Pset 5

Darian Bailey

March 7, 2018

Abstract

Computational Statistics Stat 451

1.a:9.4.a

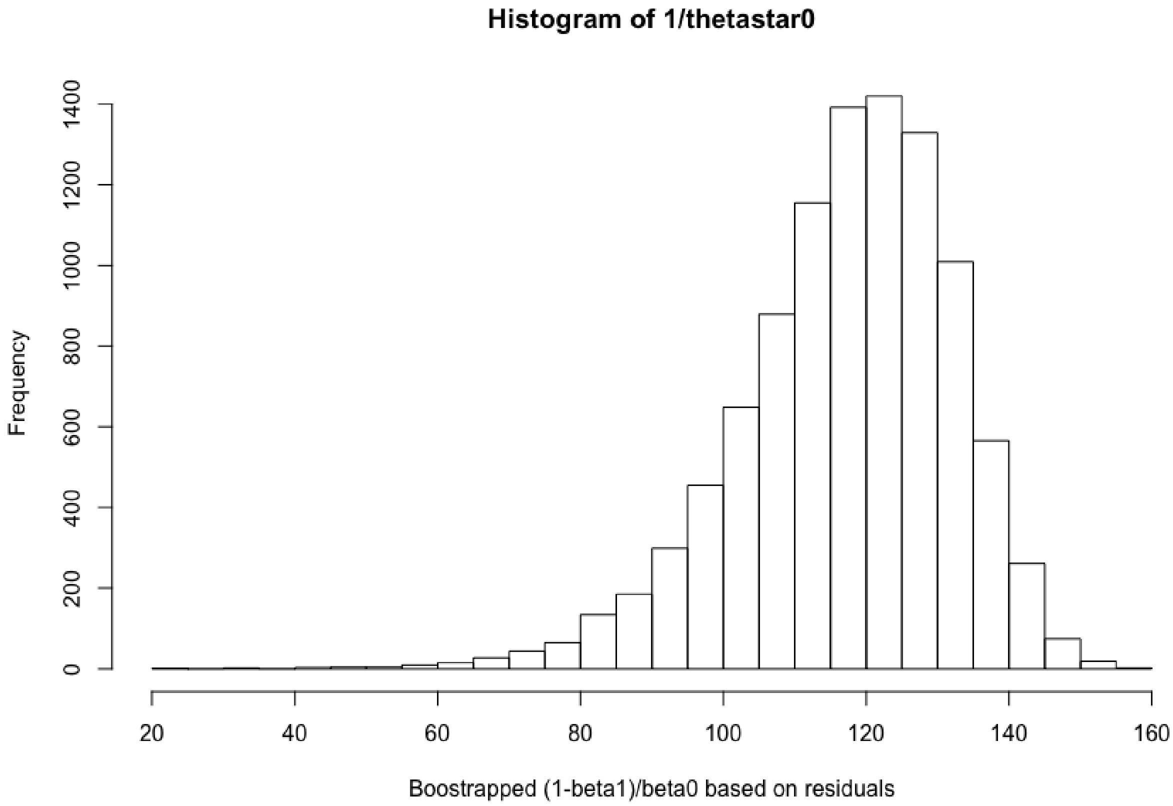
BootStrapping The Residuals:
Calculate regression and bootstrap the residuals

$$Use\ bootstrapped\ data\ to\ estimate\ 1/\hat{R} = (1 - Beta_1)/Beta_0$$
$$Estimate : 117.9$$
$$Confidence\ Interval : [83.03, 141.32]$$
$$Standard\ Error : 14.94$$
$$Histogram :$$

