

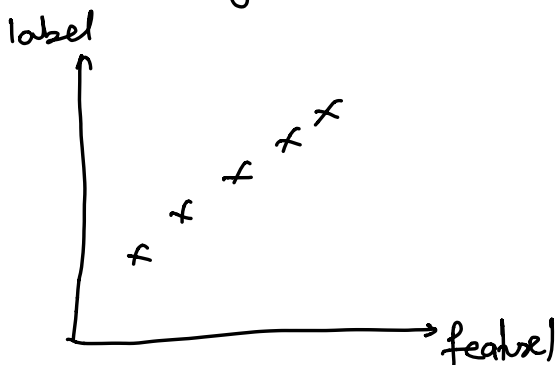
Hypothesis Testing

Doing an statistical assumption and checking whether the assumption is correct or not.

Regression based problems

Feature: any type (numerical, categorical, ordinal)

Label: numerical.



Hypothesis Testing

50_startups.csv

features: R&D, adm, mark, state

label: profit

If you want to have a good model to deploy then there must exist linear relationship between feature & label.

Hypothesis Test : R&D spend } - if they have
Profit } linear relationship

(null hypo) $\Leftrightarrow H_0$ - 0 wrong

(alt hypo) $\Leftrightarrow H_1$ - 1 correct

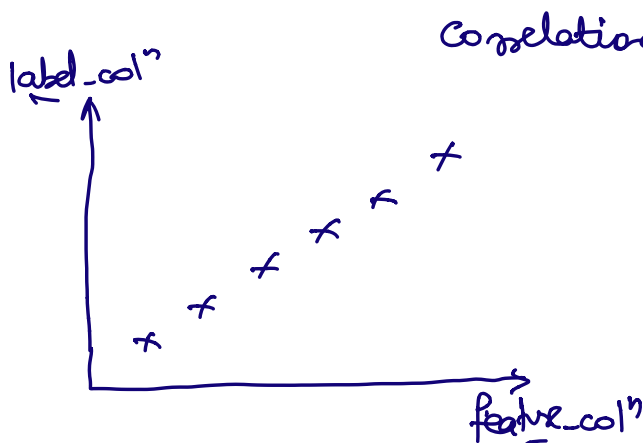
Hypothesis Test : R&D spend } - if they have
Profit } linear relationship

(null hypo) $\Leftrightarrow H_0$ - 0 wrong
(alt hypo) $\Leftrightarrow H_1$ - 1 correct

correlation test

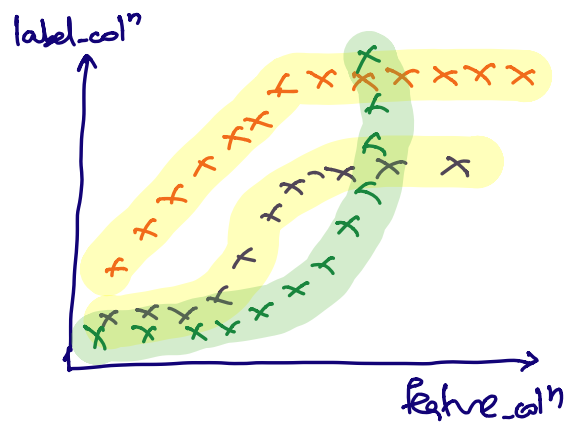
- Pearson's corr
- Spearman corr
- Kendall's Rank
- Chi-Square Test

} → fns
scipy



Linear relationship

- x Simple LR
- x Multiple LR



monotonic relationship

- x polynomial