Bootstrap and Jackknife in Regression

[redacted]

April 16, 2021

1 Introduction

In this paper, we investigate some elements of both Jackknife and bootstrapping regression models.The Jackknife is a tool for estimating standard errors and the bias of estimators. As its name suggests, the jackknife is a small, handy tool; in contrast to the bootstrap, which is then the moral equivalent of a giant workshop full of tools.

Both the jackknife and the bootstrap involve **resampling** data (*repeatedly creating new data sets from the original data*).

2 Jackknife method

Presentation of the method

The jackknife method came before **bootstrap** and was developed by **Maurice Quenouille** in **1949**. **John Tukey** expanded on the technique in **1958** and proposed the name **"Jackknife"**.[noa]

The jackknife deletes each observation and calculates an estimate based on the remaining *n-1* o them. It uses this collection of estimates to do things like estimate the bias and the standard error.[ET93]

- We'll consider the jackknife for univariate data
- Let X_1,...,X_n be a collection of data used to estimate a parameter θ
- Let $\hat{\theta}$ be the estimate based on the full data set
- Let $\hat{\theta}_i$ be the estimate of θ obtained by **deleting observation i**
- Let $\bar{\theta} = \frac{1}{n}\sum_{i=1}^{n}\hat{\theta}_i$

Estimate of the bias

Then, **the jackknife estimate of the bias** is

$$(n-1)(\bar{\theta} - \hat{\theta}) \tag{1}$$

which determine how far the average delete-one estimate is from the actual estimate)

Estimate of the standard error

The jackknife estimate of the standard error is

$$\left[\frac{n-1}{n}\sum_{i=1}^{n}(\hat{\theta}_i-\bar{\theta})^2\right]^{1/2} \tag{2}$$

Pseudo observations

Another interesting way to think about: the jackknife uses **pseudo observations**

$$PseudoObs = n\hat{\theta}-(n-1)\hat{\theta}_i \tag{3}$$

When $\hat{\theta}$ is the sample mean, the pseudo observations are the data themselves. The mean of these observations is a bias-corrected estimate of θ.

Confidence interval

We can also construct an approximate confidence interval with a confidence level *1-α*

$$\hat{\theta} \pm t_{\alpha/2;n-1}\sqrt{\frac{Var(\hat{\theta})}{n}} \tag{4}$$

with $t_{\alpha/2;n-1}$ is the appropriate quantile of a Student's law.

Jackknife example via hand and R

Set estimation

For example, if the parameter to be estimated is the population mean of x, we compute the mean $\bar{x}_i$ for each subsample consisting of all but the i-th data point:

$$\bar{x}_i = \frac{1}{n-1}\sum_{k=1,k\neq i}^{n} x_k$$

Let a sample of 3 observations x=(11,22,33)

$$\bar{x_1} = \frac{11+22}{2} = 16.5$$
$$\bar{x_2} = \frac{11+33}{2} = 22$$
$$\bar{x_3} = \frac{22+33}{2} = 27.5$$

These n estimates form an estimate of the distribution of the sample statistic if it were computed over a large number of samples. In particular, the mean of this sampling distribution $\bar{x_j}$ is the average of these n estimates:

$$\bar{\theta} = \bar{x_j} = \frac{1}{n}\sum_{k=1}^{n} x_k$$

$$\bar{x_j} = \frac{(16.5 + 22 + 27.5)}{3} = 22 = \frac{(11 + 22 + 33)}{3} = \bar{x}$$

Estimate of the standard error

Now with that, we can also follow the variance of $\bar{x}_j$

$$var(\bar{x}_j) = \frac{n-1}{n} \sum_{k=1}^{n} (\bar{x}_k - \bar{x}_j)^2$$

$$\Leftrightarrow var(\bar{x}_j) = \frac{3-1}{3}((16.5-22)^2 + (22-22)^2 + (27.5-22)^2)$$

$$\Leftrightarrow var(\bar{x}_j) = \frac{2}{3}((-5.5)^2 + 0^2 + 5.5^2) = 40.3333$$

$$sd(\bar{x}_j) = \sqrt{40.3333} = 6.3509$$

Estimate of the bias

Similarly we can also find the bias of $\bar{x}_j$,

$$B\hat{i}as(\bar{x}_j) = (n-1)(\bar{x}_j - \bar{x})$$

$$B\hat{i}as(\bar{x}_j) = (3-1)(22-22) = 0$$

Pseudo observations

As we said before, when $\hat{\theta}$ is the sample mean, the bias of this estimation is equal to 0, so the pseudo observations are the data themselves.

