

I . Short Review

1. Regression(Slope)

- From regression line, we predict the mean of Y for certain X. Think about what is “b” .

2. Inference

- Test “b” : $t = \frac{b-\beta}{\sigma_b}$, $\sigma_b = \frac{\hat{\sigma}}{s_x\sqrt{N-1}}$, $\hat{\sigma}(\text{Root MSE}) = \sqrt{\frac{SS_e}{N-K}}$,
 $SS_e = \sum(Y - \hat{Y})^2 = \sum e^2$.

3. Goodness of fit

- $SS_t = SS_b + SS_w$, $SS_t (\sum(Y - \bar{Y})^2) = SS_r + SS_e$
- In multiple regression, $r^2 \rightarrow R^2 = \frac{SS_r}{SS_t} = \frac{SS_t - SS_e}{SS_t}$, $1 - r^2 = \frac{SS_e}{SS_t}$.

4. Joint Hypothesis Test

- $H_0 : \beta_1 = \beta_2 = \dots = \beta_k = 0$
- $F = \frac{MS_r}{MSE}$, $MS_r = \frac{SS_r}{DF_r}$, $DF_r : K-1$. $MSE = \frac{SS_e}{DF_e}$, $DF_e : N-K$.
- $F_{crit} : (DF_r, DF_e)$

5. STATA

- It will be covered in the session.
- Regression DV IV₁ IV₂.....IV_k
- Anova table
- Dummy variable, and Interaction variable.

II. Problem

1. Fill the blanks

reg lreal educ age

Source	SS	df	MS	Number of obs =	872
-----+-----				F(,) =	(?)
Model	155.069798	(?)	(?)	Prob > F	
Residual	671.287419	(?)	(?)	R-squared =	(?)
-----+-----				Adj R-squared =	0.1858
Total	(?)	871	.948745369	Root MSE =	.87891

lreal	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
-----+-----					
educ	.1378977	.0100969			.1180805 .1577148
age	-.0028203	.0019003			-.0065501 .0009095
_cons	8.3219	.1724614	48.25	0.000	7.983411 8.66039

2. Test the coefficients of IVs