Capital Asset Pricing Model on Top 10 Companies Listed in NSE India using Paired Bootstrap Regression

Name: Tanmoy Paul Date: 26/12/2022

Abstract

This study focuses on the top ten companies listed on the National Stock Exchange of India (NSE), based on the weightage in Nifty. Here, we'll apply the CAPM model using paired bootstrap regression. In the CAPM model, the Beta (Systematic Risk) role is to estimate the equity's risk (its role is to measure the volatility/systematic risk of an equity/stock compared to the market as a whole). We will also estimate the alpha value to have an idea of whether the stock is undervalued, overvalued, or fairly priced. Lastly, we compute R-Squared for each of the fitted CAPM models to understand how closely alpha and beta reflect a stock's return as opposed to how much is random or due to other unobserved factors. In other words, R-Squared measures how closely the performance of a stock can be contributed to the performance of a selected benchmark index.

Introduction

The Capital Asset Pricing Model (CAPM) describes the relationship between systematic risk, and expected return for assets, particularly stocks. It is a finance model that establishes a linear relationship between the required return on an investment and risk. The model is based on the relationship between an asset's beta, the risk-free rate (typically the Treasury rate), and the equity risk premium, or the expected return on the market minus the risk-free rate.

CAPM evolved as a way to measure systematic risk(beta). The goal of the CAPM is to evaluate whether a stock is fairly valued or not. Thus Investors use CAPM as a tool for Stock selection.

Capital Asset Pricing Model

CAPM for an equity can be explained as:

$$\left(r_t - r_t^f\right) = \alpha + \beta\left(r_t^m - r_t^f\right) + \varepsilon_t$$

$$\bar{r}_t = \underbrace{\alpha + \beta \bar{r}_t^m}_{systematic\ risk} + \underbrace{\varepsilon_t}_{idiosyncratic\ risk}$$

$$E(\bar{r}_t) = \alpha + \beta \bar{r}_t^m.$$

Where, $E(\varepsilon_t)= 0$, $V(\varepsilon_t)=\sigma^2$ $\forall$ t and $Cov(\varepsilon_t, \varepsilon'_t)=0$

r_t is the return of asset/equity. Ex: return of Reliance equity.

r_t^m is the return of market index. Ex: return of Nifty 50 ETF.

r_t^f is the risk-free rate of return. Ex: return of SBI's Fixed Deposit.

$\bar{r}_t = \left(r_t - r_t^f\right)$ can be viewed as premium for taking risk with equity.

$\bar{r}_t^m = \left(r_t^m - r_t^f\right)$ can be viewed as premium for taking risk with market.

Systematic Risk is the risk which we can explain as due to market movement.

Idiosyncratic Risk is the risk very specific to stock/asset and we cannot explain.

INTERPRETATION OF α :

IF α = 0 : STOCK IS FAIRLY VALUED.

IF α > 0 : STOCK IS UNDERVALUED.

IF α < 0 : STOCK IS OVERVALUED.

INTERPRETATION OF β:

IF β = 1 : STOCK IS EXACTLY VOLATILE AS THE MARKET.

IF β > 1 : STOCK IS MORE VOLATILE THAN THE MARKET.

IF 0 < β < 1 : STOCK IS LESS VOLATILE THAN THE MARKET.

IF β = 0 : STOCK IS UNCORRELATED TO THE MARKET.

IF β < 0 : STOCK IS NEGATIVELY CORRELATED TO THE MARKET.

TOP 10 COMPANIES:

SL. No.	SYMBOL	COMPANY NAME	EQUITY CAPITAL (IN RS.)	WEIGHTAGE (%)
1	RELIANCE	Reliance Industries Ltd.	67653559330	11.36
2	HDFCBANK	HDFC Bank Ltd.	5569959026	8.53
3	ICICIBANK	ICICI Bank Ltd.	13934848622	8
4	INFY	Infosys Ltd.	21038767370	7.21
5	HDFC	Housing Development Finance Corporation Ltd.	3627111418	5.89
6	TCS	Tata Consultancy Services Ltd.	3659051373	4.19
7	ITC	ITC Ltd.	12399178702	3.61
8	KOTAKBANK	Kotak Mahindra Bank Ltd.	9928208080	3.45
9	LT	Larsen & Toubro Ltd.	2810382398	3.02
10	HINDUNILVR	Hindustan Unilever Ltd.	2349591262	2.89

SOURCE: www1.nseindia.com

These all selected companies are effective securities of the stock market of India.

ABOUT THE DATASET:

The daily adjusted closing price data on NIFTY 50(^NSEI) and all the above-selected companies is downloaded between 1-1-2022 and 16-12-2022 from Yahoo using the function `get.hist.quote()` from the `tseries` package.