$$x_{\hat{Jack}} = n\bar{x} - (n-1)\bar{x}_j$$

$$x_{\hat{Jack}} = n\bar{x} - (n-1)\bar{x}_j$$

$$x_{\hat{Jack}} = 3(22) - 2(22) = 66 - 44 = 22$$

We can compute the Jackknife method in R:

```
library(bootstrap)
x=c(11,22,33)

theta=function(x){mean(x)}

jackknife(x,theta)
```

Figure 1: Jackknife method computation in R

```
$jack.se
[1] 6.350853

$jack.bias
[1] 0

$jack.values
[1] 27.5 22.0 16.5
```

Figure 2: Jackknife method output in R

Regression application

Let's say we're doing a linear regression ($Y \sim X1 + X2$) with these 200 variables:

	X1	X2	Y
1	230.10	37.80	22.10
2	44.50	39.30	10.40
3	17.20	45.90	9.30
4	151.50	41.30	18.50
5	180.80	10.80	12.90
6	8.70	48.90	7.20
7	57.50	32.80	11.80
8	120.20	19.60	13.20
9	8.60	2.10	4.80
10	199.80	2.60	10.60
...	...	...	...

We can now find an estimation of all the previous estimators in the output with R.

resultsB0	list [4]	List of length 4
jack.se	double [1]	0.3617116
jack.bias	double [1]	0.008285467
jack.values	double [200]	2.73 2.71 2.72 2.72 2.71 2.67 ...

Figure 3: Output of the model regression in R

Using these estimators, we can deduce the variance and the confidence interval of the predictors:

	Variance	CI	CH
B0	7e-04	2.72	2.72
B1	1.6e-08	0.05	0.05
B2	5.3e-07	0.19	0.19

We find the following equation:

$$Y = 2.72 + 0.05X_1 + 0.19X_2$$

3 Bootstrap

We will discuss about the bootstrapping methodology itself and proceed to extend this to the various applications of regression such as calculating variable coefficients, confidence intervals, and component/factor analysis.

The basic approach to the bootstrap methodology is to assume that the population is to the sample as the sample is to the bootstrap samples. That is, just as population parameters can be estimated by samples from the population, sample characteristics can be estimated by randomly drawn, with-replacement, n-length subsets of the sample, also referred to as bootstrap samples. Thus, by the transitive property, population parameters can be estimated using bootstrap samples. [Fox02]

Some benefits of the Bootstrap techniques are independence from the normality assumption, better replicability, and the ability to estimate any parameter, whether previously defined or not. In fact, the only required assumption is that the sample data be representative of the population being studied. [ZT07]

Bootstrapped Parameters

So we have a sample of 5 values from a population : $\{x_1, x_2, x_3, x_4, x_5\}$ from this sample we're gonna do B samples with a simple methods for b from 1 to B and for k = 1 to 5 (construction of the bootstrap sample) Draw randomly and uniformly a data among the values of the initial sample, sampling with replacement. By example for the first Bootstrap sample:

for k=1, drawing=4→$\text{Ech}_1 = \{x_4\}$
for k=2, drawing=2→$\text{Ech}_1 = \{x_4, x_2\}$
for k=3, drawing=5→$\text{Ech}_1 = \{x_4, x_2, x_5\}$
for k=4, drawing=1→$\text{Ech}_1 = \{x_4, x_2, x_5, x_1\}$
for k=5, drawing=4→$\text{Ech}_1 = \{x_4, x_2, x_5, x_1, x_4\}$

After we have the first sample we need to take a parameter named θ we want to observe, θ is the parameter of the population, θ_0 is the parameter of the initial sample and θ_b is the parameter of the b^{th} bootstrap sample. The parameter can be the parameter of an exponential law (λ) [Cas]. We can observe the mean or the median but it can be other interesting parameter, like estimators of linear regression. Here if we want to have the mean we just need to calculate the mean so here we're gonna have:

$$\theta_1 = \sum_{i=1}^{5} \frac{x_i}{5} \tag{5}$$

After taking the parameter of one sample we need to create another B-1 samples and take the parameter of all of them. Then we're gonna do the average of all the parameters and according to the bootstrap principle, the distribution of the average with respect to the parameter of the initial sample approximates the distribution of the parameter of the initial sample with respect to the population.[Mar]

$$\hat{\theta} = \sum_{b=1}^{B} \frac{\theta_b}{B} \tag{6}$$

And we can have the standard error of this parameter:

$$se_{\hat{\theta}} = \sqrt{\frac{\sum_{b=1}^{B}(\theta_b - \hat{\theta})}{B-1}} \tag{7}$$

