Predictive Analytics
Regression and Classification
Lecture 4 : Part 1

Sourish Das

Chennai Mathematical Institute

Aug-Nov, 2020

Applications of Regression Model

- In this lecture, we will focus on some applications of **regression model**
- One of the popular application of the **regression model** is in the Statistical Finance or Quantitative Finance
- The model in Quantitative Finance is known as Capital Asset Pricing Model (CAPM).

Capital Asset Pricing Model

- The purpose of the CAPM is to evaluate if a particular asset, say a stock, is over priced or under priced or fairly priced !
- For example: Suppose the price of Reliance industry is ₹2100/-.
- As an investor, you may have the following questions?
  1. Is ₹2100 too high price? If it is already very over priced, what is the chance that it will drop?
  2. What if we invest this ₹2100 in Fixed deposit?

Capital Asset Pricing Model

- As we want to know figure out whether, the price of a stock will go up or go down; we need a base line to compare.
- Suppose $\{P_0, P_1, P_2, \cdots, P_n\}$ are the prices of stocks over a time period.
- We would be interested in change in price over period from $(t-1)$ to t, i.e.,

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} \times 100,$$

  R_t is known as simple return of stock over period $[(t-1), t]$ represents the percentage of change of stock with respect to P_{t-1}.

Capital Asset Pricing Model

- Simple Return is defined as,

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}},$$

R_t is known as simple return of stock over period $[(t-1), t]$ represents the proportion of change of stock with respect to P_{t-1}.

- log Return is defined as,

$$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right),$$

Capital Asset Pricing Model

- Simple Return can be expressed as

$$
\begin{aligned}
R_t &= \frac{P_t - P_{t-1}}{P_{t-1}} \\
&= \frac{P_t}{P_{t-1}} - 1 \\
&= e^{r_t} - 1
\end{aligned}
$$

Download data

```
> library(tseries)
> reliance<-get.hist.quote(instrument = "RELIANCE.NS"
+                          ,start="2020-07-27"
+                          ,end="2020-08-04"
+                          ,quote="AdjClose"
+                          ,provider = "yahoo")
time series ends   2020-08-03
> reliance
           Adjusted
2020-07-27  2156.20
2020-07-28  2177.70
2020-07-29  2096.65
2020-07-30  2108.85
2020-07-31  2067.10
2020-08-03  2009.00
```

Calculate Simple and log-Return

```
> ln_rt<-diff(log(reliance))
> Rt<-exp(ln_rt)-1
> cbind(ln_rt,Rt)*100
           Adjusted.ln_rt Adjusted.Rt
2020-07-28      0.9921861   0.9971246
2020-07-29     -3.7928465  -3.7218189
2020-07-30      0.5802036   0.5818900
2020-07-31     -1.9996115  -1.9797519
2020-08-03     -2.8509620  -2.8107056
```

Risk premium

- Risk premium is definded as

$$\bar{r}_t = r_t - r_f,$$

r_f is the risk-free return

- $\bar{r}_t$ is the premium that an investor earn over the return risk-free return.

Example of Risk premium

Ex Suppose you invest your money in SBI's fixed deposit and SBI gives you an interest income of 6%.

- Now instead of investing in SBI's fixed deposit scheme, suppose you invested in the equity of Reliance industries.
- Suppose the share price of reliance goes up by 15% over a period of one year. So your risk premium is 9%.
- Suppose the share price of reliance goes down by 15% over a period of one year. So your risk premium is -21%.

Risk premium of an equity vs market

- **Big Question**: Should we invest in Equity or Exchange Traded Fund (ETF)?

Note ETF's are special type of Mutual Fund. It says that its follows the market index.

- We have to model risk premium of equity as function of risk premium of market etf.
- You have either of the three instruments to invest:
  1. SBI's Fixed deposit (guranteed 6% return)
  2. Nifty 50 ETF
  3. A particular Equity, say Reliance or HDFC Bank or SBI etc.

Capital Asset Pricing Model

- CAPM for an equity can explained as:

$$(r_t - r_t^f) \quad = \quad \alpha + \beta(r_t^m - r_t^f) + \epsilon_t$$

- 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

Capital Asset Pricing Model

- CAPM for an equity can explained as:

$$\begin{aligned} (r_t - r_t^f) &= \alpha + \beta(r_t^m - r_t^f) + \epsilon_t \\ \bar{r}_t &= \underbrace{\alpha + \beta \bar{r}_t^m}_{systematic\ \ risk} + \underbrace{\epsilon_t}_{idiosyncratic\ \ risk} \end{aligned}$$

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

Capital Asset Pricing Model

- We can express it as

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

where $\mathbb{E}(\epsilon_t) = 0$, $\mathbb{V}(\epsilon_t) = \sigma^2 \ \forall \ t$ and $\mathbb{C}ov(\epsilon_t, \epsilon_{t'}) = 0$

- What α and β means?

- Suppose $\alpha = 0$, and $\beta = 1.25$, it means if market return goes up by 1% the equity return will go up by 1.25%.

- On the other hand, if means if market return goes down by 1% the equity return will go down by 1.25%.