For Risk-free interest rate, I have considered 3 month Treasury Rate. So, the risk-free rate dataset span from 1-1-2022 to 16-12-2022 and was downloaded from https://in.investing.com/rates-bonds/india-3-month-bond-yield-historical-data

Assumption of the Efficient Market:

A few assumptions of an Efficient Market are:-

1. No insider trading is allowed.
2. All publicly available information is already available to everybody and easily accessible.
3. There is no substantial tax, transaction cost, entry or exit bar.
4. No limitation on long and short positions.

Under this assumption, any non-zero α will be discovered very quickly by the people and they will take position accordingly.

As a result the $\alpha \to 0$.

Under the **Efficient Market** assumption, CAPM is

$$E(r_t) = r_t^f + \beta\left(r_t^m - r_t^f\right)$$

Deviation from Efficient Market assumption implies the price is not in the equilibrium anymore, a friction is being introduced. Due to this friction, the α will become non-zero. May be very small but definitely $\alpha \neq 0$.

SO, CAPM UNDER THE FRICTION WILL BE

$$E(r_t) = r_t^f + \alpha + \beta\left(r_t^m - r_t^f\right)$$

THE PROBLEM HAS NOW TURNED INTO A TESTING OF HYPOTHESIS PROBLEM:

$\mathbf{H_0}$**:** $\alpha = 0$ vs $H_a : \alpha \neq 0$

H_0: $\alpha = 0$ means the stock is fairly priced and the market is efficient.

H_a: $\alpha \neq 0$ means the stock is not fairly priced and the market is not efficient.

CONSIDERING THE RELIANCE STOCK DAILY ADJUSTED CLOSE PRICE DATA

For this, first we have consider the Daily adjusted close price data of Reliance stock from 1-1-2022 to 16-12-2022 to provide a brief of my analysis. .Also Nifty 50 adjusted close price data of the same time period. And downloaded the risk-free rate(3 month Treasury Bill rate) dataset.

Downloading the data sets

```
library(tseries)
library(xts)
library(ggplot2)

start_date<-"2022-01-01"
end_date<-"2022-12-17"
stock<-get.hist.quote(instrument =
"RELIANCE.NS",start=start_date,end=end_date,
                         quote = "AdjClose",provider = "yahoo")
## time series starts 2022-01-03
## time series ends   2022-12-16

nifty50<-get.hist.quote(instrument = "^NSEI",start=start_date,end=end_date,
                        quote = "AdjClose",provider = "yahoo")

## time series starts 2022-01-03
## time series ends   2022-12-16

riskfree<-read.csv("riskfree.csv")
riskfree$Date<-as.Date(riskfree$Date)
rf<-xts(riskfree$Price,riskfree$Date)
colnames(rf)<-c("risk_free")
```

Daily adjusted close price data of Reliance stock

```
head(stock)

##            Adjusted
## 2022-01-03 2396.635
## 2022-01-04 2450.722
## 2022-01-05 2462.187
## 2022-01-06 2409.246
## 2022-01-07 2428.688
## 2022-01-10 2430.682
```

Daily adjusted close price data of Nifty 50

```
head(nifty50)

##            Adjusted
## 2022-01-03 17625.70
## 2022-01-04 17805.25
## 2022-01-05 17925.25
## 2022-01-06 17745.90
## 2022-01-07 17812.70
## 2022-01-10 18003.30
```

Risk-Free Rate dataset(3 Month Treasury Bill rate)