Confidence Interval

The bootstrap method have 5 different confidence intervals:

Normal Interval

This method computes the interval from the quantiles of the t distribution, as if $\hat{\theta}$ follow a normal distribution. So for a 95% confidence interval we're gonna have:

$$(\theta_0 + t_{(n-1)0.025} se_{\hat{\theta}}, \theta_0 + t_{(n-1)0.975} se_{\hat{\theta}}) \tag{8}$$

With $se\_\hat{\theta}$ calculation in (6), n the size of the sample, and $t_{(n-1)q}$ is the quantile q of the distribution t with n - 1 degrees of freedom. And we can say $t_{(n-1)q} = -t_{(n-1)(1-q)}$.

Percentile Interval

Here it's just the Interval of the percentile so if you want a 95% confidence interval we're gonna have the 2.5% and 97.5% percentile for the interval:

$$(\hat{\theta}_{0.025}, \hat{\theta}_{0.975}) \tag{9}$$

Basic Interval

In this interval we also use the percentile but here we use them in the difference $\hat{\theta} - \theta_0$ because as we said before we can approximate the distribution $\theta_0 - \theta$ with the previous one:

$$\begin{aligned}
&(\hat{\theta}_{0.025} - \theta_0 \leq \hat{\theta} - \theta_0 \leq \hat{\theta}_{0.975} - \theta_0) \\
\Leftrightarrow\ &(\hat{\theta}_{0.025} - \theta_0 \leq \theta_0 - \theta \leq \hat{\theta}_{0.975} - \theta_0) \\
\Leftrightarrow\ &(\theta_0 - (\hat{\theta}_{0.975} - \theta_0) \leq \theta \leq \theta_0 - (\hat{\theta}_{0.025} - \theta_0)) \\
\Leftrightarrow\ &(2\theta_0 - \hat{\theta}_{0.975} \leq \theta \leq 2\theta_0 - \hat{\theta}_{0.025})
\end{aligned} \tag{10}$$

Studentised Interval

The principle is the same than the Basic Interval but here the difference is normalized by the standard deviation of $\hat{\theta}$:

$$t = \frac{\hat{\theta} - \theta_0}{s_{\hat{\theta}}} \tag{11}$$

This interval is correcting some error of the Basic Interval but it means so much we need an estimate of the bootstrap variance and one way to do that is to redo the bootstrap method with each bootstrap sample meaning lot of time.

Bias-corrected and accelerated Interval

This interval is close to the Percentile Interval but we don't choose fixed value for the percentile the method choose different percentiles taking into account the bias and skewness of the distribution. I don't understand yet all the calculation we need to do that.

A note on choosing an ideal n

We continue to treat n as the size of any given bootstrap sample, picked from the original data sample. We also observe that a small original sample size is limiting in its ability to describe, with confidence, the population it is meant to represent and as a direct result of the bootstrapping principal, a small n is limiting in its ability to describe, with confidence, the original sample it is mean to represent.

Bootstrapping is meant to be robust against small sample sizes so n need not be very large but a large number of replicates r is preferred. In fact, according to [OB21], we may desire to pick bootstrap samples of the same size as the original sample. This is meant to achieve a similar accuracy as we would if we were using the original sample. One way we can think of accuracy is as values of standard error or the ranges of our confidence intervals. Higher accuracy would imply smaller SE and slimmer ranges for our CI's. In order to test the assertion that larger bootstrapping sample sizes improves accuracy, using the 'flowers' dataset from R, we compute the SE for the bootstrapped mean of the variable Flowers at $n = 6, 8, 10, 12$ and $r = 20$. The results may be seen in Table 1.

Timing	n	Bootstrapped Standard Error
Before	6	5.2
	8	2.99
	10	3.74
	12	3.224
PFI	6	6.3
	8	4.88
	10	3.34
	12	3.9

Table 1: SE for flowers; r = 20

We see that the general trend for increasing n is decreasing standard error. Thus, we may conclude that a larger n is likely to produce a more accurate parameter. We will proceed to use

$n =$ the original sample size.

