

What is Regression Analysis?

A form of **predictive modelling technique** which investigates the relationship between a dependent (*target*) and independent variable(s) (*predictor*)

- Shows **significant relationships** btw. target and predictors
- Shows **strength of impact** of multiple predictor on a target

Coefficients

Mean change in the response variable for **one unit of change** in the predictor variable while holding other predictors in the model constant.

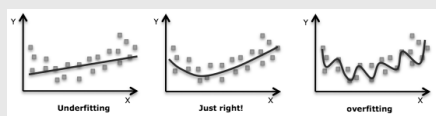
Think of them as slopes

Polynomial Regression ($y=a+b*x^2...$)

Curve that fits

Higher degree polynomial -> over-fitting risk

Polynomial Regression



Multicollinearity

Multicollinearity test: Variance inflation factors (VIF ≥ 5)

Increase the variance of the coeff. estimates (makes them sensitive)

Stepwise regression does not work as well

Doesn't affect the overall fit of the model

Doesn't produce bad predictions

SOLUTIONS

- Standardized predictors
- Removing highly correlated predictors
- Linearly combining predictors (x.e. sum)
- Different analyses: PLS or PCA

Stepwise Regression

Maximize prediction power with minimum number of predictor variables

Fits the regression model by adding/dropping co-variables one at a time

- **Standard stepwise regression** adds and removes predictors as needed for each step.

- **Forward selection** starts with most significant predictor and adds variable for each step.

- **Backward elimination** starts with all predictors and removes the least significant variable for each step.

Regularized Linear Models (Shrinkage)

Regularize linear model through constraining the weights

Regularized term added to cost function. Learning algorithm not only fits data but keeps model weights as small as possible.

Ridge (L2) | Lasso (L1) | ElasticNet (L1 & L2)

Ridge Regression

L1: adds penalty equivalent to **squ. of the magnitude** of coefficients

Minimization = LS Obj + α * (sum of squ of coefficients)

It shrinks the value of coefficients but doesn't reaches zero

Lasso Regression

L1: adds penalty equivalent to **abs. value of the magnitude** of coefficients

Minimization = LS Obj + α * (sum of abs value of coefficients)

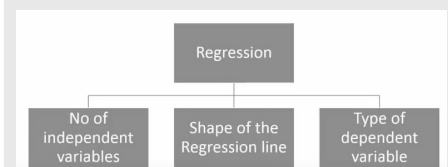
LS Obj - Least Squares objective

ElasticNet Regression

Ridge and Lasso: 'r' controls de mix ratio.

$$r*\lambda*\sum(\beta^2) + (1-r/2)*\lambda*\sum(\text{abs}(\beta))$$

Regression Types



Techniques are mostly driven by three metrics

Linear Regression ($Y=a+b*X + e$)

Straight line (regression line)

Least Square Method to best fit line

Linear relationship between predictors and target

CONS

Multicollinearity, autocorrelation, heteroskedasticity

Very sensitive to Outliers

Logistic Regression

Target binary (0/ 1): binomial distribution

Logit function

widely used for classification problems

can handle various types of relationships because it applies a non-linear log transformation to the predicted odds ratio

maximum likelihood estimates

Requires large sample sizes

Ordinal Target -> Ordinal logistic regression

Multiclass Target -> Multinomial Logistic regression.



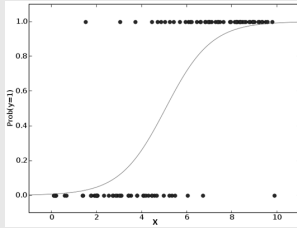
By **isantabarbara**
(isantabarbara)

cheatography.com/isantabarbara/
www.google.com

Not published yet.
Last updated 6th April, 2018.
Page 1 of 2.

Sponsored by **CrosswordCheats.com**
Learn to solve cryptic crosswords!
<http://crosswordcheats.com>

Logistic Regression



Logistic Regression

odds = $p / (1-p)$ #event prob / not event prob

$\ln(\text{odds}) = \ln(p / (1-p))$

$\text{logit}(p) = \ln(p / (1-p)) = b_0 + b_1X_1 + b_2X_2 + b_3X_3 \dots + b_kX_k$

p is the probability of presence of the characteristic of interest.



By **isantabarbara**
(isantabarbara)

cheatography.com/isantabarbara/
www.google.com

Not published yet.
Last updated 6th April, 2018.
Page 2 of 2.

Sponsored by **CrosswordCheats.com**
Learn to solve cryptic crosswords!
<http://crosswordcheats.com>