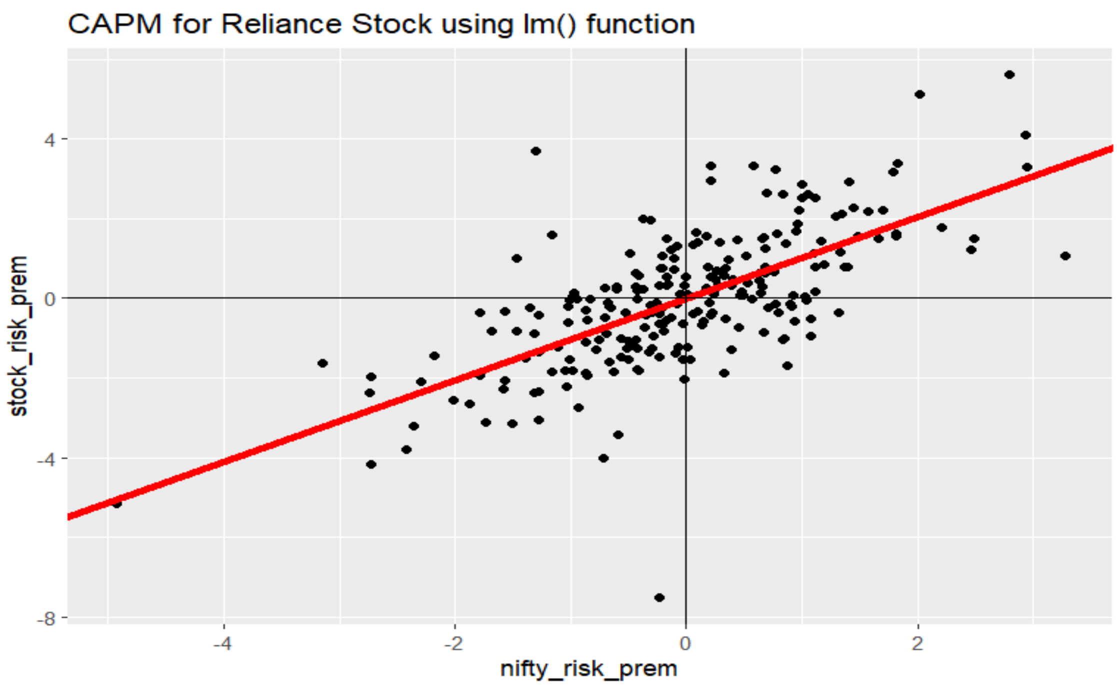
```
head(rf)

##            risk_free
## 2022-01-03      3.59
## 2022-01-04      3.60
## 2022-01-05      3.58
## 2022-01-06      3.57
## 2022-01-07      3.60
## 2022-01-10      3.59
```

Now, we calculate the log return and risk premium for Reliance stock and Nifty 50 and merge as a single data set:

```
data<-merge(nifty50,stock,rf)
data<-na.omit(data)
data$nifty_log_return=diff(log(data$Adjusted.nifty50))*100
data$risk_free=data$risk_free/100
data$stock_log_return=diff(log(data$Adjusted.stock))*100
data$nifty_risk_prem=data$nifty_log_return-data$risk_free
data$stock_risk_prem=data$stock_log_return-data$risk_free
head(data)

##            Adjusted.nifty50 Adjusted.stock risk_free nifty_log_return
## 2022-01-03         17625.70       2396.635    0.0359               NA
## 2022-01-04         17805.25       2450.722    0.0360        1.0135338
## 2022-01-05         17925.25       2462.187    0.0358        0.6716976
## 2022-01-06         17745.90       2409.246    0.0357       -1.0055808
## 2022-01-07         17812.70       2428.688    0.0360        0.3757117
## 2022-01-10         18003.30       2430.682    0.0359        1.0643476
##            stock_log_return nifty_risk_prem stock_risk_prem
## 2022-01-03               NA              NA              NA
## 2022-01-04       2.23170741       0.9775338      2.19570741
## 2022-01-05       0.46674398       0.6358976      0.43094398
## 2022-01-06      -2.17360085      -1.0412808     -2.20930085
## 2022-01-07       0.80372131       0.3397117      0.76772131
## 2022-01-10       0.08206402       1.0284476      0.04616402
```

CAPM for Reliance Stock using the lm() function

```
fit=lm(stock_risk_prem~nifty_risk_prem,data=data)
coeff<-fit$coefficients
intercept<-coeff[1]
slope<-coeff[2]
ggplot(data=data,mapping = aes(x=nifty_risk_prem,y=stock_risk_prem))+
geom_point()+
  geom_abline(intercept = intercept,slope = slope,color="red",size=1.5)+
ggtitle("CAPM for Reliance Stock using lm() function")+
  geom_vline(aes(xintercept=0))+geom_hline(aes(yintercept=0))
```




```
summary(fit)

##
## Call:
## lm(formula = stock_risk_prem ~ nifty_risk_prem, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -7.3051 -0.7555 -0.0540  0.7934  5.0180
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.01479    0.08319   0.178    0.859
## nifty_risk_prem  1.02514    0.07416  13.823   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.266 on 230 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.4538, Adjusted R-squared:  0.4514
## F-statistic: 191.1 on 1 and 230 DF,  p-value: < 2.2e-16
```

We got, α (Intercept)=0.01479 and p-value is 0.859, so at 5% level of significance we accept H_0: $\alpha=0$ i.e. Reliance stock is fairly priced.

