**F statistics for robust sample- and gene-wise tests**

The robust F-test is conducted similarly to the standard F-test by comparing two nested models. Then the test is made robust by adding a variance inflation factor Δ to the standard F-statistics:

For example, in the gene-wise test, two linear regression models are applied to the moment estimate values (separately for the mean, variance, skewness, and kurtosis): the full model contains batch as a predictive factor, whereas the reduced model contains only a constant intercept. The robust F-statistic is computed as follows with a suitably selected Δ:

RSS0: Residual Sum of Squares for the reduced model

RSS1: Residual Sum of Squares for the full model

df1= #batch; df0= 1; n= #genes \* #batch

The Δ in the equation is selected heuristically as the total degrees of freedom times the squared 5% of the average moment estimates. This variance inflation factor reduces the statistical significance when the effect size of the batch is much smaller than the expression level of the normalized data.