Problem 1: Noise pollution due to cars

Question a)

First check the assumptions of the ANOVA model:

- Normality of the groups:

```
> Ps <- c(
    shapiro.test(noise[fuel == "diesel" & category ==
"commercial"])$p.value,
    shapiro.test(noise[fuel == "diesel" & category ==
"passenger"])$p.value,
    shapiro.test(noise[fuel == "ethanol" & category ==
"commercial"])$p.value,
    shapiro.test(noise[fuel == "ethanol" & category ==
"passenger"])$p.value,
    shapiro.test(noise[fuel == "gasoline" & category ==
"commercial"])$p.value,
    shapiro.test(noise[fuel == "gasoline" & category ==
"passenger"])$p.value
)
Ps
> Ps
0.3574334 0.2930456 0.5325407 0.7372111 0.1878295 0.7897125
```

All p-values are greater than 0.1, so we can't reject at any significant level the normality of the groups.

- Homoscedasticity:

```
> bartlett.test(noise ~ fuel_category)

        Bartlett test of homogeneity of variances

data:  noise by fuel_category
Bartlett's K-squared = 7.968, df = 5, p-value = 0.158
```

The p-value is greater than 0.1, so we can't reject at any significant level the homoscedasticity of the groups.

Now that I have checked the assumptions, I can perform the ANOVA test:

Complete ANOVA model with interactions:

$$ X_{ijk} = mu + tau_i + beta_j + gamma_{ij} + eps_{ijk} $$ with:

- mu: overall mean

- tau_i: effect of fuel i
- $beta_j$: effect of category j
- $gamma_{ij}$: interaction effect of fuel i and category j
- $eps_{ijk} \sim \mathcal{N}(0, \sigma^2)$: error term

```
> summary(aov(noise ~ fuel + category + fuel:category))
               Df Sum Sq Mean Sq F value   Pr(>F)
fuel            2  375.8  187.91  16.040 7.28e-07 ***
category        1   21.8   21.84   1.864   0.1749
fuel:category   2   63.8   31.91   2.723   0.0699 .
Residuals     114 1335.5   11.72
```

From the summary I can see that:

- Test 1:

  $ H0: \gamma_i = 0, i = 1,...,6 $ vs $H1: (H0)^c $

  that is:

  - H0: the effect of the fuel doesn't significantly affect the noise pollution
  - H1: the effect of the fuel significantly affects the noise pollution

  -> the p-value for this test is 0.0699: I reject at 10% the null hypothesis but not at 5% and 1%.

  -> we don't have strong evidence that the interaction term has effect

- Test 2: $H0: tau_i = 0, i = 1,2,3$ vs $H1: (H0)^c$

  that is

  - H0: the effect of the fuel doesn't significantly affect the noise pollution
  - H1: the effect of the fuel significantly affects the noise pollution

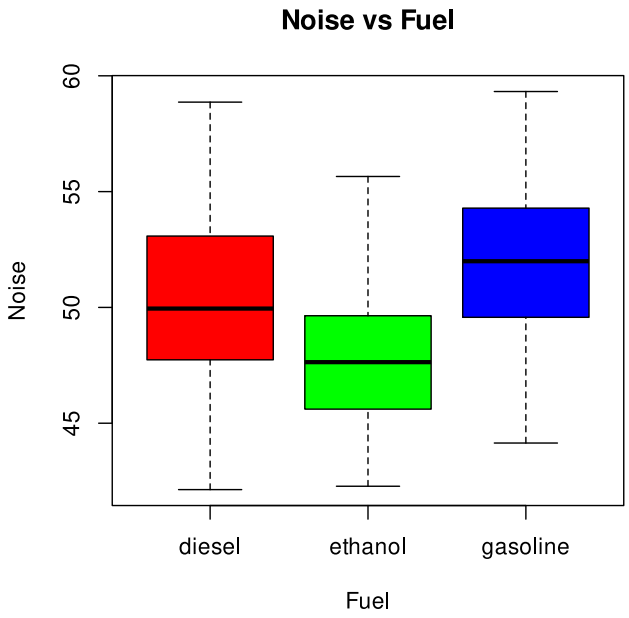
  -> the p-value for this test is ~ 7e-07: I reject at 1% the null hypothesis so I can conclude that the effect of the fuel significantly affects the noise pollution

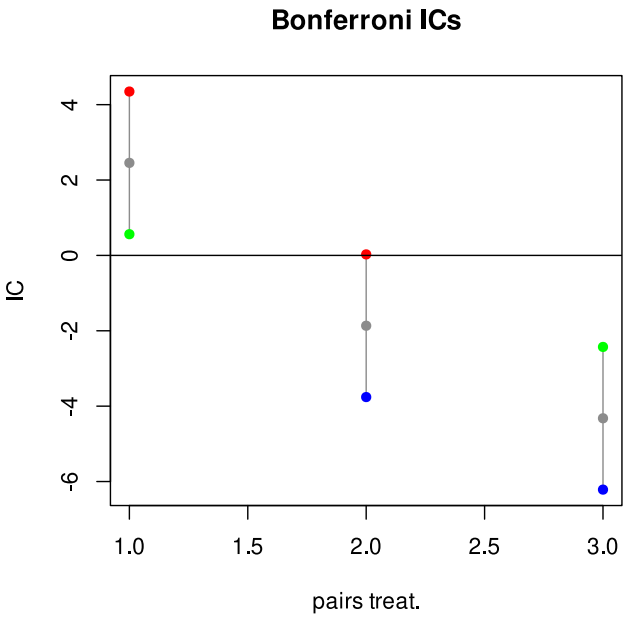
- Test 3: $ H0: beta_i = 0, i = 1,2$ vs $H1: (H0)^c$

  that is

  - H0: the effect of the category doesn't significantly affect the noise pollution
  - H1: the effect of the category significantly affects the noise pollution

  -> the p-value for this test is ~ 0.1749: I reject at any significant level the null hypothesis so I can conclude that the effect of the category doesn't significantly affect the noise pollution


Noise vs Fuel
Noise
60
55
50
45
diesel
ethanol
gasoline
Fuel


Bonferroni ICs
IC
4
2
0
-2
-4
-6
1.0
1.5
2.0
2.5
3.0
pairs treat.