Here, the test is conducted under the assumption $\varepsilon \sim N(0, \sigma^2 I_n)$.So, we need to check the four assumption associate with Linear Regression Model:

Assumptions:

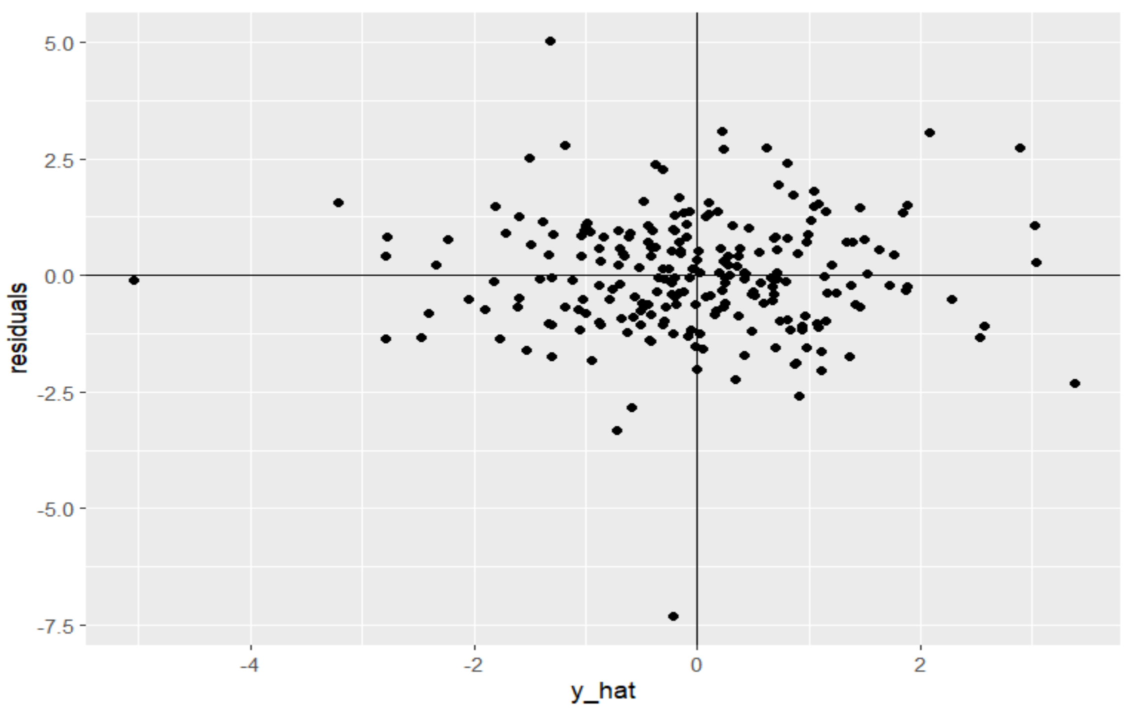
1. Linearity: Data is in linear hyper-plane.
2. Homoscedasticity: The variance of residual is same for any value of x i.e. V(ε_i)=σ^2 ∀ i
3. Independence: Cov($\varepsilon_i, \varepsilon_j$)=0 ∀ i≠j (randomness).
4. Normality: ε follows N(0, $\sigma^2 I_n$).

Checking the assumption of the linear model:

linearity assumption

```
residuals=fit$residuals
y_hat=fit$fitted.values
y_hat_res=as.data.frame(cbind(y_hat,residuals))
ggplot(data=y_hat_res,mapping = aes(x=y_hat,y=residuals))+geom_point()+
geom_vline(aes(xintercept=0))+geom_hline(aes(yintercept=0))
```

Residuals vs Predicted Return(y_hat)




```
cor(y_hat_res$y_hat,y_hat_res$residuals)

## [1] -2.589984e-17
```

I have plotted residuals vs predicted Return(y_hat) and there is no pattern and also correlation between y_hat and residual is almost 0. So, **assumption of linearity is okay i.e. data is in linear hyper plane**

Checking Homoskedasticity assumption

```
library(lmtest)

bptest(fit)

##
##  studentized Breusch-Pagan test
##
## data:  fit
## BP = 0.03277, df = 1, p-value = 0.8563
```

H_0: Homoscedasticity is present($\sigma_i^2 = \sigma^2 \; \forall \; i$)

Vs

H_1: Heteroscedasticity is present($\sigma_i^2 \neq \sigma^2$ for at least one i)
Where σ_i^2 is the variance of the residuals.

