Chapter 5 Homework

Kaylynn Hiller

2026-02-09

Setup

```
-- 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 becor
here() starts at /Users/kaylynnhiller/Desktop/Stats1
```

```
Attaching package: 'psych'
```

```
The following objects are masked from 'package:ggplot2':

    %+%, alpha
```

General notes on the model-comparison workflow

For most questions below, you will compare:

- **Model C (constrained / null model)**: a model where the intercept is fixed (often to 0 or to a known/reference value).
- **Model A (augmented model)**: a model where the intercept is estimated from the data.

For a one-parameter intercept-only Model A,

$$\text{Model A: } Y_i = b_0 + \varepsilon_i$$

and for a fixed-intercept Model C (often b_0 fixed to some value c),

$$\text{Model C: } Y_i = c + \varepsilon_i.$$

The key quantities:

- SSE is sum of squared errors.
- $\text{PRE} = \dfrac{\text{SSE}_C - \text{SSE}_A}{\text{SSE}_C}$.
- For a one-parameter comparison (typically $P_A - P_C = 1$), the usual model-comparison F statistic is:

$$F = \frac{(\text{SSE}_C - \text{SSE}_A)/(P_A - P_C)}{\text{SSE}_A/(n - P_A)}.$$

Part A: Hand-calculation practice (small, self-contained datasets)

A1. Schizophrenia age of onset

Assume the mean age of onset of schizophrenia in a sample of $n = 50$ persons was $\bar{y} = 21.5$ years and the variance was $s^2 = 2.5$ (years^2).

1. Specify the model that uses the mean to describe the data (Model A), in both words and in equation form.

   -
   $$\text{Model A: } Y_i = b_0 + \varepsilon_i$$

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

2. What is the SSE for this model?

   - $SSE = s^2 * (n - 1)$
   - 122.5

(Use the definitional relationships between variance, sums of squares, and SSE for an intercept-only model.)

```
2.5*(50-1)
```

```
[1] 122.5
```

A2. Multiculturalism vs. colorblindness (paired data)

Ryan, Hunt, Weible, Peterson, and Casas (2007) investigated Americans' interethnic ideologies. They asked college students to indicate the extent to which they believed that multiculturalism and colorblind ideologies would improve interethnic relations in the U.S.

Assume you have data for 10 Black participants:

Subjid	MC	CB	Diff (MC – CB)
1	5.0	4.0	1.0
2	4.0	3.5	0.5
3	6.0	4.0	2.0
4	6.0	6.0	0.0
5	5.0	4.5	0.5
6	4.5	5.0	-0.5
7	5.0	3.0	2.0
8	6.5	4.5	2.0
9	7.0	5.0	2.0
10	6.5	3.0	3.5

A2a. By hand

1. Compute **by hand** the mean, variance, and standard deviation of the difference variable. Use definitional formulas.

   - mean = 1.3
   - variance = 1.46
   - standard deviation = 1.21

2. Compute SSE_C for the simple model that predicts no difference (i.e., that the difference is 0). This is Model C.

- 30

3. Compute SSE_A for the simple model that best describes the differences based on the data (Model A).

   - $SSE = s^2 * (n - 1)$ = 1.46*(10-1) = 13.1

4. Compute PRE and F.

   - PRE = 0.563
   - F = 11.61

5. Specify the null hypothesis tested by comparing these two models.

   - the null hypothesis is that there is no difference in the two ideologies
   - $H_0 : b_0 = 0$

6. Indicate whether you would reject or fail to reject Model C / the null hypothesis.

   - We would reject the null hypothesis since our values of PRE and F are greater than the critical values of PRE and F at 95% significance and a 99% significance.

7. Write up your findings in journal style.

   - Our findings indicate that there is likely a difference in how these participants viewed how multiculturalism and colorblind ideologies would improve interethnic relations in the U.S.

