

Hypothesis Testing: Mathematics

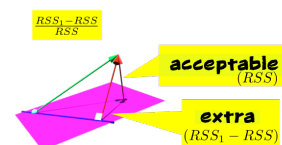
Let say,

RSS = residual sum of squares in the base model (first model, the model which we know is a good fit) which is acceptable,

RSS_1 = Residual sum of squares in restricted model.

Now we want to see whether $(RSS_1 - RSS)$ is significantly larger compared to RSS . We have to compare this deviation relatively.

So, the most intuitive test statistic should be $\frac{(RSS_1 - RSS)}{RSS}$. Now, if this value is large then we believe the extra error is too much and if it is small then we may accept the restricted model.



Now, if we carry out the idea of *likelihood ratio test* then after some linear algebraic jugglery we can find that our intuitive test statistic is precisely the same test statistic what we get theoretically.

Now what is the distribution of this statistic, specially under null hypothesis? It is needed to find P-value, critical values etc. In general it is not easy problem, but under the Gaussianity assumption we find that it has a scaled version of F distribution (if we divide by suitable degrees of freedom it becomes F statistic), under null hypothesis it will be a central F and under alternative it will be a non-central F.