Regression

21st Oct, 2019

What if Analysis

"Stress Testing."

How long can you survive if 1000s of customers show up?

In REGRESSION,

$$y = \beta_0 + \beta_1 x + \varepsilon$$

$\uparrow x$ by 1% → How much $\uparrow$ in y?

$\beta_1 \times 100\%$.

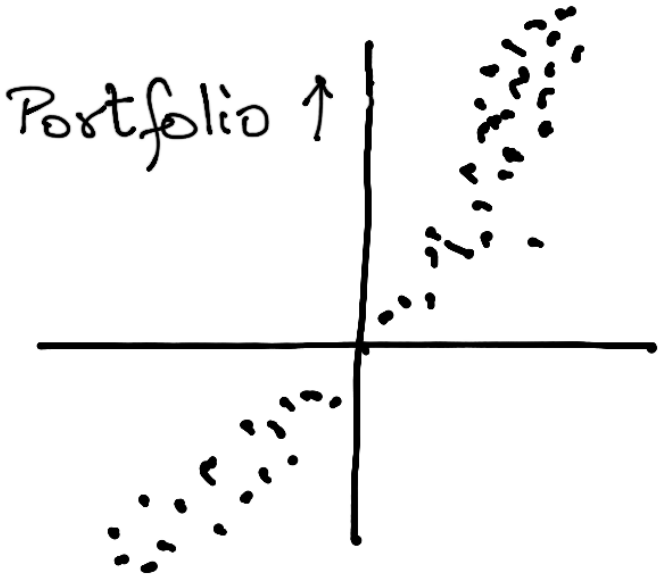
can be determined easily $\left[\begin{array}{l} x_0 \rightarrow y_0 \\ x_1 = 1.01\, x_0 \\ \rightarrow y_1 \end{array}\right.$

What if Analysis helps in case of black box models.

Consider GP Regression

$$y = f(x) + \varepsilon \quad , \quad f \sim GP(\mu(x), \Sigma(x, x'))$$

Slope varies. Hence change in y due to a change in x cannot be easily determined.



Asset Pricing Model tells you to fit linear regression

But in 1ST Quadrant market return is generally not linear.

Hence GP regression gives a more localised estimate.

How does it work for Multiple Linear Reg?

$$y = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip} + \varepsilon \quad i = 1, \ldots, n$$

$$y = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \cdots + \hat{\beta}_p x_p$$

Ex2

y = Salary

x_1 = Gender

x_2 = Education

x_3 = Race / Caste / Religion

x_4 = School

FOCUS ON ONE VARIABLE AT A TIME.

	x_1	x_2	x_3	x_4		
Candidate 1	(M	DS	WC	MIT)	x_0	y_0
Candidate 2	(F	DS	WC	MIT)	x_1	y_1

Check if assumptions are statistically significant.

BOOTSTRAP

- Draw many many samples
- Check salary for each gender
- Check if salary faces gender discrimination

Change variables to see how the output changes.

So, what is What if analysis ??

Google at your own risk.