Chapter 4 Homework

Kaylynn Hiller

2026-02-02

```
-- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
v dplyr     1.1.4     v readr     2.1.5
v forcats   1.0.1     v stringr   1.6.0
v ggplot2   4.0.0     v tibble    3.3.1
v lubridate 1.9.4     v tidyr     1.3.2
v purrr     1.2.1
-- Conflicts ------------------------------------------ tidyverse_conflicts() --
x dplyr::filter() masks stats::filter()
x dplyr::lag()    masks stats::lag()
i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to becom
here() starts at /Users/kaylynnhiller/Desktop/Stats1

Rows: 31 Columns: 9
-- Column specification --------------------------------------------------------
Delimiter: ","
chr (1): _OBSTAT_
dbl (8): AGE, WEIGHT, RUNTIME, RSTPULSE, RUNPULSE, MAXPULSE, OXYGEN, GROUP

i Use `spec()` to retrieve the full column specification for this data.
i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

Homework questions

1. Sampling distributions and statistical surprise

Write brief answers.

- What is a sampling distribution of a statistic (like a mean), and what is it useful for?

- The sampling distribution is a theoretical distribution of taking repeated samples of a certain size from a population and calculating a statistic from those samples such as the mean. These are useful because we can compare a sample we have to the sampling distribution and make conclusions. We can use a sampling distribution to estimate whether a null hypothesis is likely true.

- How can a sampling distribution be used to evaluate whether an observed value is "surprising" under a null hypothesis?
  - A sampling distribution can be used to evaluate an observed value by comparing the observed value to the sampling distribution. If the observed value is higher or lower than the majority of the sampling distribution (such as 95% of the distribution) we can reject the null hypothesis. If the observed value is similar to the majority of the values in the sampling distribution we would fail to reject the null hypothesis.

2. Model comparison test: is the mean resting pulse 72?

Using the Fitness data, evaluate whether the sample of participants is consistent with a population mean of `RSTPULSE = 72` using a **model comparison**.

2a. Fit Model C and Model A

Use these models:

- **Model C (null):** the outcome is fixed at 72 with no free parameters

$$Y_i = 72 + \varepsilon_i.$$

- **Model A (mean model):** estimate the mean from the data

$$Y_i = b_0 + \varepsilon_i.$$

```
Call:
lm(formula = rstpulsedev ~ 0, data = fitness)

Residuals:
   Min     1Q Median     3Q    Max 
 -32.0  -24.0  -20.0  -13.5    4.0 

No Coefficients
```

```
Residual standard error: 20 on 31 degrees of freedom
```

```
Call:
lm(formula = rstpulsedev ~ 1, data = fitness)

Residuals:
    Min      1Q  Median      3Q     Max 
-13.742  -5.742  -1.742   4.758  22.258 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)   -18.26       1.49  -12.26 3.29e-13 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 8.294 on 30 degrees of freedom
```

2b. Compute PRE and F

Compute:

$$\text{SSE}(\cdot) = \sum_i e_i^2$$

(use `deviance()` for an `lm` object)

$$\text{PRE} = \frac{\text{SSE}(C) - \text{SSE}(A)}{\text{SSE}(C)}$$

$$F^* = \frac{\text{PRE}/(P_A - P_C)}{(1 - \text{PRE})/(n - P_A)}$$