Here, I have used Breusch-Pagan test. The p-value is 0.8563. So, we accept the null hypothesis: **Homoscedasticity is present.**

Testing Randomness assumption i.e. cov(ei,ej)=0

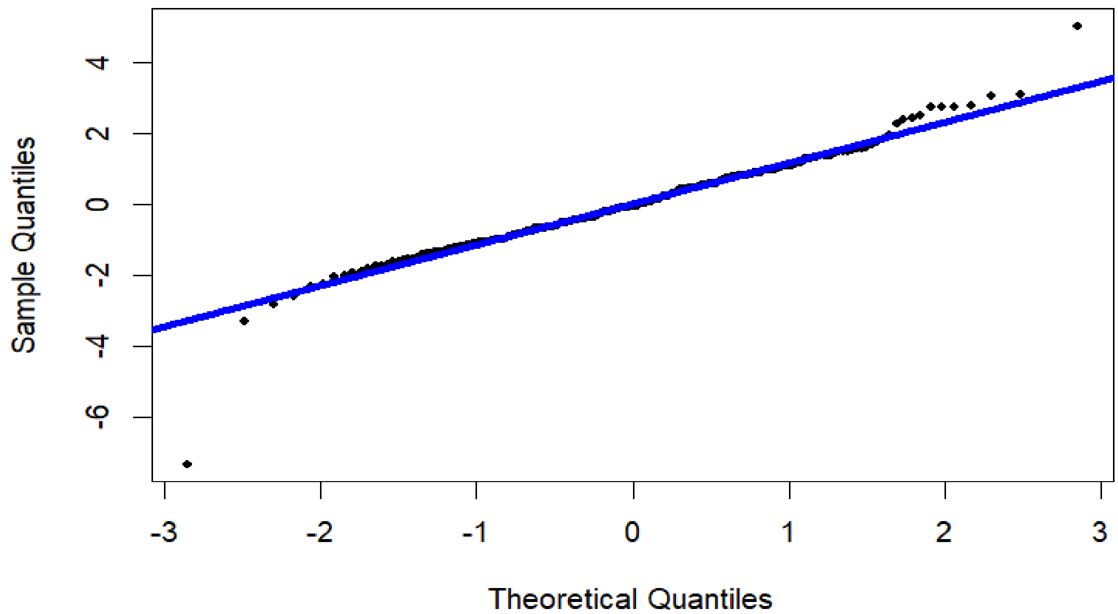
```
library(randtests)

##
## Attaching package: 'randtests'

## The following object is masked from 'package:tseries':
##
##     runs.test

bartels.rank.test(residuals)

##
##  Bartels Ratio Test
##
## data:  residuals
## statistic = -0.38343, n = 232, p-value = 0.7014
## alternative hypothesis: nonrandomness
```

I have used Bartels ratio test .The p-value is 0.7014, so at 5% level of significance we accept the Null hypothesis. So, the residuals are random i.e. **cov(ei,ej)=0**

Checking Normality assumption i.e. ε ~ Normal distribution

First we check the Normality assumption of the residuals using QQ-plot.

```
qqnorm(residuals,pch=20,main="")
qqline(residuals,lwd=4,col="blue")
```



Some points is deviating from the QQ-line. So, the residuals may not follow Normal distribution. We go for further checking

```
library(stats)
shapiro.test(as.vector(residuals))

##
##  Shapiro-Wilk normality test
##
## data:  as.vector(residuals)
## W = 0.95091, p-value = 4.414e-07
```

H_0: residuals(ε) follows Normal distribution

Vs

H_1: residuals(ε) does not follows Normal distribution

Here, I have used Shapiro-wilk test. The p-value is 4.414e-07. So, we reject the null hypothesis. So, **The residuals(ε) does not follow Normal distribution.**

BOOTSTRAPPING

In Capital Asset Pricing Model we see that the residuals(ε) does not follow Normal distribution. That's why we take the help of bootstrapping so that we don't need Normality assumption. Here, we will use **paired bootstrapping regression**

```
library(pacman)

pacman::p_load(data.table,fixest)
n=dim(data)[1]
M=10000

x=as.matrix(cbind(1,as.matrix(data$nifty_risk_prem)))
y=as.matrix(data$stock_risk_prem)
d=data.table(x=x,y=y)
d=na.omit(d)
X=as.matrix(d[,c(1,2)])
Y=as.matrix(d[,c(3)])
beta=solve(t(X)%*%X)%*%t(X)%*%Y ## OLS estimates


slope_dt=rep(0,M)  # storage for the slope estimates
intercept_dt=rep(0,M) # storage for the intercept estimates

## Resample and computing bootstrap estimates
for(i in 1:M){  # Bootstrapping M no. of time
  boot=d[sample(n,n,replace=TRUE)]  # Resample data with replacement
  boot=na.omit(boot)
  boot_x=as.matrix(boot[,c(1,2)])
  boot_y=as.matrix(boot[,c(3)])

  # computing OLS estimates for each resample data
  boot_beta=solve(t(boot_x)%*%boot_x)%*%t(boot_x)%*%boot_y
  intercept_dt[i]=boot_beta[1] # saving 10,000 intecept
  slope_dt[i]=boot_beta[2]  # saving 10,000 slope

}
boot_mean_intercept=mean(intercept_dt)
boot_mean_intercept  # bootstrap estimate of intercept(alpha)

## [1] 0.01470077

boot_SE_intercept=sd(intercept_dt)
boot_SE_intercept  # std. error for intercept(alpha)

## [1] 0.08444739

crit_val_intercept=quantile(intercept_dt,c(0.025,0.975))
crit_val_intercept  # alpha 95% CI
```

```
##       2.5%      97.5%
## -0.1540372  0.1783785
```

```
boot_mean_slope=mean(slope_dt)
boot_mean_slope  # bootstrap estimate of slope(beta)
```

```
## [1] 1.025845
```

```
boot_SE_slope=sd(slope_dt)
boot_SE_slope  # std. error for slope(beta)
```

```
## [1] 0.07600509
```

```
crit_val_slope=quantile(slope_dt,c(0.025,0.975))
crit_val_slope  # beta 95% CI
```

```
##      2.5%     97.5%
## 0.8743811 1.1766148
```