(You can optionally use R at the end to check arithmetic, but do the main work by hand.)

```
diff_tibble <- tibble(
  subjid = c(1:10),
  diff = c(1,.5,2,0,.5,-.5,2,2,2,3.5)
)

summary_stats <- diff_tibble |>
  summarize(mean_diff = mean(diff),
            var_diff  = var(diff),
            sd_diff   = sd(diff))
summary_stats
```

```
# A tibble: 1 x 3
  mean_diff var_diff sd_diff
      <dbl>    <dbl>   <dbl>
1       1.3     1.46    1.21
```

```r
diff_tibble <- diff_tibble |>
  mutate(diff_2 = (diff)^2)

SSE_C <- sum(diff_tibble$diff_2)

SSE_A <- summary_stats$var_diff * 9

PRE <- (SSE_C - SSE_A)/SSE_C

PC <- 0
PA <- 1

n <- nrow(diff_tibble)

FSTAT <- (PRE/(PA - PC)) / ((1 - PRE)/(n - PA))
FSTAT
```

```
[1] 11.61069
```

A2b. In R (reproducing your values)

Recreate the analysis in R and verify the quantities above (mean, variance/SD, SSE_C, SSE_A, PRE, F, and p).

```r
model_c <- lm(diff ~ 0, data = diff_tibble)

model_a <- lm(diff ~ 1, data = diff_tibble)

sse_c <- deviance(model_c)
sse_a <- deviance(model_a)

pre <- (sse_c - sse_a)/sse_c

n <- nrow(diff_tibble)

fstat <- (pre/(PA - PC)) / ((1 - pre)/(n - PA))
fstat
```

```
[1] 11.61069
```

```
p <- pf(fstat, 1, 9, lower.tail = FALSE)
```

A2c. Confidence interval

Construct a 95% confidence interval for the multiculturalism – colorblindness difference score in this sample.

- 0.55 - 2.05

```
lower_bound <- summary_stats$mean_diff - (1.96*(summary_stats$sd_diff/sqrt(n)))
upper_bound <- summary_stats$mean_diff + (1.96*(summary_stats$sd_diff/sqrt(n)))
```

A2d. Power planning with `pwr`

A researcher wants to replicate the finding with a new sample of 10 participants.

1. Use results from this sample and the `pwr` package to estimate the number of participants needed to replicate your observed effect (at $\alpha = .05$).

    - 10 participants (5 in each group)

2. Use the same package to estimate the number of participants needed to detect a **medium** effect size (according to Cohen's very arbitrary definition…).

    - 128 participants (64 in each group)

```
f <- sqrt(PRE / (1 - PRE))
f
```

```
[1] 1.135815
```

```
pwr.anova.test(
  k = 2,
  f = f,
  sig.level = 0.05,
  power = 0.80
)
```

6

```
    Balanced one-way analysis of variance power calculation

            k = 2
            n = 4.254046
            f = 1.135815
     sig.level = 0.05
        power = 0.8

NOTE: n is number in each group
```

```r
pwr.anova.test(
  k = 2,
  f = 0.25,
  sig.level = 0.05,
  power = 0.80
)
```

```
    Balanced one-way analysis of variance power calculation

            k = 2
            n = 63.76561
            f = 0.25
     sig.level = 0.05
        power = 0.8

NOTE: n is number in each group
```

A3. Gym BMI (one-sample mean test)

A researcher wonders whether people who go to the gym in January have above-normal BMIs, that is, above 24.9 (upper level of "normal," according to the CDC). She obtains the BMI scores below for a randomly selected group of 10 gym users:

Participant	BMI
1	24.0
2	27.0

Participant	BMI
3	25.0
4	27.0
5	27.0
6	23.0
7	22.0
8	30.0
9	26.0
10	29.0

1. Specify Models C and A and the null hypothesis for this statistical test. What is this test commonly called?

   - Model C:
   $$\text{BMI}_i = 24.9 + \varepsilon_i$$
   - Model A:
   $$\text{BMI}_i = b_0 + \varepsilon_i$$
   - **Null hypothesis:** $H_0 : b_0 = 24.9$
   - One sample t-test

2. Conduct the analysis **by hand**.

```
bmi_tibble <- tibble(
  id = c(1:10),
  bmi = c(24,27,25,27,27,23,22,30,26,29)
)

summary_stats2 <- bmi_tibble |>
  summarize(mean_bmi = mean(bmi),
            var_bmi  = var(bmi),
            sd_bmi   = sd(bmi))
summary_stats2
```

```
# A tibble: 1 x 3
  mean_bmi var_bmi sd_bmi
     <dbl>   <dbl>  <dbl>
1       26    6.44   2.54
```

```
bmi_tibble <- bmi_tibble |>
  mutate(bmi_2 = (bmi - 24.9)^2)

SSE_C2 <- sum(bmi_tibble$bmi_2)
```

```
SSE_A2 <- summary_stats2$var_bmi * 9

PRE2 <- (SSE_C2 - SSE_A2)/SSE_C2

n2 <- nrow(bmi_tibble)

FSTAT2 <- (PRE2/(PA - PC)) / ((1 - PRE2)/(n2 - PA))
FSTAT2
```

[1] 1.877586

3. Conduct the analysis in R.

4. Provide a statistical conclusion.

 - We fail to reject the null hypothesis.