Since the bootstrapping method is random, the SE, and any other characteristic of the distribution, may change over each iteration. This is why we see a very low SE for $n = 8$ and $Timing = Before$. The expectation is that these values would converge for higher number of replicates r. We see in Table 2 that this is likely true, as we increase r to 2000.

Timing	n	Bootstrapped Standard Error
Before	6	4.77
	8	4.06
	10	3.66
	12	3.38
PFI	6	5.12
	8	4.32
	10	3.81
	12	3.6

Table 2: SE for flowers; r = 2000

Bootstrap Regression with Random Regressors (X)

There are two basic bootstrap sampling approaches to regression. The Random-X approach pulls bootstrap samples, as described above, directly from from the set of observations. Each element of the bootstrap sample consists of the response variable Y_i and the predictors X_{ij} for $i =$ *the observation index in the original sample* and $j =$ *the predictor variable index*. As before, this element is randomly drawn from the initial dataset, with replacement. n many elements are drawn this way and added to the bootstrap sample.

We will apply this method with the previous linear regression ($Y \sim X1 + X2$) with these 200 variables:

	X1	X2	Y
1	230.10	37.80	22.10
2	44.50	39.30	10.40
3	17.20	45.90	9.30
4	151.50	41.30	18.50
5	180.80	10.80	12.90
6	8.70	48.90	7.20
7	57.50	32.80	11.80
8	120.20	19.60	13.20
9	8.60	2.10	4.80
10	199.80	2.60	10.60
...	...	...	...

We have two different approach to do the bootstrap on a linear regression, we can resample the observations or resample the residuals.

Resampling the observations

It's not that complex to understand what we do here we just need to realise a draw with replacement. Here we can see we have two times 139 in the 4^{th} and 9^{th} row. It's gonna be our Bootstrap sample.

	X1	X2	Y
73	26.80	33.00	8.80
152	121.00	8.40	11.60
83	75.30	20.30	11.30
139	43.00	25.90	9.60
111	225.80	8.20	13.40
16	195.40	47.70	22.40
162	85.70	35.80	13.30
108	90.40	0.30	8.70
139	43.00	25.90	9.60
147	240.10	7.30	13.20
...	...	...	...

Then the model is adjusting for all the new bootstrap sample. And after that we just need to estimate each coefficient (the Intercept β_1 and β_2) like we did in the equation (6). Then we're gonna have also the 5 confidence Interval we saw before.

Bootstrap Regression with Fixed- X

We briefly describe a Fixed-X bootstrap regression. In the Random-X bootstrap regression, we randomly draw observations, with replacement, from the original dataset, to include the associated predictor observation. This approach implies that the regressor variables be treated as random. [Tex21].

If instead we wish to treat X as fixed, as in an experimental design, we can do so with an established procedure in bootstrapped regression. Instead of randomly drawing observations from the original dataset as in Random-X, we concern ourselves with the sample error or sample residuals of a standard regression.

In this approach we need to have the residuals thanks to the linear regression, it's given by the real value of Y-the fitted value of Y $\Leftrightarrow Y - \hat{Y}$. So like we can see in this table:

	X1	X2	Y	fitted	residuals
1	230.10	37.80	22.10	20.56	1.54
2	44.50	39.30	10.40	12.35	-1.95
3	17.20	45.90	9.30	12.34	-3.04
4	151.50	41.30	18.50	17.62	0.88
5	180.80	10.80	12.90	13.22	-0.32
6	8.70	48.90	7.20	12.51	-5.31
7	57.50	32.80	11.80	11.72	0.08
8	120.20	19.60	13.20	12.11	1.09
9	8.60	2.10	4.80	3.71	1.09
10	199.80	2.60	10.60	12.55	-1.95
...	...	...	...	...	...

The principle of this method is to resampling the residuals and add the new value of the residuals to the fitted value to have the bootstrap sample. So here in the table, the bootstrap sample is in the column newY:

	X1	X2	Y	fitted	residuals	newY
1	230.10	37.80	22.10	20.56	-3.21	17.35
2	44.50	39.30	10.40	12.35	0.68	13.02
3	17.20	45.90	9.30	12.34	-5.31	7.02
4	151.50	41.30	18.50	17.62	-0.47	17.14
5	180.80	10.80	12.90	13.22	-1.16	12.06
6	8.70	48.90	7.20	12.51	-0.20	12.31
7	57.50	32.80	11.80	11.72	-5.31	6.41
8	120.20	19.60	13.20	12.11	0.24	12.34
9	8.60	2.10	4.80	3.71	0.24	3.95
10	199.80	2.60	10.60	12.55	-1.04	11.51
...	...	...	...	...	...	...