BootStrapping The Cases:
Bootstrap the original data

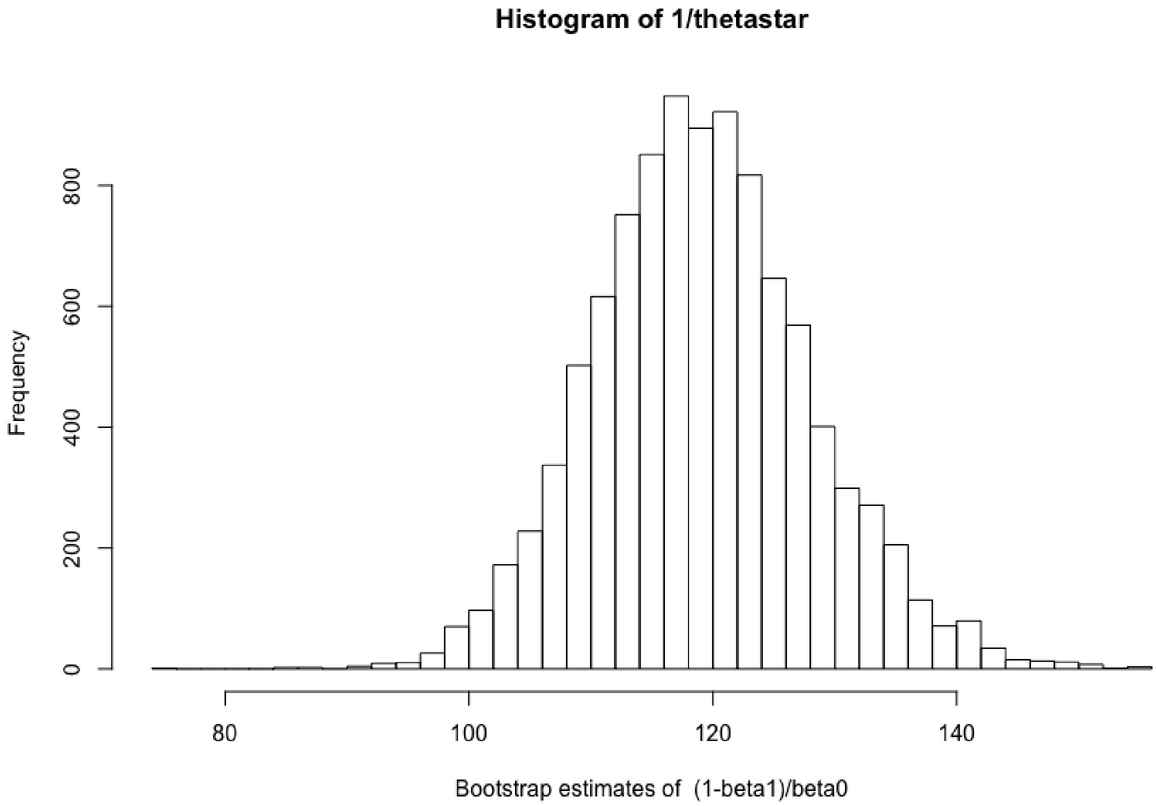
$$Use\ bootstrapped\ data\ to\ estimate\ 1/\hat{R} = (1 - Beta_1)/Beta_0$$
$$Estimate : 119.14$$
$$Confidence\ Interval : [102.5352, 137.5675]$$
$$Standard\ Error : 8.9$$
$$Histogram :$$

We can see above that bootstrapping the cases provides an estimate with a smaller standard error, less of a skew, and a smaller confidence interval. This makes sense because the model might not provide an appropriate fit for R and S and the residuals might not have a constant variance. For the following questions I'll be using the case estimate only



1.b:9.4.b

$$
\begin{aligned}
&Bias\ Corrected\ Estimate: \\
\bar{\theta}^* - (\bar{\theta}^* - \hat{\theta}) &= 119.14 \\
&Standard\ Error: \\
d(R_*)/\sqrt{B} &= 0.0009
\end{aligned}
$$



1.c:9.4.c

Nested Bootstrap

$$1) Generate\ Bootstrap\ pseudo\ data$$
$$2) For\ j = 1, 2..B_0$$
$$a) Generate\ X^*_{ji}.....X^*_{jB}\ iid(X_1...B_0)$$
$$b) Compute\ R_0(X^{**}_{jk}, F_j)$$
$$c) Get\ (\gamma_1, \gamma_2)\ = (\hat{F}_1^{-1}(\alpha/2), \hat{F}_0^{-1}(1 - \alpha/2))$$
$$d) Calculate\ quantile\ using\ \hat{F}_0^{-1}(\gamma_1), \hat{F}_0^{-1}(\gamma_2)$$

$$95\%\ Confidence\ Interval : [102.22, 129.69]$$

2.a:9.5.a

95% confidence intervals

$$BC_\alpha$$
$$Bootstrap\ mean\ and\ follow\ instructions\ on\ page\ 294-296$$
$$Stomach\ Cancer\ CI(log) = [2.522551\%, 97.52347\%] = [4.316303, 5.626966]$$
$$Stomach\ Cancer\ CI = [74.89, 278.66]$$
$$Breast\ Cancer\ CI(log) = [0.8140988\%, 94.65553\%] = [5.306486, 7.267744]$$
$$Breast\ Cancer\ CI = [201.64, 1433.31]$$

$$Bootstrap\ t$$
$$Bootstrap\ mean\ and\ follow\ instructions\ on\ page\ 296-297$$
$$Using\ Var(\hat{\theta})\ as\ variance\ estimator$$
$$Stomach\ Cancer\ CI(log) = [2.5\%, 97.5\%] = [70.77219,\ 324.55934]$$
$$Breast\ Cancer\ CI(log) = [2.5\%, 97.5\%] = [73.18363, 1656.03275]$$

2.b:9.5.b

Hypothesis test

place both data in one set and sample from it. $Calculate\ CI\ a$

Calculate test stat T$_{new} = mean(x1) - mean(x2)$

2.c:9.5.c

Bad CI

$$Simple\ Bootstrap\ CI\ on\ log\ data$$
$$[57.2338, 558.4158]$$
$$Simple\ Bootstrap\ CI\ on\ original\ data$$
$$[435.2417, 1235.7594]$$

We can see above that these confidence intervals are smaller and give smaller tails