5. Write up your findings in journal style.

 - Our findings indicate that there is likely a no difference in the bmi of this sample compared to the upper end of the bmi normal score of 24.9.

```
bmi_tibble <- bmi_tibble |>
  mutate(bmi_dev = bmi - 24.9)
model_c2 <- lm(bmi_dev ~ 0, data = bmi_tibble)

model_a2 <- lm(bmi_dev ~ 1, data = bmi_tibble)

sse_c2 <- deviance(model_c2)
sse_a2 <- deviance(model_a2)

pre2 <- (sse_c2 - sse_a2)/sse_c2

n2 <- nrow(bmi_tibble)

fstat2 <- (pre2/(PA - PC)) / ((1 - pre2)/(n2 - PA))
fstat2
```

[1] 1.877586

```
p2 <- pf(fstat2, 1, 9, lower.tail = FALSE)
```

Part B: Fitness dataset (model comparison, confidence intervals, power)

B0. Read the data

Read in the `fitness.csv` dataset from the course project `data/` folder.

```
fitness <- read_csv(here("data", "fitness.csv"))
```

B1. Resting pulse vs. "normal" (72 bpm)

The variable `RSTPULSE` contains resting pulse rate for each participant in the fitness study. The normal pulse rate for men is 72.

B1a. 95% CI for b_0 in Model A

Construct a 95% confidence interval for b_0 in Model A.

- Do this **by hand** using a critical value (e.g., $t_{\alpha/2,\,df}$), and then confirm with R.

  - 50.8 - 56.7
  - 50.69951 - 56.78436 (difference due to rounding)

```
summary_stats3 <- fitness |>
  summarize(mean_pls = mean(RSTPULSE),
            var_pls  = var(RSTPULSE),
            sd_pls   = sd(RSTPULSE))
summary_stats3
```

```
# A tibble: 1 x 3
  mean_pls var_pls sd_pls
     <dbl>   <dbl>  <dbl>
1     53.7    68.8   8.29
```

```
n3 <- nrow(fitness)

lower_bound3 <- summary_stats3$mean_pls - (1.96*(summary_stats3$sd_pls/sqrt(n3)))
upper_bound3 <- summary_stats3$mean_pls + (1.96*(summary_stats3$sd_pls/sqrt(n3)))
```

- Make note of what in R output corresponds to your hand calculation.

For reference, the by-hand CI for an intercept estimate is:

$$\hat{b}_0 \pm t_{\alpha/2,\,df} \cdot SE(\hat{b}_0).$$

```
t.test(fitness$RSTPULSE)$conf.int
```

```
[1] 50.69951 56.78436
attr(,"conf.level")
[1] 0.95
```

B1b. Journal-style summary

Write a complete summary:

1. briefly explain the data;

   - The dataset `fitness.csv` contains resting pulse rate (`RSTPULSE`) for a sample of men.

2. set up the analysis (describe the comparison without using "Model A/C" language if you prefer);

   - A commonly cited "normal" resting pulse rate for men is 72. We want to assess whether this sample looks consistent with that reference value.

3. report mean (and SD), test statistic (t or F), PRE, p, and the confidence interval; and

   - mean = 53.7
   - SD = 8.3
   - F = 150.2
   - PRE = 0.83
   - p = 3.3e-13
   - CI = 50.7 - 56.8

4. state statistical and substantive conclusions.

