Regression and Time Series Model

-Gaurav Suryawanshi

# Simple Linear Regression

|  |  |
| --- | --- |
| Heading | Description or value |
| Simple Regression Formula |  |
| Regression Coefficients | and |
| Expectations and variance |  |
| Cost function |  |
| Value of coefficients, beta\_0 and beta\_1 |  |
| Sxx and Sxy |  |
| Variance of coefficients |  |
| Residual |  |

## Best Estimators

|  |  |
| --- | --- |
| Heading | Description or value |
| SSres |  |
| SSt |  |
| Expected value of SSres |  |
| Unbiased Estimator of Variance or Residual Mean Square Average MSres |  |
| Standard Error in Regression |  |

## Hypothesis Testing

|  |  |
| --- | --- |
| Heading | Description or value |
| Standard t-test |  |
| t-test, NID stands for Normal Independent |  |
| Modified t-test as sigma is unknown | Follows distribution  Dof=Dof(=n-2  Reject Null Hypothesis if |
| t-test for Intercerpt |  |
| Standard error |  |
| ANOVA Test. SS\_r stands for Regression | Reject null hypothesis if |
| Expectation of MSres |  |

## Interval Testing

|  |  |
| --- | --- |
| Heading | Description or value |
| 100(1-) Confidence Interval of |  |
| Confidence Interval in sigma |  |
| Estimation of mean Response |  |
| Sampling Distribution with |  |
| Confidence Interval of mean Response |  |
| Prediction of new values interval |  |
| Coefficient of Determination |  |

## Maximum Likelihood Estimators

|  |  |
| --- | --- |
| Heading | Description or value |
| Generic Function |  |
| Estimators Here Variance estimator is biased. |  |

## Jointly Distributed Model

|  |  |
| --- | --- |
| Heading | Description or value |
| Generic Function | is correlation coefficient |
| Maximum Likelihood Parameters r is for is the measure of linear association b/w y & x |  |
| Hypothesis Testing for correlation |  |

# Multiple Linear Regression

|  |  |
| --- | --- |
| Heading | Description or value |
| Generic Function |  |
| Vectors |  |
| Method Of least Squares. |  |
| Sxx |  |
| Covariance Defination |  |
| Properties of LS Operators |  |
| Variance property in vector where A is const |  |
| 04Estimation of |  |

## Maximum Likelihood Estimators

|  |  |
| --- | --- |
| Heading | Description or value |
| Generic Function |  |
| Variance Estimate |  |

## Hypothesis Testing

|  |  |
| --- | --- |
| Heading | Description or value |
| Best Estimates |  |
| Hypothesis (ANOVA) | for at least one  If values of is large then it’s likely that is true Reject if |
| Variance Estimate |  |

## Hypothesis Testing on Individual Variables

|  |  |
| --- | --- |
| Heading | Description or value |
| Basic Hypothesis | If is not rejected, we can delete the regression variable  Where is the diagonal element of and  Reject if |
| Testing for a set of Regressor Variables | where is and is is the full model  is the reduced model  Here Fo follows non central F distribution with non-centrality parameter |
| Testing General Linear Hypothesis | Here only r equations out of m in TB=0 are independent.  Let we have a reduced model using r equations  Get  Then  Reject if  Or |
| Testing Equality of Regression model | Let , we want to check if  Then obtain a reduced model such that  Then use  And apply t-test |

# Residual Analysis

|  |  |
| --- | --- |
| Heading | Description or value |
| R2 test |  |
| Basic Residuals |  |
| Standardized Residuals | Large values tend to be outliers |
| Studentized Residuals |  |
| PRESS/Jack Knife Residuals | Use where it is fitted value of response based on all obs but (i) |
| R student | Estimate with data removed |

# Time Series

## Definitions And Basics

|  |  |
| --- | --- |
| Heading | Description or value |
| Definition | A time series is generated from uncorrelated variables with 0 mean and fixed variance called white noise |
| Implementation | Where  =Time series at time t=t  =White noise, random error added at deterministic point  =Seasonal or repetition trend |
| Mean  Auto Covariance Function |  |
| Auto Covariance (ACVF) | Auto Covariance (t1, t2) =  Where h is the time period of seasonality |
| Auto Correlation (ACF) | Auto Correlation (t1,t2)=  here |
| Properties of ACF | R is a p.s.d |
| Positive Semidefinite Function |  |

### Weakly and Strongly Stationary

|  |  |
| --- | --- |
| Heading | Description or value |
| Weakly Stationary | 1. is independent of t 2. is independent of each h   Usually implies there is no trend in the series |
| Strongly Stationary | If joint distribution of and are same i.e |

## Different types of Time series

### Random Walk

|  |  |
| --- | --- |
| Heading | Description or value |
| Formulae |  |
| Mean and ACVF | here t<s |

### Linear Process

|  |  |
| --- | --- |
| Heading | Description or value |
| Formulae [WN process]  Weakly stationary  If normally distributed -> Strongly Stationary | , |
| Mean and ACVF |  |
| Generic Linear Process |  |
| Mean and ACVF |  |
| Barlett’s Formula about | If  Then |

### Auto Regression

|  |  |
| --- | --- |
| Heading | Description or value |
| Formulae [AR(1)] | Where & |
| Mean and ACVF |  |
| Formulae [AR(p)] | Where & |

### Moving Average (MA) Process

|  |  |
| --- | --- |
| Heading | Description or value |
| Formulae (MA (1))  Stationary in nature |  |
| Mean and other stats |  |
| ACF |  |
| Formulae (MA (q)) |  |

### Auto-Regressive Moving Average (ARMA) Process

|  |  |
| --- | --- |
| Heading | Description or value |
| Formulae (ARMA (p,q))  Stationary in nature |  |
| ARMA(1,1) |  |
| Mean and Variance |  |
| ACVF ARMA(1,1) |  |
| ARIMA Process | is ARIMA(p,d,q) if |
| Correlation |  |

## Trend Estimation

|  |  |
| --- | --- |
| Heading | Description or value |
| Estimation of Trend in Absence of Seasonality |  |
| Moving average method |  |
| Exponential Smoothing Method |  |
| Poly Fit |  |
| Estimation of Trend & Seasonality |  |
| Estimation of Seasonality | First Estimate then, use  Where jth period and k=1 to d  Then |
| Sample Auto Covariance Function (h is called lag) |  |
| Sample Variance |  |
| Sample Mean |  |
| Estimation of Future values |  |
| Durban Levinson Algorithm |  |
| Innovation algorithm |  |

## Causality, Invertibility, Pacf And Model Accuracy

|  |  |
| --- | --- |
| Heading | Description or value |
| Causality (A process is casual if \_) | Here B is the backshift operator and |
| Invertibility (A process is invertible if \_) | Here B is the backshift operator and |
| Generic PACF | If normally distributed then |
| PACF in Time series | Where |
| Akaike Information Criterion | Where k is the number of estimated parameters  is the maximum likelehood  Minimum is the best |
| Bayesian Information Criterion | Lower is prefered |