Then the model is adjusting for all the new bootstrap sample but the fitted value don't change just the value of the residuals is resample.

A Brief Case Study in R

We employ boostrapping methods to analyze mortality data from the sinking of the Titanic. The response variable is binary on $\{0, 1\}$ for survived or not. Thus, we employ a logistic regression with a binomial distribution to determine odds of surviving or not. The regressors vary from Age, Sex, Fare, etc. The data was retrieved from [Sta21]. The regression is performed in the same way as was done previously but with the different model. We will not focus on results of the regression itself as that is outside the scope of this paper.

Instead, we will first derive the statistical model through standard logistic regression and then we will perform the bootstrap to compare results and evaluate certain aspects of the bootstrap as it pertains to this study.

Figure 4 shows the summary of the logistic regression from R. We proceed to perform the bootstrap regression using the logistic model.

```
> summary(logit)

Call:
glm(formula = Survived ~ Age + Fare + SibSp, family = binomial(link = "logit"),
    data = train)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.9086  -0.9417  -0.7938   1.2223   1.9113

Coefficients:
             Estimate Std. Error z value Pr(>|z|)
(Intercept) -0.113944   0.207852  -0.548  0.58355
Age         -0.024922   0.006203  -4.018 5.87e-05 ***
Fare         0.019907   0.002917   6.824 8.88e-12 ***
SibSp       -0.317767   0.098713  -3.219  0.00129 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 964.52  on 713  degrees of freedom
Residual deviance: 880.17  on 710  degrees of freedom
  (177 observations deleted due to missingness)
AIC: 888.17

Number of Fisher Scoring iterations: 5
```

Figure 4: Logistic Regression

The model coefficients are calculated for each bootstrap sample and recorded in a matrix which is a concatenation of vectors whose individual elements are the bootstrap coefficients calculated for each sample. To illustrate how the bootstrap method works, we show histograms for the many calculated coefficients for the regressor Age at $r = 100, 1000, 10000$.

