|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Assumption | | | Implications | | | Assessment | | |
| Name | Mathematical Representation | Description |  | | | Evidence – Tool | How to Use | Mitigation |
| MLR1  (Gauss-Markov 1) | + | Linearity in Coefficients.  Weak assumption. | OLS Estimators Consistent. For large samples, | OLS Estimators Best Linear Unbiased Estimators (BLUE). | ,  The OLS coefficients are normally distributed, so you can conduct t-tests, etc.   on the values of the coefficients. | The model is linear. Input variables can be of any form: log, polynomial, etc. only coefficients are assumed to be linear. | | |
| MLR2  (Gauss-Markov 2) |  | Random sample of data. Independent and Identically Distributed (iid). Strong. | Discern from sampling technique as described by researcher.  Check resid\_v\_leverage | Look for outliers with influence and leverage. Cook’s distance lines. | Look hard at those points. Exclude carefully with strong evidence of errant data only. |
| MLR3  (Gauss-Markov 3) | There is no | No perfect co-linearity.  Weak assumption. | VIF Variance Inflation Factor  vif(model) | If VIF is close to 1, you have multi-colinearity. | Remove one of the variables from the model. |
| MLR4  (Gauss-Markov 4) |  | Zero conditional mean for errors. A strong assumption. |  |  |  |
| MLR4’ |  | All X uncorrelated with error. | Residuals v. Fitted values.  Scale-location plot.  plot(model1) gives both | Show that exogeneity is realistic. This preserves consistency of estimators. | Claim an assumption of exogeneity and that model is associative and not causal. Just looking gor best fit line. |
| MLR5  (Gauss-Markov 5) |  | Homoskedasticity of errors. Constant error variance across all x. A strong assumption. |  | Residuals v. Fitted values.  Scale-location plot.  plot(model1) gives both  Breusch-Pagan Test  bptest(model1)  White Test | R v FV: look for different width of bands as a function of x.  BP: Regresses residuals on Xs and does F-test on model.  White: regresses error residuals on fitted values. | Use White standard errors, which are robust to homoskedasticity.  coeftest(model3,   vcov = vcovHC)  (se.model1 = sqrt(diag(vcovHC(model1)))) |
| MLR6 CLM  (Classical Linear Model) |  | Errors are normally distributed. |  |  | Hist(residuals)  Q-Q plot – look for perfect diagonal line, no tails  Shapiro-Wilk test of normality of residuals | hist(model3$residuals)  shapiro.test(model1$residuals) | Large data set: rely on asymptotics of OLS est. Need 1-5.  Small data sets: transform skewed variable, try x^2. If resi\_v\_fitted shows curvature have zcm violation too.  Bootstrap errors. |