STAT 331

1.Summary

2.Model selection

1. use pair plots (y, continuous variable xi)

2. forward/backward/step selection (Model V, P10)

NA 的数值用平均值代替, 如果是分类变量的话，可以用平均数值或者其他方法代替

3. ASA 越小模型越好 & PRESS statistic 越小越好（3.2 & 3.3 计算方法p18）

Hii = H = (X’X)^-1X’

3.Model diagostics

Residual vs fitted value plot

Ei = yi – mui\_hat 验证:1. Mean of zero, E(ei)>0 for I = 1,2,3,,,,n

2.constant variance, car(ei) = mu\_hat^2 for…

3. independent Cov(ei,ej) = 0 for all I nq j

Standardized residual ri = ei / sigema\_hat. ~(mle, approximately N(0,1)) （相片1）

Standentized residual ri\_hat =

Normal qq-plot anallity, about assumption of normality

搜搜residual plot to prove assumption in MLRM

3.model diagonostics

1.residual vs fitted value

2.normal qq-plot

1. Outliers

Residual 特别大，就是outliers 有两个标准

1. Leverage points(with high leverage)(x 超出范围) 有两个标准 （相片2）（residual module IV p4）
2. Inplnential points