In class activity 05

2024-09-17

1. Using data from the blood pressure example run two models using the `lm()` function, one that includes age and sex as covariates and one that includes age, sex, and the interaction.

```
lm_1 <- lm(formula = blood_pressure ~ age + sex, data = blood)
lm_2 <- lm(formula = blood_pressure ~ age + sex + age * sex, data = blood)
```

1. Calculate the BIC for both models and compare them.

```
bic_age_sex <- dim(blood)[1] * log(x = mean(lm_1$residuals ^ 2)) +
  4 * log(x = dim(blood)[1])

bic_int <- dim(blood)[1] * log(x = mean(lm_2$residuals ^ 2)) +
  5 * log(x = dim(blood)[1])

print(c(bic_age_sex,bic_int))
```

```
## [1] 481.7062 470.5836
```

1. Write a conclusion stating which model you would choose from the two and interpret the values of the parameters.
2. Fit a model that adds height as a predictor (with age, sex, and the interactions). Calculate the BIC, is this new model better than the one that doesn't include height?

```
lm_3 <- lm(formula = blood_pressure ~ age + sex + height + age * sex,
           data = blood)

bic_int_height <- dim(blood)[1] * log(x = mean(lm_3$residuals ^ 2)) +
  6 * log(x = dim(blood)[1])

print(c(bic_age_sex,bic_int, bic_int_height))
```

```
## [1] 481.7062 470.5836 475.3681
```