   - Our results suggest we should reject the null hypothesis. This means that our sample of men has a statistically different resting heart rate than the commonly cited reference value of 72 in the population.

```r
fitness <- fitness |>
  mutate(rst_dev = RSTPULSE - 72)

model_c3 <- lm(rst_dev ~ 0, data = fitness)
model_a3 <- lm(rst_dev ~ 1, data = fitness)

summary(model_c3)
```

```
Call:
lm(formula = rst_dev ~ 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
```

```r
summary(model_a3)
```

```
Call:
lm(formula = rst_dev ~ 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
```

```r
sse_c3 <- deviance(model_c3)
sse_a3 <- deviance(model_a3)
```

```
pre3 <- (sse_c3 - sse_a3) / sse_c3
pre3
```

```
[1] 0.8335267
```

```
fstat3 <- (pre3/(PA - PC)) / ((1 - pre3)/(n3 - PA))
fstat3
```

```
[1] 150.2091
```

```
fcrit3 <- qf(.95, 1, 30)
p3 <- pf(fstat3, 1, 30, lower.tail = FALSE)

fcrit3
```

```
[1] 4.170877
```

```
p3
```

```
[1] 3.286974e-13
```

B1c. Estimate of true PRE

If you wanted to replicate this study, what would be your best guess for the *true* value of PRE?

Using the unbiased formula, it would be 0.8381779.

```
df1 <- 1
df2 <- n3 - df1 - 1

PRE_unbiased <- (fstat3 * df1) / (fstat3 * df1 + df2)
PRE_unbiased
```

```
[1] 0.8381779
```

B1d. Power for a replication at the same n

Given your estimate of true PRE, if you replicated with the same sample size, what would be your power to detect the effect? Use `pwr.f2.test()`. The power would be low to detect this due to the small sample size.

A common translation between PRE and Cohen's f^2 for this model-comparison setup is:

$$f^2 = \frac{\text{PRE}}{1 - \text{PRE}}.$$

```
true_PRE3 <- 0.06   # conservative estimate for planning replication
true_PRE3
```

```
[1] 0.06
```

```
f2_true3 <- true_PRE3 / (1 - true_PRE3)
f2_true3
```

```
[1] 0.06382979
```

```
power_replication <- pwr.f2.test(
  u = 1,
  v = n3 - 2,
  f2 = f2_true3,
  sig.level = 0.05
)

power_replication
```

```
        Multiple regression power calculation

              u = 1
              v = 29
             f2 = 0.06382979
      sig.level = 0.05
          power = 0.2748052
```

14

B2. Did running increase pulse rate? (within-person change)

Using the same dataset, compare resting pulse (RSTPULSE) to post-run pulse (RUNPULSE). Did running increase pulse rate?

Hint: create a new variable such as:

$$\text{change} = \text{RUNPULSE} - \text{RSTPULSE}.$$

B2a. 95% CI for b_0 in Model A

Construct a 95% CI for b_0 in your Model A for the change score. Do it by hand, then confirm with R. (111.8881 - 119.9184)

```
fitness <- fitness |>
  mutate(pulse_change = RUNPULSE - RSTPULSE)

summary_stats4 <- fitness |>
  summarize(mean_plsc = mean(pulse_change),
            var_plsc  = var(pulse_change),
            sd_plsc   = sd(pulse_change))
summary_stats4
```

```
# A tibble: 1 x 3
  mean_plsc var_plsc sd_plsc
      <dbl>    <dbl>   <dbl>
1      116.     120.    10.9
```

```
lower_bound4 <- summary_stats4$mean_plsc - (1.96*(summary_stats4$sd_plsc/sqrt(n3)))
upper_bound4 <- summary_stats4$mean_plsc + (1.96*(summary_stats4$sd_plsc/sqrt(n3)))

lower_bound4
```

```
[1] 112.0498
```

```
upper_bound4
```

```
[1] 119.7566
```

```
t.test(fitness$pulse_change)$conf.int
```

```
[1] 111.8881 119.9184
attr(,"conf.level")
[1] 0.95
```

B2b. Estimate of true PRE

If you wanted to replicate this study, what would be your best guess for the true value of PRE?

Using the unbiased formula, it would be 0.9917248.

```
model_c4 <- lm(pulse_change ~ 0, data = fitness)
model_a4 <- lm(pulse_change ~ 1, data = fitness)

sse_c4 <- deviance(model_c)
sse_a4 <- deviance(model_a)

pre4 <- (sse_c4 - sse_a4)/sse_c4

fstat4 <- (pre4/(PA - PC)) / ((1 - pre)/(n3 - PA))
fstat4
```

```
[1] 38.70229
```

```
df1 <- 1
df2 <- n3 - df1 - 1

PRE_unbiased2 <- (fstat4 * df1) / (fstat4 * df1 + df2)
PRE_unbiased2
```

```
[1] 0.5716541
```

B2c. Required effect size for power $= .80$ at n $= 100$

How big would true PRE need to be to yield power = .80 in a future study with $n = 100$ participants? Use `pwr.f2.test()`.

