.01,1.99	-1.97,1.93
CI	1.43,2.34	1.44,2.32
Coverage	0.95	0.95

Table: Bootstrap Quantiles and Confidence Intervals

Note that this version of bootstrap is not transformation invariant. Hence building confidence interval for the log-odds without transforming is a better option.

Percental Interval(PI)

Percentile Interval Algorithm

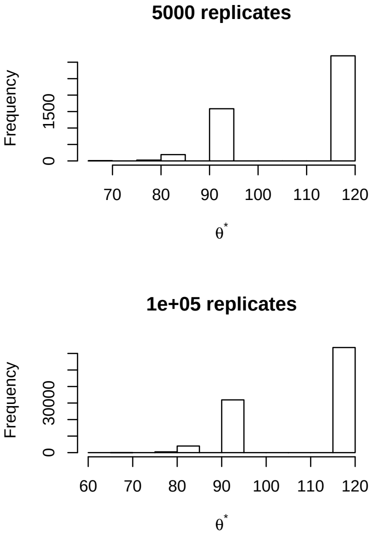
i. Draw B = 5000 and 50000 bootstrap samples from the data with replacement
ii. Compute bootstrap replicates (odds ratio).
iii. obtain quantiles of the replicates.

Results

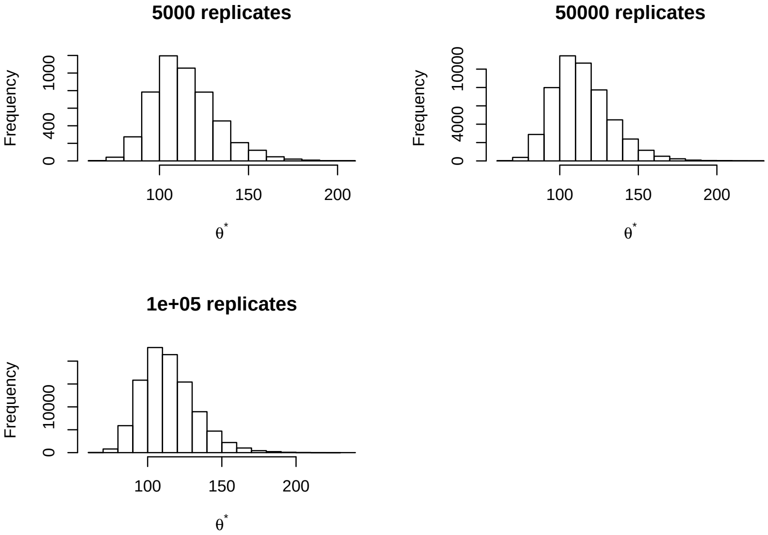
Replicates	5000	50000
CI	1.43,2.37	1.45,2.34
Coverage	0.95	0.95

Table: Confidence Intervals

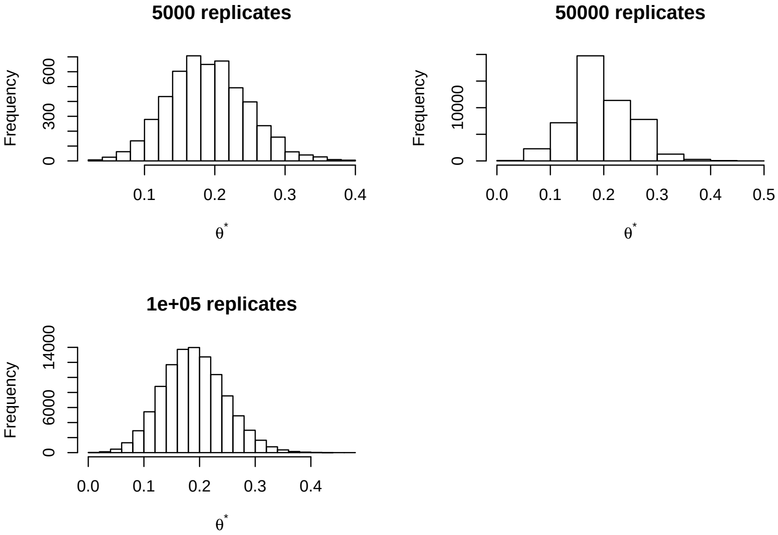
Distribution of Bootstrap Replicates (Non Parametric)



Distribution of Bootstrap Replicates (Parametric)



Distribution of Bootstrap Replicates (Probability)



Distribution of Bootstrap Replicates (Odds Ratio)