Simple tests

$$\beta \in \hat{\beta} \pm s.e.(\hat{\beta}) \cdot t_{crit}$$
$$F_{k,n-k-1} = \frac{SSE/k}{RSS/(n-k-1)} = \frac{R^2/k}{(1-R^2)(n-k-1)}$$
$$t = \frac{\hat{\beta}_2 - \beta_2^0}{s.e.(\hat{\beta}_2)}, df = n-2$$
$$F_{q,n-k} = \frac{(SSR_r - SSR_{ur})/q}{SSR_{ur}/(n-k)} = \frac{(R_{ur}^2 - R_r^2)/q}{(1-R_{ur}^2)/(n-k)}$$
r - restricted, ur - unrestricted,

q - number of linear restrictions on coefs
$$\chi^2(1) = \frac{n}{2}\log\frac{SSR_2}{SSR_1}$$
Chow breakpoint test
$$F_{k(m-1),n-mk} = \frac{(SSR_0 - \sum SSR_i)/(k(m-1))}{\sum SSR_i/(n-mk)} \tag{1}$$

Formulas

$$\hat{\beta}_1 = \bar{Y} - \hat{\beta}_2\bar{X}; \hat{\beta}_2 = \frac{\sum(X_i - \bar{X})(Y_i - \beta Y)}{\sum(X_i - \bar{X})^2}$$
$$\hat{\beta}_2 = \frac{\sum X_i Y_i}{\sum X_i^2} \text{ - no intercept}$$
$$SST = \sum(Y_i - \bar{Y})^2, \text{ sum of squares total}$$
$$SSE = \sum(\hat{Y}_i - \bar{Y})^2, \text{ sum of squares explained}$$
$$SSR = \sum(\hat{Y}_i - Y_i)^2, \text{ sum of squares}$$
$$R^2 = \frac{SSE}{SST} = \frac{SST - SSR}{SST}; r_{Y,\hat{Y}} = \sqrt{R^2} \text{ - correlation}$$
$$R_{adj}^2 = 1 - \frac{SSR/(n-k)}{SST/(n-1)}$$
$$p = F(Z);\; Z = \beta_1 + \beta_2 X_3 + \cdots + \beta_k X_k$$
$$F(Z) = \frac{1}{1+e^{-Z}}; f(Z) = \frac{dp}{dZ} = \frac{1}{\sqrt{2\pi}}e^{-0.5Z^2}$$
$$Logit = \prod Pr(Y = Y_j | X_j, \beta) \text{ - max over } \beta$$
$$P(Y=1) = F(Y); P(Y=0) = 1 - F(Y) \tag{2}$$

White test heteroscedascticty test

$$H_0 : \text{homoscedastic} \tag{3}$$

- Regress against variables, their squares, and their cross-products, no dup
- Calculate nR^2 using R^2 from this regression. It is $\sim \chi_{n-1}$, where n - number of regressors, including constant.

Breusch-Pagan heteroscedascticty test

Needs less variables.

$$H_0 : \text{homoscedastic} \tag{4}$$

- Regress against variables
- Calculate nR^2 using R^2 from this regression. It is $\sim \chi_{n-1}$, where n - number of regressors, including constant.

Goldfeld-Quandt heteroscedascticty test

Split into three ranges: n_1 with smallest X, n_2 with highest X, and the rest.

$$F_{n_2-k,n_1-k} = \frac{RSS_2/(n_2-k)}{RSS_1/(n_1-k)} \tag{5}$$

Maximum likelihood estimation

$$f(Y_j) = \frac{1}{\sigma\sqrt{2\pi}}e^{-0.5\left(\frac{Y_i - \beta_1 - \beta_2 X_i}{\sigma}\right)}$$
$$\log L = log\left(\prod f(Y_i)\right) = n\log\left(\frac{1}{\sigma\sqrt{2\pi}}\right) - \frac{\sigma^{-2}}{2}Z \tag{6}$$
$$R_{\text{pseudo}}^2 = 1 - \frac{\log L}{\log L_0}; LR = 2(\log L - \log L_0)$$
where L_0 is the likelihood with only intercept

Gauss-Markov conditions

1. $\mathbb{E}(u_i) = 0$
2. $\sigma_{u_i}^2 = \sigma_{u_j}^2$
3. u_i and u_j are independent $\iff i \neq j$
4. u_i is independent of X

4th condition violation

- Reasons: measurement errors, endogenity
- Consequences: inconsistent OLS estimators, standard stats wrongly calculated, tests invalid
- Remedial measures: instrumental variables. Find another variable independently dsitributed with the disturbance term, but correlated with the endogenous explanatory variable
- Detection: Durbin-Wu-Hausman test (standard or Davidson-McKinnon version a a)

MLE tests

Test H_0 if all explanatory variables are 0.

$$LR \sim \chi_{k-1}^2 \tag{7}$$

where k - number of parameters estimated (and $k-1$ - number of explanatory variables). Also, F-test:

$$LR = 2(L_{ur} - L_r) \sim \chi_q^2 \tag{8}$$