0.07412344

```
f2_needed <- pwr.f2.test(
  u = 1,
  v = 98,
  sig.level = 0.05,
  power = 0.80
)$f2

f2_needed
```

```
[1] 0.08005759
```

```
PRE_needed <- f2_needed / (1 + f2_needed)
PRE_needed
```

```
[1] 0.07412344
```

B3. Oral contraceptives and systolic blood pressure (Rosner example)

Ten women are studied to determine if oral contraceptives increase systolic blood pressure (mm Hg). Baseline blood pressure is measured before and then again several months after oral contraceptives are started. The following **changes** in systolic blood pressure are observed:

13, 3, -1, 9, 7, 7, 6, 4, -2, 2

A summary table (from SAS) reports:

Mean	Std Dev	USS	CSS	T	Prob >
4.8000000	4.5655716	418.0000000	187.6000000	3.3246511	0.0089

Here, **USS** ("uncorrected sum of squares") represents squared error around the model $Y_i = 0 + \varepsilon_i$ (i.e., Model C with intercept fixed at 0). **CSS** ("corrected sums of squares") represents squared error around $Y_i = b_0 + \varepsilon_i$ (Model A with estimated intercept).

B3a. Interpret the output using model comparison

Use this output to determine whether blood pressure significantly increased after oral contra-
ceptives were begun.

1. Specify Models A & C (equations are encouraged).

   - $D_i = \text{BPBEFORE}_i - \text{BPAFTER}_i$.
   - **Model C:** $D_i = 0 + \varepsilon_i$ (no free parameters)
   - **Model A:** $D_i = b_0 + \varepsilon_i$ (estimate the mean difference)

2. Specify the null hypothesis.

   - $H_0 : b_0 = 0$ there is no increase in blood pressure.

3. Calculate PRE and F.

   - F = 11.0533
   - PRE = 0.5511962

4. State your statistical and substantive one-sentence conclusions.

   - Due to the low p-value, we can conclude that there is a statistically significant
     change in blood pressure before and after taking oral contraceptives in this study.

```
fstat5 <- (3.3246511)^2

df1 <- 1
df2 <- 9

PRE5 <- (fstat5 * df1) / (fstat5 * df1 + df2)
PRE5
```

```
[1] 0.5511962
```

B3b. Estimate of true PRE

If you wanted to replicate this study, what would be your best guess for the true value of
PRE?

Using the unbiased formula, it would be 0.5801253.

```
df1 <- 1
df2 <- 10 - df1 - 1

PRE_unbiased3 <- (fstat5 * df1) / (fstat5 * df1 + df2)
PRE_unbiased3
```

```
[1] 0.5801253
```

B3c. Sample size for power targets

What sample size would you need to achieve power = .80? What sample size would you need to achieve power = .90? Use `pwr.f2.test()`.

1. n = 125
2. n = 167

```
f2 <- 0.0638
u <- 1
sig.level <- 0.05

sample_80 <- pwr.f2.test(u = u, f2 = f2, sig.level = sig.level, power = 0.80)
n_80 <- ceiling(sample_80$v + u + 1)  # convert v to n
n_80
```

```
[1] 125
```

```
sample_90 <- pwr.f2.test(u = u, f2 = f2, sig.level = sig.level, power = 0.90)
n_90 <- ceiling(sample_90$v + u + 1)  # convert v to n
n_90
```

```
[1] 167
```

B4. Power and design choices (GPA example)

In a particular year, a graduate program admitted 24 new students. The admissions committee wants to interpret the students' first-year GPA.

B4a. Approximate power for a "moderate" effect

Graduate students in this program are required to maintain a 3.0 average. The committee plans to compare:

- Model C: GPA fixed at 3.0
- Model A: GPA estimated as the students' mean

If the true difference corresponds to what the text refers to as a "moderate" effect in the social sciences, what are the chances that the committee will correctly conclude Model A is better than Model C if they use $\alpha = .05$?

1. Calculate approximate power using the book approach.
2. Verify using `pwr` (it will not match exactly, but should be close).

```
Fcrit <- qf(0.95, 1, 23)
Fcrit
```

```
[1] 4.279344
```

```
lambda <- 1.4375
power_book <- 1 - pf(Fcrit, 1, 23, ncp = lambda)
power_book
```

```
[1] 0.209715
```

```
f2 <- 0.0625
u <- 1
v <- 24 - u - 1

power_pwr <- pwr.f2.test(u = u, v = v, f2 = f2, sig.level = 0.05)$power
power_pwr
```

```
[1] 0.2161431
```

B4b. Two ways to increase power without increasing n

Given that they cannot admit additional students, state two things they could do to increase power.

