

properties of OLS Matrix

Sum of Squared Residuals

$$(y - X\hat{\beta})'(y - X\hat{\beta})$$

$$y'y - \hat{\beta}'X'y - y'X\hat{\beta} + \hat{\beta}'X'X\hat{\beta}$$

$$y'y - 2\hat{\beta}'X'y + \hat{\beta}'X'X\hat{\beta}$$

Minimise the SSR

$$\partial(SSR)/\partial\hat{\beta} = -2X'y + 2X'X\hat{\beta} = 0$$

from the minimum we get: "normal equation"

$$(X'X)\hat{\beta} = X'y$$

Solve for OLS estimator $\hat{\beta}$; by pre multiplying both sides by $(X'X)$

$$(X'X)^{-1}(X'X)\hat{\beta} = (X'X)^{-1}X'y$$

by definition, $(X'X)^{-1}(X'X) = I$

$$I\hat{\beta} = (X'X)^{-1}X'y$$

$$\hat{\beta} = (X'X)^{-1}X'y$$

Properties

The observed values of X are uncorrelated with the residuals.

$X'e = 0$ implies that for every column x_k of X, $x_k'e = 0$.

substitute in $y = X\hat{\beta} + e$ into normal equation

$$(X'X)\hat{\beta} = X'(X\hat{\beta} + e)$$

$$(X'X)\hat{\beta} = (X'X)\hat{\beta} + X'e$$

$$X'e = 0$$

The sum of the residuals is zero.

If there is a constant, then the first column in X (i.e. X_1) will be a column of ones. This means that for the first element in the $X'e$ vector (i.e. $X_{11}e_1 + X_{12}e_2 + \dots + X_{1n}e_n$) to be zero, it must be the case that $\sum e_i = 0$.

The sample mean of the residuals is zero.

$$\bar{e} = \sum e_i / n = 0.$$

properties of OLS Matrix (cont)

The regression hyperplane passes through the means of the observed values (X and y).

This follows from the fact that $e = 0$. Recall that $e = y - X\hat{\beta}$. Dividing by the number of observations, we get $\bar{e} = \bar{y} - \bar{x}\hat{\beta} = 0$. This implies that $\bar{y} = \bar{x}\hat{\beta}$. This shows that the regression hyperplane goes through the point of means of the data.

The predicted values of y are uncorrelated with the residuals.

$$\hat{e}'e = (X\hat{\beta})'e = b'X'e = 0$$

The mean of the predicted Y's for the sample will equal the mean of the observed Y's: $\bar{y} = \bar{y}$

The Gauss-Markov Theorem: Proof

that $\hat{\beta}$ is an unbiased estimator of β

$$\hat{\beta} = (X'X)^{-1}X'y = (X'X)^{-1}X'(X\beta + \epsilon)$$

$$\hat{\beta} = (X'X)^{-1}X'\epsilon$$

given $(X'X)^{-1}X'X = I$

$$E[\hat{\beta}] = E[\beta] + E[(X'X)^{-1}X'\epsilon] = \beta + (X'X)^{-1}X'E[\epsilon]$$

where $E[X'\epsilon] = 0$

$$E[\hat{\beta}] = \beta$$

Proof that $\hat{\beta}$ is a linear estimator of β .

$$\hat{\beta} = \beta + (X'X)^{-1}X'\epsilon; \text{ where } (X'X)^{-1}X' = A$$

$$\hat{\beta} = \beta + A\epsilon \Rightarrow \text{linear equation}$$

Heteroskedasticity

consequence: the statistics used to test hypotheses under Gauss-Markov assumptions are not valid in the presence of heteroskedasticity.

Valid estimator (any form)

