

Classwork 13B: Ramsey RESET Test as a Function in R

In this classwork, you'll conduct the Ramsey RESET test for functional misspecification for our `students` dataset and a model that's linear in variables. Then you'll write a function `ramsey` which will be able to take *any* model and dataset and run the Ramsey RESET test on it.

Run this code to get started:

```
library(tidyverse)
library(gapminder)
```

```
students <- read_csv("https://raw.githubusercontent.com/cobriant/students_dataset/main/students.csv")
```

1. Take `students` and estimate the model $final_grade = \beta_0 + \beta_1 sex + \beta_3 failures + \beta_4 romantic + \beta_6 absences + u$. Interpret the estimates for the coefficients.
2. Using `ggplot` and intuition, which explanatory variables might have a nonlinear relationship with `final_grade`? That is, do you think it makes sense to include any squared terms or interactions in the model above?
3. In your own words, explain how the Ramsey RESET test works for functional misspecification.
4. Take `students` and conduct the first step in the Ramsey RESET test: take the linear model from question 1 and add a column `yhat` to `students` that is the fitted values from that regression.

```
# students %>%
#   mutate(yhat = __)
```

5. Take your answer to question 4 and use it to estimate the new model: $final_grade = \beta_0 + \beta_1 sex + \beta_2 failures + \beta_3 romantic + \beta_4 absences + \beta_5 yhat^2 + u$. Conduct a hypothesis test on β_5 by looking at the p-value from `broom::tidy()`. What are the results from the Ramsey test?

```
# students %>%
#   mutate(yhat = __) %>%
#   lm(__ + I(yhat^2), data = .) %>%
#   broom::tidy() %>%
#   slice_tail(n = 1) %>%
#   select(p.value) %>%
#   mutate(nonlinear_detected = __ < .05)
```

6. Use your code from question 5 to write a function to conduct the Ramsey test on any model and dataset.

```
# ramsey <- function(dataset, linear_model) {  
#   __ %>%  
#   mutate(yhat = __) %>%  
#   lm(paste0(__, " + I(yhat^2)"), data = .) %>%  
#   broom::tidy() %>%  
#   slice_tail(n = 1) %>%  
#   select(p.value) %>%  
#   mutate(nonlinear_detected = __ < .05)  
# }
```

Make sure you can call your function like this:

```
# students %>%  
#   ramsey("final_grade ~ sex + failures + romantic + absences")
```

And also like this:

```
# gapminder %>%  
#   ramsey("lifeExp ~ gdpPerCap + year")
```