The could reduce error or increase alpha.

B4c. Try-out question: setting up a conditional model comparison

The committee wants to know whether GRE-Verbal, GRE-Math, and undergraduate GPA are useful predictors of first-year GPA.

We have not yet learned how to *test* these conditional models, but set up the comparison:

- Model C: predicts the mean
    - **Model C:** $FYGPA_i = b_0 + \varepsilon_i$
- Model A: predicts GPA conditionally using all 3 predictors
    - **Model C:** $FYGPA_i = b_0 + b_1, \text{GRE\_Verbal}_i + b_2, \text{GRE\_Math}_i + b_3, \text{UGPA}_i + \varepsilon_i$
    - $n - P_A$ = 20
    - $P_A - P_C$ = 3

Specify both models and indicate $n - P_A$ and $P_A - P_C$ for this comparison.

Part C: Your own data (leave code blocks empty)

These questions refer to the model comparison you tested previously using your own dataset.

C1. Confidence interval for b_0

Identify the 95% confidence interval for b_0 in Model A. Explain what this interval represents.

(2.155336 - 2.244099)

This is where in a normal distribution based on our standard deviation and mean from the sample where 95% of our data would lie.

```
gss2024 <- readRDS(file = here::here("data", "gss2024.rds"))

mydata <- gss2024 |>
  select(pillok) |>
  drop_na() |>
  haven::zap_labels()

mydata <- mydata |>
```

```
    mutate(pill_dev = pillok - 2)

model_c6 <- lm(pill_dev ~ 0, data = mydata)
model_a6 <- lm(pill_dev ~ 1, data = mydata)

summary(model_c6)
```

```
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
```

```
summary(model_a6)
```

```
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
```

```
t.test(mydata$pillok)$conf.int
```

```
[1] 2.155336 2.244099
```

```
attr(,"conf.level")
[1] 0.95
```

```
sse_c6 <- deviance(model_c6)
sse_a6 <- deviance(model_a6)

pre6 <- (sse_c6 - sse_a6) / sse_c6
pre6
```

```
[1] 0.03540141
```

```
n6<- nrow(mydata)
fstat6 <- (pre6/(PA - PC)) / ((1 - pre6)/(n6 - PA))
fstat6
```

```
[1] 77.87881
```

```
fcrit6 <- qf(.95, 1, 2123)
qf(.95, 1, 99)
```

```
[1] 3.937117
```

```
qpre <- function(p,
                 df1,
                 df2) {

  f <- qf(p = p,
          df1 = df1,
          df2 = df2)

  crit_pre <- f / (f + df2/df1)

  return(crit_pre)

}

PRE_crit <- qpre(p = .95, df1 = 1, df2 = 2122)

p6 <- pf(fstat6, 1, 2123, lower.tail = FALSE)
```

C2. Journal-style summary

Write a complete journal-style summary for your data.

Due to our p-value, we reject the null hypothesis that the average response to this question was 2.

C3. Estimate of true PRE

If you wanted to replicate this study, what would be your best guess for the true value of PRE?

```
df1 <- 1
df2 <- 2123 - df1 - 1

PRE_unbiased6 <- (fstat6 * df1) / (fstat6 * df1 + df2)
PRE_unbiased6
```

```
[1] 0.03541751
```

C4. Sample size for power = .80

Given your estimate of true PRE, what sample size would you need to achieve power = .80? Use `pwr.f2.test()`.