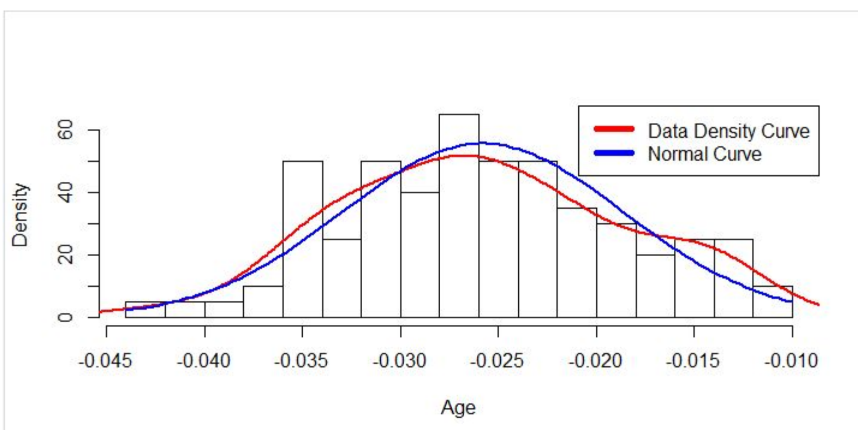

Figure 5: Age, $r = 100$

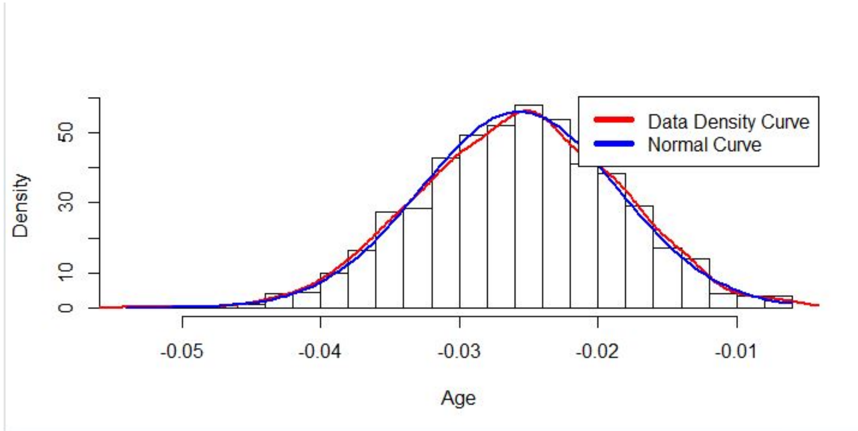

Figure 6: Age, $r = 1000$

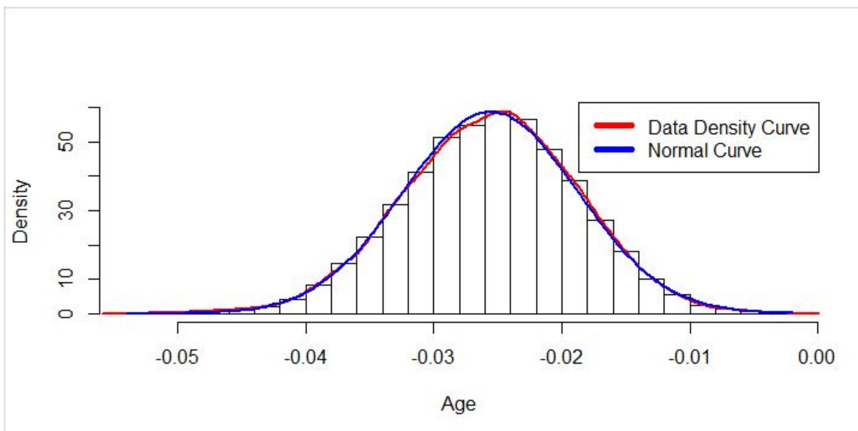

Figure 7: Age, $r = 10000$

It can easily be seen how, as the number of bootstrap samples increases per calculation, the sample of coefficients continues to converge on a normal distribution. The red curve converges on the blue curve. So, we may conclude that the sample coefficients have a normal distribution. This means we may use models and parameters that depend on normality.

Further, Figure 8 shows the histogram and QQ-Plot of t^*.

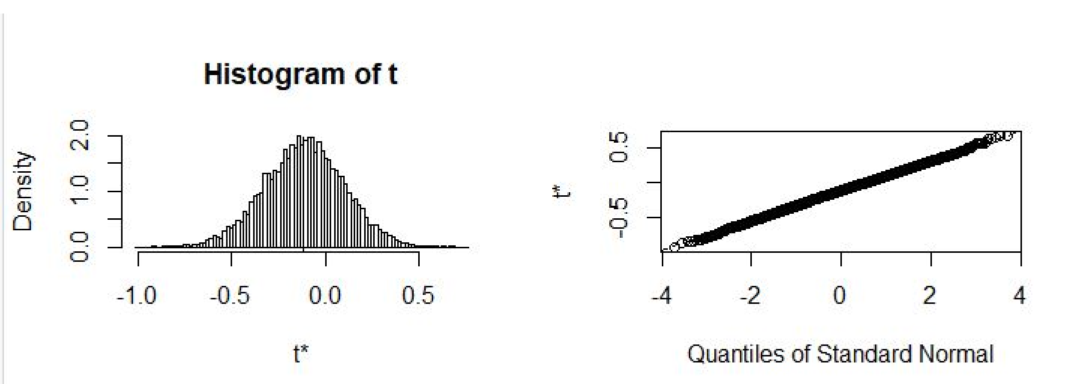

Figure 8: Bootstrapped Logistic Model Summary

We can see the histogram of t^* is very close to normal and the QQ-plot supports this.

Figure 9 shows the bootstrapped coefficients derived from the bootstrapped model. We can compare this to the original results above. The coefficients $t1^*$ = intercept through $t4^*$ = SibSp, are extremely close to equivalent. Additionally, the bias incurred is relatively small.

```
> coef.boot

ORDINARY NONPARAMETRIC BOOTSTRAP


Call:
boot(data = train, statistic = myLogitCoef, R = 10000, formula = logit$formula)


Bootstrap Statistics :
       original        bias    std. error
t1* -0.11394444  0.0023269005 0.214918766
t2* -0.02492193 -0.0006046442 0.006824368
t3*  0.01990657  0.0007152021 0.004753524
t4* -0.31776702 -0.0076755835 0.102165522
```

