1. Clarification About Homework 1

Question 5: We want “percentage variation explained by speed” and “intercept of the regression model” and “coefficient of speed” separately but not “percentage variation explained by speed” and “percentage variation explained by intercept” and “percentage variation explained by speed coefficient”

Question 10: In option A “ error term = 1- confidence level”

Question 11: Choose the closet answer; set Facebook = 45 as a constant when calculating

Question 20: See below

2. Diagnostic Plots

- Residuals vs. Fitted
  This plot shows if residuals have non-linear relationships. If you find equally spread residuals around a horizontal line without distinct patterns, that is a good indication you don’t have non-linear relationships or non-constant variances.
- Normal QQ plot
  This plot shows whether the residuals are normally distributed
  https://stats.stackexchange.com/questions/101274/how-to-interpret-a-qq-plot

3. Log Transformation and Interpretation

In a regression model $y \sim x$:

If we say, x increases by 1 unit, we simply mean $x \rightarrow x + 1$

If we say, x increases by 1%, we simply mean $x \to (1 + 1\%)x$ or $logx \to logx + 0.01$; The former is an approximation while the latter is the accurate formula in this course

The same terminology applies to y

Notice: $x \to (1 + 1\%)x \Leftrightarrow log\ x \to log(1 + 1\%)x = log\ x + log1.01 \approx log\ x\ +\ 0.01$

In homework 1 from Question 17 to Question 20, please use $\boldsymbol{logx \to logx + 0.01}$ to solve the problems

Linear-Linear Model:

$y = b_0 + b_1 x$ and $x \to x + 1$

$$y(x + 1) - y(x) = \big(b_0 + b_1(x + 1)\big) - (b_0 + b_1 x) = b_1$$

As x increases by 1 unit, y increases by b_1 units

Linear-Log Model:

$y = b_0 + b_1 log\ x$ and $x \to (1 + 1\%)x$ or $log\ x \to log\ x\ + 0.01$

$$\begin{aligned} y(1.01x) - y(x) &= \big(b_0 + b_1 log(1.01x)\big) - (b_0 + b_1 log\ x) \\ &= b_1 log(1.01) \approx 0.01 b_1 \end{aligned}$$

$$\begin{aligned} &y(log\ x + 0.01) - y(log\ x) \\ &\qquad = \big(b_0 + b_1(log\ x + 0.01)\big) - (b_0 + b_1 log\ x) = 0.01 b_1 \end{aligned}$$

As x increases by 1%, y increases by $0.01b_1$ units

Log-Linear Model:

$log\ y = b_0 + b_1 x \rightarrow y = e^{b_0 + b_1 x}$ and $x \rightarrow x + 1$

$$\frac{y(x+1)}{y(x)} = \frac{e^{b_0 + b_1(x+1)}}{e^{b_0 + b_1 x}} = e^{b_1} \Leftrightarrow y \rightarrow (1 + (e^{b_1} - 1))y$$

As x increases by 1 unit, y increases by $100(e^{b_1} - 1)\%$

Log-Log Model:

$log\ y = b_0 + b_1 log\ x \rightarrow y = e^{b_0 + b_1 log\ x} = e^{b_0} x^{b_1}$ and $x \rightarrow (1 + 1\%)x$ or $log\ x \ \rightarrow log\ x + 0.01$

$$\frac{y(log\ x + 0.01)}{y(log\ x)} = \frac{e^{b_0 + b_1(log\ x + 0.01)}}{e^{b_0 + b_1 log\ x}} = e^{0.01 b_1}$$

$$\Leftrightarrow y \rightarrow \left(1 + (e^{0.01 b_1} - 1)\right) y$$

As x increases by 1%, y increases by $100(e^{0.01 b_1} - 1)\%$