Paired Bootstrap Estimates of alpha and beta

Here, we have sum up the findings for Reliance stock

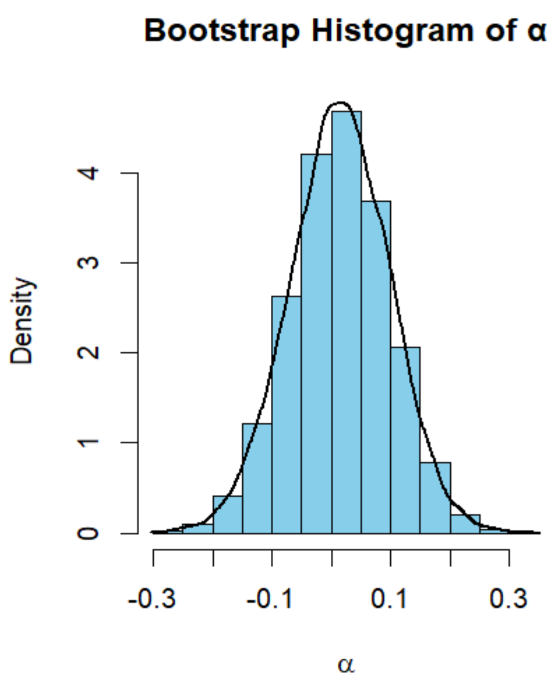
	Estimate	Std. Error	2.5%	97.5%
alpha	0.0147	0.0844	-0.1540	0.1784
beta	1.0258	0.0760	0.8744	1.1766

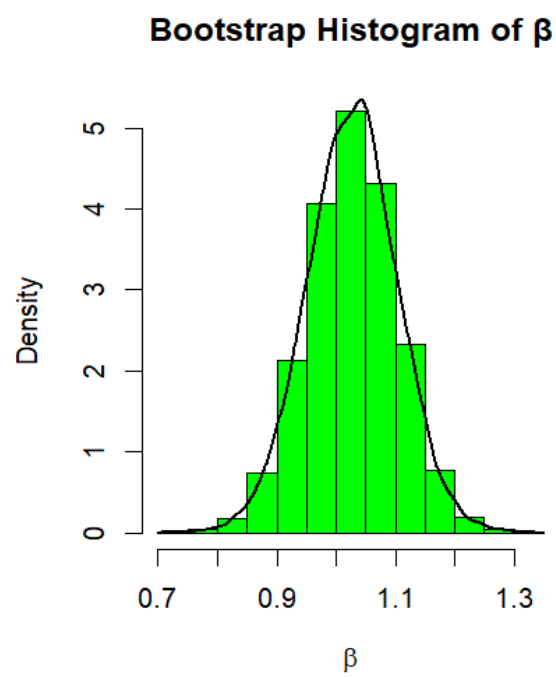
95% Confidence Interval for alpha does include zero.So, at 5% level of significance we accept the Null hypothesis i.e. the stock is fairly priced and the market is efficient.