Figure 9: Bootstrapped Logistic Model Summary

Finally, we can also consider a normal confidence interval for each of the three coefficients and the intercept. Since the sample coefficients appear most normal at $r = 10000$, we can produce traditional confidence intervals using the computed bootstrap mean and standard error. The results may be seen in Figures 10 and 11.

```
> boot.ci(coef.boot, type="norm", index=3)
BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
Based on 10000 bootstrap replicates

CALL :
boot.ci(boot.out = coef.boot, type = "norm", index = 3)

Intervals :
Level      Normal
95%   ( 0.0098,  0.0286 )
Calculations and Intervals on Original Scale
> boot.ci(coef.boot, type="norm", index=4)
BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
Based on 10000 bootstrap replicates

CALL :
boot.ci(boot.out = coef.boot, type = "norm", index = 4)

Intervals :
Level      Normal
95%   (-0.5108, -0.1082 )
Calculations and Intervals on Original Scale
```

Figure 10: Fare and SibSp CIs

```
> boot.ci(coef.boot, type="norm", index=1)
BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
Based on 10000 bootstrap replicates

CALL :
boot.ci(boot.out = coef.boot, type = "norm", index = 1)

Intervals :
Level      Normal
95%   (-0.5366,  0.3034 )
Calculations and Intervals on Original Scale
> boot.ci(coef.boot, type="norm", index=2)
BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
Based on 10000 bootstrap replicates

CALL :
boot.ci(boot.out = coef.boot, type = "norm", index = 2)

Intervals :
Level      Normal
95%   (-0.0377, -0.0109 )
Calculations and Intervals on Original Scale
> boot.ci(coef.boot, type="norm", index=3)
BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
Based on 10000 bootstrap replicates
```

Figure 11: Intercept and Age CIs

4 Conclusion: Comparison of bootstrap and jackknife

The main application for the Jackknife is to **reduce** bias **and evaluate** variance **for an estimator**. It can also be used to:

- Find the standard error of a statistic
- Estimate precision for an estimator θ

Some specific differences with the Bootstrap:

- The bootstrap **requires a computer and is about ten times more computationally intensive**. The Jackknife can (at least, theoretically) be performed by hand.
- Estimate precision for an estimator θ
- The bootstrap is **conceptually simpler** than the Jackknife. The Jackknife requires n repetitions for a sample of n (for example, if you have 10,000 items then you'll have 10,000 repetitions), while the bootstrap requires the number of repetitions we chose.
- In most cases, the Jackknife **doesn't perform as well** the Bootstrap.
- The Jackknife is **more conservative** than bootstrapping, producing slightly larger estimated standard errors.
- The Jackknife gives the **same results** every time, because of the small differences between replications. The bootstrap gives **different results** each time that it's run.
- The Jackknife tends to **perform better** for confidence interval estimation for pairwise agreement measures.
- Bootstrapping performs better for **skewed distributions.**
- The Jackknife is more suitable for **small original data samples**.

Summary

Obviously, we can continue to derive results using the bootstrap methods ad infinitum. There have been many books dedicated to the bootstrap alone and countless scholarly articles.

Some of the main takeaways are that we can use bootstrapping in regression without concern over normality or worrying over sample size. We can produce thousands of replicates using only the original data sample and do not have to concern ourselves with gathering data, an arduous process, more than once or in performing an experiment more than once.

The only assumption we need concern ourselves with - in no way a humble one - is does the original sample correctly reflect the population from which it was drawn? If it does not, our results won't either. If it does, then we have an extremely powerful tool with which to make predictions and to draw inference from a population.

References

[Cas] Bruno Castanier. Estimation des quantités d'interêt en cas de faible information.

[ET93] Bradley Efron and Robert J. Tibshirani. *An Introduction to the Bootstrap*. Springer US, Boston, MA, 1993.

[Fox02] John Fox. Bootstrapping regression models. *An R and S-PLUS Companion to Applied Regression*, pages 1–14, 2002.

[Mar] Philippe Marchand. La méthode du bootstrap.

[noa] The Jackknife - Confidence Intervals, Bootstrapping, and Plotting.

[OB21] Jeremy Orloff and Jonathan Bloom. Bootstrap confidence intervals, 2021.

[Sta21] Kaggle Staff. Titanic - machine learning from disaster, 2021.

[Tex21] Bootstrapping regression models, 2021.

[ZT07] Linda R. Zientek and Bruce Thompson. Applying the bootstrap to the multivariate case: Bootstrap component/factor analysis. *Behavior Research Methods*, pages 318–325, 2007.