$$\sum [(x_1 - \bar{x})^2 u_i^2] / [SST_x]$$

$$SST_x = \sum (x_1 - \bar{x})^2$$

Robust Standard error

$$\text{Var}(\hat{\beta}_j) = \sum [r_{ij}^2 \hat{u}_i^2] / [SSR^2]$$



Inference	
Normality Assumption:	zero mean and Variance $Var(u) = \sigma^2$
T-test:	$(\hat{\beta}_j - \beta_j) / se(\hat{\beta}_j) \sim t_{n-k-1} = df$
$H_0: \beta_j = 0$	used in testing hypotheses about a single population parameter as in .
Test statistic	$t_{\hat{\beta}_j} = (\hat{\beta}_j - \beta_j) / se(\hat{\beta}_j) \sim t_{n-k-1}$ $t = (\text{estimate} - \text{hypothesised value}) / \text{standard error}$
Alternative Hypothesis/one sided	
$H_1: \beta_j > 0$	$t_{\hat{\beta}_j} > c [c @ 5\%]$
$H_1: \beta_j < 0$	$t_{\hat{\beta}_j} < -c [c @ 5\%]$
Two sided	
$H_1: \beta_j \neq 0$	$ t_{\hat{\beta}_j} > c [c @ 2.5\%]$
If H_0 , rejected	x_j is statistically significant, (significantly different from zero), @ the 5% level
if H_0 , not rejected	x_j is statistically insignificant @ the 5% level
P-value	smallest significant level at which the null hypotheses would be rejected
Confidence Interval	
	$\hat{\beta}_j \pm c \cdot se(\hat{\beta}_j)$ where c is 97.5 percentile in a t_{n-k-1} distribution
CI given; @ 5% significant level	$H_0: \beta_j = a_j$ is rejected against $H_1: \beta_j \neq a_j$; if a_j is not in the 95% confidence interval
$H_0: \beta_1 < \beta_2 \Leftrightarrow \beta_1 - \beta_2 < 0$	$t = (\hat{\beta}_1 - \hat{\beta}_2) / se(\hat{\beta}_1 - \hat{\beta}_2)$
$se(\hat{\beta}_1 - \hat{\beta}_2) = \sqrt{Var(\hat{\beta}_1 - \hat{\beta}_2)}$	$Var(\hat{\beta}_1 - \hat{\beta}_2) = Var(\hat{\beta}_1) + Var(\hat{\beta}_2) - 2Cov(\hat{\beta}_1, \hat{\beta}_2)$
alternative to calculating $se(\hat{\beta}_1 - \hat{\beta}_2)$	Let $\theta = \beta_1 - \beta_2$; $\beta_1 = \theta + \beta_2$

Inference (cont)	
$H_0: \theta = 0, H_1: \theta < 0$	Substituting $\beta_1 = \theta + \beta_2$ into the model we obtain
$\beta_0 + \theta x_1 + \beta_2(x_1 + x_2) + \beta_3 x_3 + u$	
F Test	$F = [(SSR_r - SSR_{ur}) / q] / [SSR_{ur} / (n - k - 1)]$
q	= number of restrictions
$n - k - 1 = df_{ur}$	= $df_r - df_{ur}$
R^2 F stat	$SSR = SST(1 - R^2)$ $F = [(R^2_{ur} - R^2_r) / q] / [1 - R^2_{ur} / (df_{ur})]$
remember to not square the R value thats already been done	
Overall significance of the regression	
Testing joint exclusion	$[R^2 / R] / [(1 - R^2) / (n - k - 1)]$

Data Scaling	
Changes:	
if X_j is * by c	Its coefficient is / by c
If dependant variable is * by c	ALL OLS coefficients are * by c
neither t nor F statistics are affected	
Beta coefficients	obtained from an OLS regression after the dependant and independent variables have been transformed into z-scores

Dummy Variables	
Dummy/Binary Variables	= yes/no variables
= take on the values 0 and 1 to identify the mutually exclusive classes of the explanatory variables.	
= leads to regression models where the parameters have very natural interpretations	
Given: $wage = \beta_0 + \beta_1 \text{female} + \beta_2 \text{edu} + u$	
To solve for ∂_0 :	
$\partial_0 = E(wage \text{female}, \text{edu}) - E(wage \text{male}, \text{edu})$	
where level of education is the same	
Graphically ∂_0	an intercept shift
= male intercept = β_0	



Dummy Variables (cont)

female intercept= $\beta_0 + \alpha_0$

dummy variable trap= when both dummy variables (male & female) are included; resulting in perfect collinearity

If a qualitative variable has m levels; then (m-1) dummy variables are required and each of them takes value 0 and 1.

Hypothesis test

Test whether the two regression models are identical:

$$H_0 : \beta_2 = \beta_3 = 0$$

$$H_1 : \beta_2 \neq 0 \text{ and/or } \beta_3 \neq 0.$$

Acceptance of H_0 indicates that only single model is necessary to explain the relationship.

Test is two models differ with respect to intercepts only and they have same slopes

$$H_0 : \beta_3 = 0$$

$$H_1 : \beta_3 \neq 0.$$

Treating a **quantitative variable** as qualitative variable increases the complexity of the model.

The degrees of freedom for error are reduced.

Can effect the inferences if data set is small



By **dsjac3**
cheatography.com/dsjac3/

Not published yet.
Last updated 31st October, 2016.
Page 3 of 3.

Sponsored by **ApolloPad.com**
Everyone has a novel in them. Finish Yours!
<https://apollopad.com>