Bootstrap Histogram of alpha and beta:

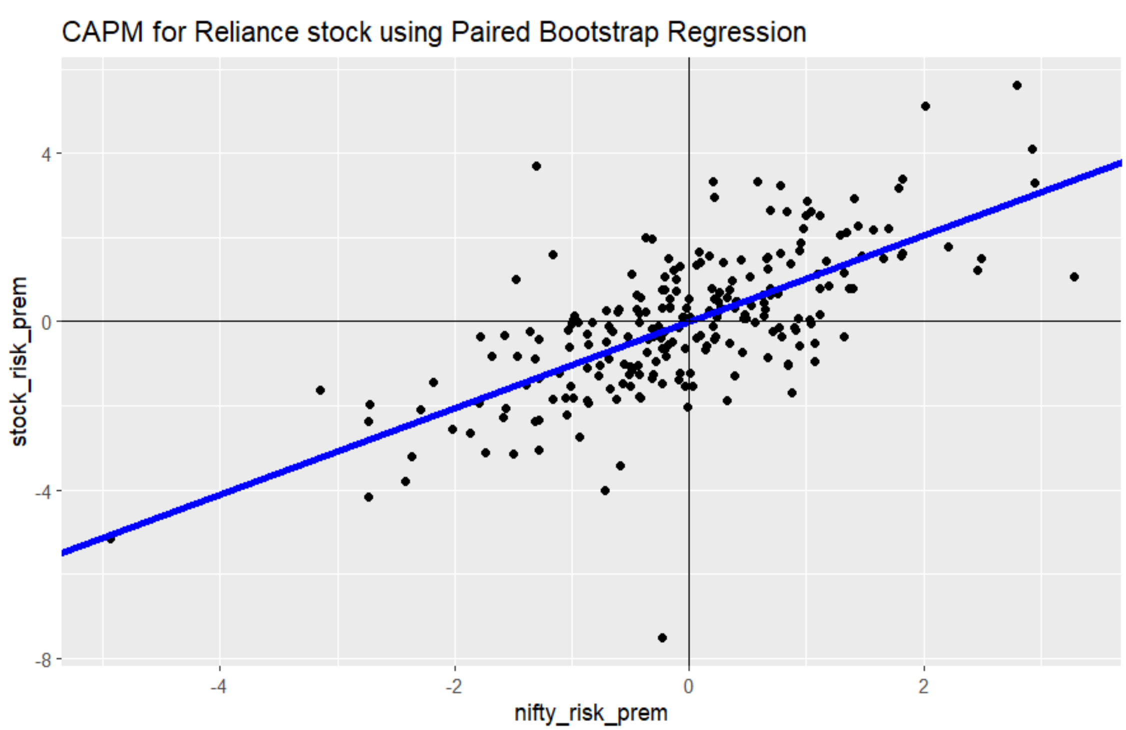
The main advantages of using Bootstrapping is we can create Histogram of Bootstrap estimates.Thus histogram provides an estimate of the shape of the distribution of Bootstrap estimates from which we can compute standard error of the estimates without having any prior idea about the sampling distribution of the estimates or any prior assumption.

```
par(mfrow=c(1,2))
hist(intercept_dt,main="Bootstrap Histogram of α", prob=TRUE,
xlab=expression(alpha),breaks=20,col="skyblue")
lines(density(intercept_dt),lwd=2)

hist(slope_dt,main="Bootstrap Histogram of β", prob=TRUE,
xlab=expression(beta),breaks=20,col="green")
lines(density(slope_dt),lwd=2)
```






```
ggplot(data=data,aes(nifty_risk_prem,stock_risk_prem))+geom_point()+
  geom_abline(slope = boot_mean_slope,intercept=boot_mean_intercept,
color="blue",size=1.5)+
ggtitle("CAPM for Reliance stock using Paired Bootstrap Regression")+
  geom_vline(aes(xintercept=0))+geom_hline(aes(yintercept=0))
```



PAIRED BOOTSTRAP R-SQUARED CONFIDENCE INTERVAL:

, R-Squared measures how closely the performance of a stock can be contributed to the performance of a selected benchmark index(ex. Nifty 50).

$$R^2 = 1 - \frac{Unexplained\ Variance}{Total\ Variance}$$

$$R^2 = 1 - \frac{Residual\ sum\ of\ square(SSE)}{Total\ sum\ of\ square(SST)}$$

$$R^2 = 1 - \frac{\sum(y_i - \hat{y})^2}{\sum(y_i - \bar{y})^2}$$

In OLS method we get one value for R^2 ,but in the paired Bootstrap method ,for each bootstrap sample we compute the R^2. So, here we get 10,000 R^2 values. Thus we can give a 95% confidence interval for R^2

```
R_squared_dt=rep(0,M)  # storage for 10,000 R-squared value
y=data$stock_risk_prem
x=data$nifty_risk_prem
y=na.omit(y)
x=na.omit(x)
y_mean=sum(y)/length(y)
for(i in 1:M)
{ # Computing predicted Y for each pair of intercept and slope
  y_hat=intercept_dt[i]+(slope_dt[i]*x)
  SSresidual=sum((y-y_hat)^2)  # Computing Residual sum of square
  SStotal=sum((y-y_mean)^2)  # computing Total sum of square
  R_squared_dt[i]=1-(SSresidual/SStotal)  # computing R-squared

}
crit_val_Rsquared=quantile(na.omit(R_squared_dt),c(0.025,0.975))
crit_val_Rsquared  ## R-Squared 95% CI
```

```
##      2.5%     97.5%
## 0.4353869 0.4536554
```