Capital Asset Pricing Model

- We can express it as

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

where $\mathbb{E}(\epsilon_t) = 0$, $\mathbb{V}(\epsilon_t) = \sigma^2 \;\; \forall \;\; t$ and $\mathbb{C}ov(\epsilon_t, \epsilon_{t'}) = 0$

- What α and β means?

- Suppose $\alpha = 0$, and $\beta = 0.85$, it means if market return goes up by 1% the equity return will go up by 0.85%.

- On the other hand, if means if market return goes down by 1% the equity return will go down by 0.85%.

Capital Asset Pricing Model

- We can express it as

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

where $\mathbb{E}(\epsilon_t) = 0$, $\mathbb{V}(\epsilon_t) = \sigma^2 \;\; \forall \;\; t$ and $\mathbb{C}ov(\epsilon_t, \epsilon_{t'}) = 0$

- What α and β means?

- So β is a measure of systematic risk.

- Now let us try to understand what is the α?

Capital Asset Pricing Model

- We can express it as

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

where $\mathbb{E}(\epsilon_t) = 0$, $\mathbb{V}(\epsilon_t) = \sigma^2 \;\; \forall \;\; t$ and $\mathbb{C}ov(\epsilon_t, \epsilon_{t'}) = 0$

- Now let us try to understand what is the α?

- Suppose $\beta = 1$ and $\alpha = 0.01$. It means if the market return is 0%. Still the equity goes up by 0.01%. That means even if market is flat; the equity has gone up. Why?

- Because the equity is intrinsically undervalued. More purchase pressure makes the price to go up and hence a positive return for the stock; though market is flat.

Capital Asset Pricing Model

- We can express it as

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

where $\mathbb{E}(\epsilon_t) = 0$, $\mathbb{V}(\epsilon_t) = \sigma^2 \;\; \forall \;\; t$ and $\mathbb{C}ov(\epsilon_t, \epsilon_{t'}) = 0$

- Now let us try to understand what is the α?
- If $\alpha > 0$; the stock is **undervalued**.
- If $\alpha < 0$; the stock is **overvalued**.
- If $\alpha = 0$; the stock is **fairly valued**.

Capital Asset Pricing Model

- Now I am making certain assumptions. These assumptions are known as the assumptions of **Efficient Market**.

1. No insider-trading is allowed.

2. All publicly available informations are already available to evrybody and easily accessible.

3. There is no substantial tax, transaction cost, entry or exit bar.

4. No limitations on long and short positions.

Capital Asset Pricing Model

- Under these assumption; any non-zero α will be discovered very quickly by the people and they will take positions accordingly
- As a result the $\alpha \longrightarrow 0$
- Under the **Efficient Market** assumptions, CAPM is

$$\begin{aligned}\mathbb{E}(\bar{r}_t) &= \beta \bar{r}_t^m \\ \mathbb{E}(r_t - r_t^f) &= \beta(r_t^m - r_t^f)\end{aligned}$$

$$\boxed{\mathbb{E}(r_t) = r_t^f + \beta(r_t^m - r_t^f)} \tag{1}$$

Capital Asset Pricing Model

The outcome of CAPM under **Efficient Market** assumptions:

- Under the **Efficient Market** assumptions, nobody will be able to make more return consistently than the market !!
- The best you can do is your performance will be at best average and/or similar to that of **market return**.
- So why invest in individual equity? You invest in either Mutual Fund or Market ETF.

Capital Asset Pricing Model

- Let's look into the **Efficient Market** assumptions once more:

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

Capital Asset Pricing Model

- Deviation from **Efficient Market** assumptions implies the price is not in the equilibrium anymore; a friction is being introduced.
- Due to this friction; the α will become non-zero. **Maybe very small. But definitely** $\alpha \neq 0$.
- So CAPM under the friction will be:

$$\boxed{\mathbb{E}(r_t) = r_t^f + \alpha + \beta(r_t^m - r_t^f)} \tag{2}$$

Capital Asset Pricing Model

- The problem has now turned into a testing of hypothesis problem:

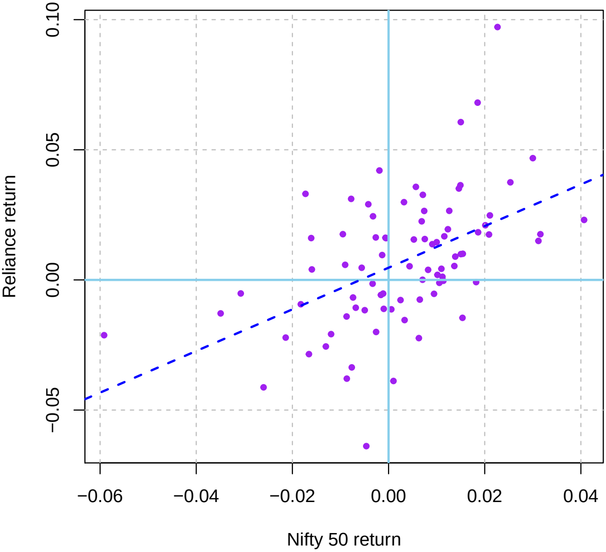
$$H_0 : \ \alpha = 0 \ \ vs \ \ H_a : \ \alpha \neq 0$$

- $H_0 : \ \alpha = 0$ 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.
- Let's look into the data.