```
[1] 150.2091
```

2c. Identify "surprising" values from the book tables and/or by using `qf()`.

Fill in the quantities the tables depend on:

- number of participants: n = 31
- number of parameters in Model C: P_C = 0
- number of parameters in Model A: P_A = 1
- new parameters: $P_A - P_C$ = 1
- unused parameters (residual df for Model A): $n - P_A$ = 30

Then, using the book tables and/or `qf()`:

- critical PRE at $\alpha = .05$: PREcrit = .122
- critical F at $\alpha = .05$: Fcrit = 4.17

2d. Conclusions

- Statistical conclusion: Do you reject or retain the null hypothesis at $\alpha = .05$? Explain using the logic of "surprising under the null."
  - We reject the null hypothesis. The f stat we calculated for this test was 150.2. This is larger than the critical value of 4.17. An f stat over a value of 4.17 would be surprising under the null since this means the observed sample mean is far from the null-hypothesized mean of 72. Since our fstat is high, our sample mean is very different from the estimated mean and the estimated difference is larger than a difference that would occur due to random error.
- Substantive conclusion: Give a one-sentence interpretation in plain language.
  - This sample of men has an average resting heart rate statistically different from the population mean of 72 beats per minute.

3. Alternative F formula using SSR and SSE(A)

Let

$$\text{SSR} = \text{SSE}(C) - \text{SSE}(A)$$

$$F^* = \frac{\text{SSR}/(P_A - P_C)}{\text{SSE}(A)/(n - P_A)}$$

3a. Compute F using SSR and SSE(A)

3b–3e. Interpretation

Write brief answers.

- What does the numerator SSR$/(P_A - P_C)$ represent?
  - This represents the mean squares reduced. This is the average proportional reduction in error per parameter added.
- What does the denominator SSE$(A)/(n - P_A)$ represent?
  - This represents the mean square error. This is the average proportional reduction in error that could be obtained by adding all possible remaining parameters.
- What does F^* represent conceptually?
  - F represents the ratio of explained variance per added parameter to unexplained variance per unused parameter
- Compute t for this F^* (use the relationship between t and F when the numerator df is 1).
  - -12.25598

```
[1] 12.25598
```

4. Does running increase pulse rate?

Using the Fitness dataset, compare `RSTPULSE` (resting pulse) to `RUNPULSE` (post-run pulse).

4a. Set up Model C and Model A

Use a model comparison where the null corresponds to **no mean change** from resting to post-run.

One convenient approach:

1. Create a difference score $D_i = \text{RUNPULSE}_i - \text{RSTPULSE}_i$.
2. Compare:
   - **Model C:** $D_i = 0 + \varepsilon_i$ (no free parameters)

$$D_i = 0 + \varepsilon_i.$$

- **Model A:** $D_i = b_0 + \varepsilon_i$ (estimate the mean difference)

$$D_i = b_0 + \varepsilon_i.$$

```
Call:
lm(formula = d ~ 0, data = fitness)

Residuals:
   Min     1Q Median     3Q    Max 
  92.0  111.0  116.0  123.5  136.0 

No Coefficients

Residual standard error: 116.4 on 31 degrees of freedom



Call:
lm(formula = d ~ 1, data = fitness)

Residuals:
     Min       1Q   Median       3Q      Max 
-23.9032  -4.9032   0.0968   7.5968  20.0968 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  115.903      1.966   58.95   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 10.95 on 30 degrees of freedom
```

4b. Compute PRE and F, then compare to critical values

```
[1] 3475.443
```

From the book tables at $\alpha = .05$:

- PREcrit = .122
- Fcrit = 4.17

4c. Conclusions

- Statistical conclusion (reject/retain the null).
  - We reject the null hypothesis.
- Substantive conclusion (plain-language interpretation).
  - Our sample mean difference between resting and running heart rate is statistically much higher than the estimated difference of 0.

5. With your own data

Return to your own dataset and the model comparison you set up previously. In this section, please use

5a. Fit your models

```
Call:
lm(formula = pill_dev ~ 0, data = mydata)

Residuals:
   Min     1Q Median     3Q    Max 
    -1     -1      0      1      2 

No Coefficients

Residual standard error: 1.061 on 2123 degrees of freedom
```

```
Call:
lm(formula = pill_dev ~ 1, data = mydata)

Residuals:
    Min      1Q  Median      3Q     Max 
-1.1997 -1.1997 -0.1997  0.8003  1.8003 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  0.19972    0.02263   8.825   <2e-16 ***
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.043 on 2122 degrees of freedom
```

5b. Report PRE, F*, and t*

```
[1] 77.87881
```

```
[1] 8.824897
```

5c. Use tables or functions to identify critical values and write conclusions

```
[1] 3.845843
```

```
[1] 2.239893e-18
```

Our f stat (77.9) is much higher than the critical value of 3.85. Our p-value is very small and essentially 0. This means that our sample is statically much different from our estimated value of 2. We reject the null hypothesis that the Americans on average agree (2) that teenagers should have access to birth control.