PAIRED BOOTSTRAP R-SQUARED CI = (0.435 , 0.454)

INTERPRETATION OF R^2 CONFIDENCE INTERVAL:

43.5-45.4% of the Reliance Stock movement can be explained by the movements in Nifty50 Index.

CAPM FOR TOP 10 LISTED COMPANIES IN NSE

Using the same paired bootstrap regression method, I have estimated alpha and beta of capital asset pricing model for the top 10 listed companies in NSE. Also calculated r-squared confidence interval using bootstrap method. In CAPM, r-squared is generally interpreted as the proportion of equity/stock's movement that can be explained by the movements in a market index (ex. Nifty50)

COMPANY	ALPHA	BETA	R-SQUARED 95% CI	ALPHA 95% CI	BETA 95% CI
RELIANCE	0.0147	1.0258	(0.435, 0.454)	(-0.1540, 0.1784)	(0.8744, 1.1766)
HDFCBANK	0.0259	1.1268	(0.523, 0.552)	(-0.1217, 0.1808)	(0.9400, 1.3586)
ICICIBANK	0.0594	1.0096	(0.574, 0.589)	(-0.0654, 0.1790)	(0.8861, 1.1296)
INFY	-0.0984	1.0772	(0.423, 0.449)	(-0.2742, 0.0651)	(0.8733, 1.2903)
HDFC	0.0020	1.2743	(0.559, 0.583)	(-0.1467, 0.1598)	(1.0850, 1.5109)
TCS	-0.0825	0.9046	(0.433, 0.456)	(-0.2249, 0.0627)	(0.7401, 1.0645)
ITC	0.1725	0.5853	(0.227, 0.257)	(0.0308, 0.3168)	(0.4309, 0.7279)
KOTAKBANK	-0.0113	0.9974	(0.452, 0.469)	(-0.1648, 0.1409)	(0.8518, 1.1306)
LT	0.0465	1.1024	(0.549, 0.564)	(-0.0909, 0.1876)	(0.9776, 1.2320)
HINDUNILVR	0.0396	0.7610	(0.250, 0.272)	(-0.1389, 0.2206)	(0.6076, 0.9121)

INTERPRETATION OF THE TABLE:

For example : If we consider TCS stock, we have α=-0.0825 and β=0.9046

$$E\left(\bar{r}_t{}^{tcs}\right) = -0.0825 + 0.9046 * \bar{r}_t^{nifty50}$$

As a result, CAPM tells us that TCS stock is overvalued with beta value less than 1 which means TCS stock is less volatile than the market. 43.3-45.6% of the TCS Stock movement can be explained by the movements in Nifty50 Index.

Conclusion

Here, we see how Bootstrapping rescue us when the Normality assumption of the linear model is not true. Using paired bootstrap regression we estimated alpha and beta value for CAPM models and also computed 95% Confidence Interval for R-squared.

An investor uses beta to gauge how much risk is associated with a particular stock. A conservative investor should look for a stock with lower beta ($\beta < 1$) like TCS, ITC, HINDUNILVR. In bull market, beta value greater than 1.0 will tend to produce above-average returns but will also produce larger losses in a down market. so, if someone want to take more risk and earn more then he/she should grab stock like HDFC, LT, HDFCBANK. If one wish to replicate the broader market in his/her portfolio, a beta of 1.0 would be ideal. Then he/she should hold stock like RELIANCE, KOTAKBANK, INFY. One should also see the corresponding R-squared CI for each of the fitted CAPM models to understand how closely alpha and beta reflect a stock's return as opposed to how much is random or due to other unobserved factors.

Capital Asset Pricing Model (CAPM) is the one of the broadly used method to evaluate the stock return. It clarifies the risk return relationship and is used to understand the pricing of risky stocks. The main focus is to learn how to make the highest return with the same level of risk managing portfolio risk that is based on CAPM. The result found that Capital Asset Pricing Model (CAPM) support the linear structure and it is good model for the explanation of stock return which guide the investor to take better investment decision.

Reference:

www1.nseindia.com

en.wikipedia.org/wiki/Capital_asset_pricing_model

www.investopedia.com/terms/c/capm.asp

https://tradingfuel.com/nifty-50-stock-list-2022/