216

```
f2 <- PRE_unbiased6 / (1 - PRE_unbiased6)
f2
```

```
[1] 0.03671797
```

```
u <- 1
f2 <- PRE_unbiased6 / (1 - PRE_unbiased6)
sig.level <- 0.05
power <- 0.80

result <- pwr.f2.test(u = u, v = NULL, f2 = f2, sig.level = sig.level, power = power)
result
```

```
   Multiple regression power calculation

          u = 1
          v = 213.6984
         f2 = 0.03671797
  sig.level = 0.05
      power = 0.8
```

```
n_required <- result$v + u + 1
n_required
```

```
[1] 215.6984
```

C5. Sketch the PRE distributions

Draw a chart that shows the distribution of PRE under both the H_0 and H_1 hypotheses. Label α, β, power, and the surprise point.

```
library(ggplot2)

PRE_obs <- pre6            # observed PRE
PRE_true <- PRE_unbiased6 # estimated true PRE
sd_pre <- 0.02             # approximate spread
alpha <- 0.05

# Grid of PRE values
pre_values <- seq(-0.02, 0.15, length.out = 500)

# Density under H0
dens_H0 <- dnorm(pre_values, mean = 0, sd = sd_pre)
df_H0 <- data.frame(PRE = pre_values, Density = dens_H0)

# Density under H1
dens_H1 <- dnorm(pre_values, mean = PRE_true, sd = sd_pre)
df_H1 <- data.frame(PRE = pre_values, Density = dens_H1)

# --- H0 plot ---
plot_H0 <- ggplot(df_H0, aes(x = PRE, y = Density)) +
  geom_line(color = "black", size = 1.2) +
  geom_vline(xintercept = PRE_crit, linetype = "dashed", color = "red") +
```
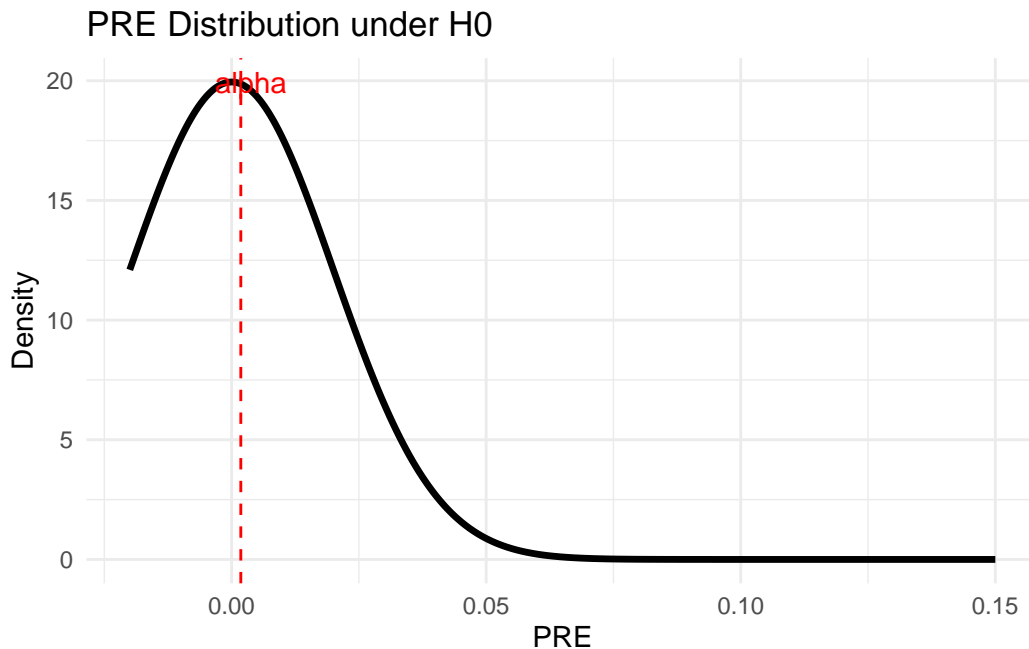
```
  labs(title = "PRE Distribution under H0",
       x = "PRE", y = "Density") +
  annotate("text", x = PRE_crit + 0.002, y = max(dens_H0), label = "alpha", color = "red") +
  theme_minimal()

# --- H1 plot ---
plot_H1 <- ggplot(df_H1, aes(x = PRE, y = Density)) +
  geom_line(color = "blue", size = 1.2) +
  geom_vline(xintercept = PRE_crit, linetype = "dashed", color = "red") +
  geom_vline(xintercept = PRE_obs, linetype = "dotted", color = "purple") +
  labs(title = "PRE Distribution under H1",
       x = "PRE", y = "Density") +
  annotate("text", x = PRE_crit + 0.002, y = max(dens_H1), label = "Power region", color = "d
  annotate("text", x = PRE_obs, y = max(dens_H1), label = "Observed PRE", color = "purple", v
  theme_minimal()

# --- Display plots ---
plot_H0
```

PRE Distribution under H0


```
plot_H1
```

PRE Distribution under H1