Multicollinearity

- Effect of multicollinearity on testing:

$\downarrow t_j = \frac{b_j - \beta_j}{\sqrt{Var(b_j)}\uparrow}$ vs. t_c Slide 15

Slide 18:

- F-test based on an auxilliary regression:

$H_0: \beta_j = 0, \forall j = 2, \ldots, k$ (no multicollinearity)
$H_1: \beta_j \neq 0, \exists j = 2, \ldots, k$ (multicollinearity)

- Variance inflation factor:

$1 \leq VIF_j \leq \infty$

no multicollinearity

perfect multicollinearity

$R_j^2 = ?$

Rule of thumb: $VIF_j > 10$, serious multicollinearity issues that need to be addressed.

$R_j^2 > 0.9$, whereas ideally it should be 0.

- Tolerance:

$$Tol_j = \frac{1}{VIF_j} = 1 - R_j^2$$

Rule of thumb: $Tol_j < 0.1$, serious multi-collinearity issues that need to be addressed.

Example 1 (PDF, pp. 1,3-5)

b) Multicollinearity:

- R^2 versus $p(b_j)$. ✓
- Correlation coefficients:

$p(r_{exp,alco}) = 0.0011 < \alpha = 0.05$

$r^2_{exp,alco} = (-0.5510)^2 = 0.3036$

$r^2_{exp,alco} = 0.3036 < R^2 = 0.7381$ ✓

Klein's rule of thumb: $r^2_{x_j x_l} \geq R^2$

- F-tests based on auxilliary regressions:

M_{AE}: $\widehat{exp}_i = c_1 + c_2\, tobacco_i + c_3\, alco_i$

$R^2_{exp} = 0.3280$

$H_0: \gamma_j = 0, \forall j = 2,3$ no multicol.

$H_1: \gamma_j \neq 0, \exists j = 2,3$ multicol.

$$F = \frac{R^2_{exp} / (k-2)}{(1-R^2_{exp})/(n-k+1)} = \frac{0.3280/(4-2)}{(1-0.3280)/(32-4+1)} = \underline{7.078}$$

from M_B

$F_{c\,(m_1 = k-2 = 2,\ m_2 = n-k+1 = 29,\ \alpha = 0.05)} = 3.33$

$F > F_c$, we reject H_0 at $\alpha = 0.05$ and conclude from H_1.

Variable exp_i causes multicollinearity.

HW: Do the other two auxilliary regressions.

- Calculation of VIF_j and Tol_j:

$$VIF_j = \frac{1}{1-R_j^2}$$
$$VIF_{exp} = \frac{1}{1-R_{exp}^2} = \frac{1}{1-0.3280} = \underline{\underline{1.488}} < 10 \checkmark$$

$$Tol_j = 1-R_j^2 = \frac{1}{VIF_j}$$
$$Tol_{exp} = \frac{1}{1.488} = \underline{\underline{0.672}} > 0.1 \checkmark$$

Stata: commands ESTAT VIF and COLLIN.

Example 2 (PDF, pp. 5-8) in Stata only.