Solution to: Some more work with Linear Splines

Method 1:

Let

X1 = age - 65

X2 = 0 if age < 65
age - 65 if age ≥ 65

X3 = 0 if age < 80
Age - 80 if age ≥ 80

Our model is: $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon$

Lets re-write our model for the three age ranges:

Age < 65

$$\begin{aligned} Y &= \beta_0 + \beta_1(age - 65) + \varepsilon \\ &= \beta_0 - 65\beta_1 + \beta_1 age + \varepsilon \\ &= \beta_0^{\ 1} + \beta_1 age + \varepsilon \end{aligned}$$

65 ≤ Age < 80

$$\begin{aligned} Y &= \beta_0 + (\beta_1 + \beta_2)(age - 65) + \varepsilon \\ &= \beta_0 - 65\beta_1 - 65\beta_2 + (\beta_1 + \beta_2)age + \varepsilon \\ &= \beta_0^{\ 2} + (\beta_1 + \beta_2)age + \varepsilon \end{aligned}$$

Age ≥ 80

$$\begin{aligned} Y &= \beta_0 + (\beta_1 + \beta_2)(age - 65) + \beta_3(age - 80) + \varepsilon \\ &= \beta_0 - 65\beta_1 - 65\beta_2 - 80\beta_3 + (\beta_1 + \beta_2 + \beta_3)age + \varepsilon \\ &= \beta_0^{\ 3} + (\beta_1 + \beta_2 + \beta_3)age + \varepsilon \end{aligned}$$

Now we can interpret our regression coefficients:

$\beta_1 =$ Expected change in total medical expenditures per year for persons under 65 years of age.

$\beta_1 + \beta_2 =$ Expected change in total medical expenditures per year for persons aged 65 to 80.

$\beta_1 + \beta_2 + \beta_3 =$ Expected change in total medical expenditures per year for persons over 80.

$\beta_0^{\ 1} =$ Y-intercept for the linear function for persons under 65 years of age.

$\beta_0^{\ 2} =$ Y-intercept for the linear function for persons aged 65 to 80.

$\beta_0^{\ 3} =$ Y-intercept for the linear function for persons over 80.

Question 1: Is the rate at which total medical expenditures increase with age different for persons less than 65 years of age compared to persons who are 65 to 80 years of age?

Statistical Answer: Test $H_0 : \beta_2 = 0$

Question 2: Is the rate at which total medical expenditures change with age different for persons 65 to 80 years of age compared to persons over 80?

Statistical Answer: Test $H_0 : \beta_3 = 0$

Method 2:

Let

X1 = 0 if age > 65
age – 40 if age ≤ 65

X2 = 0 if age < 65 or age ≥ 80
age - 65 if 65 ≤ age < 80

X3 = 0 if age < 80
Age - 80 if age ≥ 80

Our model is: $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon$

Lets re-write our model for the three age ranges:

Age < 65
$$\begin{aligned} Y &= \beta_0 + \beta_1(age - 40) + \varepsilon \\ &= \beta_0 - 40\beta_1 + \beta_1 age + \varepsilon \\ &= \beta_0^{1} + \beta_1 age + \varepsilon \end{aligned}$$

65 ≤ Age < 80
$$\begin{aligned} Y &= \beta_0 + \beta_2(age - 65) + \varepsilon \\ &= \beta_0 - 65\beta_2 + \beta_2 age + \varepsilon \\ &= \beta_0^{2} + \beta_2 age + \varepsilon \end{aligned}$$

Age ≥ 80
$$\begin{aligned} Y &= \beta_0 + \beta_3(age - 80) + \varepsilon \\ &= \beta_0 - 80\beta_3 + \beta_3 age + \varepsilon \\ &= \beta_0^{3} + \beta_3 age + \varepsilon \end{aligned}$$

Now we can interpret our regression coefficients:

$\beta_1 =$ Expected change in total medical expenditures per year for persons under 65 years of age.

$\beta_2 =$ Expected change in total medical expenditures per year for persons aged 65 to 80.

$\beta_3 =$ Expected change in total medical expenditures per year for persons over 80.

$\beta_0^1 =$ Y-intercept for the linear function for persons under 65 years of age.

$\beta_0^2 =$ Y-intercept for the linear function for persons aged 65 to 80.

$\beta_0^3 =$ Y-intercept for the linear function for persons over 80.

Question 1: Is the rate at which total medical expenditures increase with age different for persons less than 65 years of age compared to persons who are 65 to 80 years of age?

Statistical Answer: Test $H_0 : \beta_2 = \beta_1$

Question 2: Is the rate at which total medical expenditures change with age different for persons 65 to 80 years of age compared to persons over 80?

Statistical Answer: Test $H_0 : \beta_3 = \beta_2$