Capital Asset Pricing Model

```
> library(tseries)
> start_date<-"2020-04-06"
> end_date<-"2020-08-06"
> rel<-get.hist.quote(instrument = "RELIANCE.NS"
+                     ,start=start_date,end=end_date
+                     ,quote="AdjClose",provider = "yahoo")
time series starts 2020-04-07
time series ends   2020-08-05
> nifty<-get.hist.quote(instrument = "^NSEI"
+                     ,start=start_date,end=end_date
+                     ,quote="AdjClose",provider = "yahoo")
time series starts 2020-04-07
time series ends   2020-08-05
> data <-merge(nifty,rel)
> rt<-diff(log(data))
> head(rt*100)
           Adjusted.nifty Adjusted.rel
```

Capital Asset Pricing Model

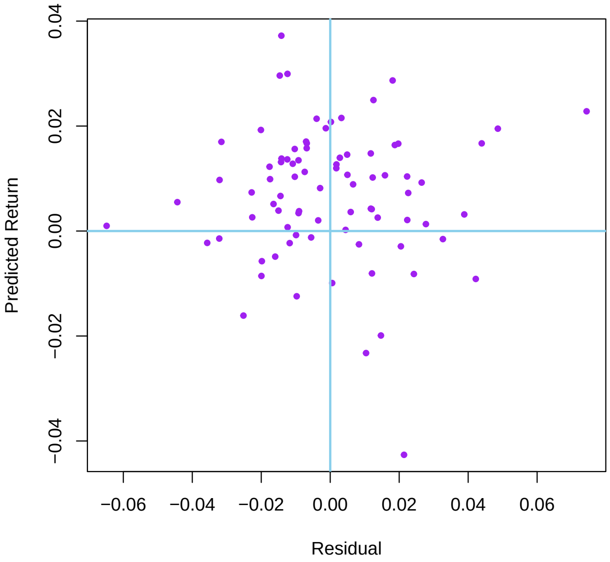


```
Call:
lm(formula = Adjusted.rel ~ Adjusted.nifty, data = rt)

Residuals:
      Min         1Q     Median         3Q        Max
-0.064845 -0.014158 -0.003207  0.012484  0.074344

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)    0.004686   0.002477   1.892   0.0622 .
Adjusted.nifty 0.799907   0.155627   5.140  1.9e-06 ***
---
Signif. codes:  0
```

Note that the test is conducted under the assumption $\epsilon \sim \mathcal{N}(0, \sigma^2 \mathbf{I}_n)$. Let's check these out



Perhaps, assumptions of linearity is okay!

Rank test for Randomness

```
> library(randtests)
> randtests::bartels.rank.test(resid)

        Bartels Ratio Test

data:  resid
statistic = -0.66282, n = 82, p-value = 0.5074
alternative hypothesis: nonrandomness
```

Looks like, assumptions of randomness is okay!

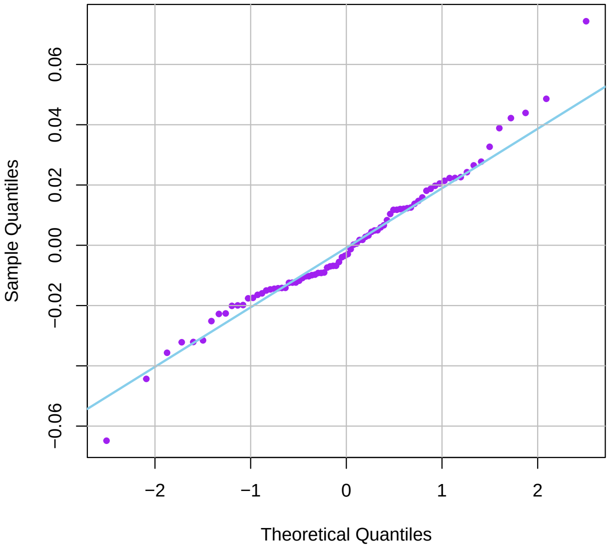
Breusch-Pagan Test against heteroskedasticity

```
> library(lmtest)
> lmtest::bptest(CAPM)

        studentized Breusch-Pagan test

data:  CAPM
BP = 0.028894, df = 1, p-value = 0.865
```

Looks like, assumptions of homoscadasticity is okay!



Is the assumptions of normality okay?

Kolmogorv-Smironov Test to check Normality

```
> stats::ks.test(resid,'pnorm')

        One-sample Kolmogorov-Smirnov test

data:  resid
D = 0.47415, p-value < 2.2e-16
alternative hypothesis: two-sided
```

Oops!! assumptions of Normality is not correct! What to do?

In the next part...

- We will discuss how we can use the nonparametric Bootstrap Regression to slavage the situation of Non-Gaussian distribution.

Thank You

sourish@cmi.ac.in