One-way ANOVA Homework

Caleb Rother

2025-10-22

```r
# remove # from line below if you need to install UNKstats package
# remotes::install_github("CalebRother/UNKstats")

library(tidyverse)
```

```
## -- 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    3.3.0
## v lubridate 1.9.4     v tidyr     1.3.1
## v purrr     1.2.0
## -- 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 become errors
```

```r
library(UNKstats)
```

Question 1:

Researchers want to test a new drug when given at different dosages. They are also including a placebo. They are measuring the score on a psychological examination, and are hoping to find the most effective dose to reduce score. Run the proper tests to determine if each group is unique and what effects they have.

```r
drug <- tibble(
  Dose = rep(c("Placebo","Low","Med","High"), each = 12),
  Score = c(
    # Placebo
    0.90,0.82,1.00,1.12,0.74,0.93,1.02,1.21,0.84,0.91,1.08,0.86,
    # Low
    0.10,0.00,-0.10,0.20,0.05,-0.05,0.15,0.00,-0.10,0.12,0.06,0.01,
    # Med
    -0.30,-0.42,-0.22,-0.50,-0.35,-0.27,-0.31,-0.46,-0.38,-0.20,-0.36,-0.29,
    # High
    0.40,0.32,0.50,0.36,0.25,0.46,0.41,0.52,0.35,0.30,0.44,0.39
  )
)
```

1. State the null and alternative hypotheses.

2. Are these data normal? How can you tell?

3. Which dosage/dosages are most effective at reducing the score?

```
## # A tibble: 2 x 6
##   term         df  sumsq meansq statistic   p.value
```

```
##   <chr>     <int>  <dbl>  <dbl>     <dbl>     <dbl>
## 1 Dose          3 10.9   3.62        339.  2.10e-30
## 2 Residuals    44  0.470 0.0107       NA  NA
```

4. What might the biological implications of this test be?

Question 2:

Professors at UNK are debating if labs should be kept in Biology 105. They want to see if exam scores are higher in classes taught with only lecture, only lab, or both. Run the proper tests to determine if each group is unique and what effects they have.

```r
teach <- tibble(
  Method = rep(c("Lab","Both","Lecture"), each = 15),
  Exam = c(
    # Lab
    76,79,82,77,75,80,81,78,76,79,77,83,74,80,78,
    # Both
    84,85,82,88,83,86,84,87,81,86,85,83,84,89,82,
    # Lecture
    80,82,79,83,81,80,82,81,84,78,83,80,81,79,82
  )
)
```

1. State the null and alternative hypotheses.
2. Are these data normal?
3. Which teaching method is most effective? How do you know this?

```
## # A tibble: 2 x 6
##   term         df sumsq meansq statistic  p.value
##   <chr>     <int> <dbl>  <dbl>     <dbl>    <dbl>
## 1 Method        2  297. 148.        30.1  7.72e-9
## 2 Residuals    42  207.   4.93      NA   NA
```

4. What might the teaching implications of this test be?

Question 3:

Researchers are testing weight gain drugs in comparison to each other. They are also comparing the effects of these drugs to a placebo. Run the proper tests to determine if each group is unique and what effects they have.

```r
diet <- tibble(
  Diet = rep(c("A","B","C", "Placebo"), each = 14),
  WeightGain = c(
    # A
    1.1,0.9,1.5,1.2,0.8,1.0,1.4,0.9,1.0,1.2,1.5,0.9,1.3,1.4,
    # B
    1.3,1.5,1.2,0.8,1.6,1.5,1.0,2.0,1.5,1.4,1.6,1.2,1.5,1.4,
    # C
    2.1,2.0,1.6,2.3,2.1,2.0,1.6,2.1,2.5,1.7,2.2,2.1,2.3,2.7,
    # Placebo
    0.2,0.2,-0.2,0.0,-0.8,-.4,0.4,0.2,1.0,0.0,-0.1,-0.2,-0.7,0.6
  )
)
```

1. State the null and alternative hypotheses.

2. Are these data normal?

```
diet |>
  group_by(Diet) |>
  summarise(p_value = shapiro.test(WeightGain)$p.value)
```

```
## # A tibble: 4 x 2
##   Diet    p_value
##   <chr>     <dbl>
## 1 A         0.228
## 2 B         0.527
## 3 C         0.457
## 4 Placebo   0.952
```

```
## # A tibble: 2 x 6
##   term         df sumsq meansq statistic   p.value
##   <chr>     <int> <dbl>  <dbl>     <dbl>     <dbl>
## 1 Diet          3 31.3  10.4        88.3  2.11e-20
## 2 Residuals    52  6.15  0.118      NA   NA
```

3. Interpret the meaning of the assigned letters. How were they found, and what do they mean?
4. Which diet plan was most effective for gaining weight?

Question 4:

A field study is measuring time-to-first-foraging (minutes) for four bird colonies. Run the proper tests to determine if each group is unique.

```
set.seed(4517)

n_per <- 18
Colonies_data <- data.frame(
  Colony = rep(c("North","East","South","West"), each = n_per),
  Minutes = c(
    round(rexp(n_per, rate = 1.0), 1),
    round(rexp(n_per, rate = 1.3), 1),
    round(rexp(n_per, rate = 0.7), 1),
    round(rexp(n_per, rate = 1.0), 1)
  )
)
```

1. State the null and alternative hypotheses.
2. Are these data normal?

```
## # A tibble: 2 x 6
##   term         df sumsq meansq statistic  p.value
##   <chr>     <int> <dbl>  <dbl>     <dbl>    <dbl>
## 1 Colony        3  4.35  1.45       5.33  0.00234
## 2 Residuals    68 18.5   0.272     NA    NA
```

3. Interpret the meaning of the assigned letters. How were they found, and what do they mean?
4. Which colony/colonies were the fastest foragers?

Question 5:

A team of behavioral ecologists recorded the number of aggressive interactions among four populations of prairie voles raised in different social environments. They want to know if social upbringing influences aggression levels in male voles.

```r
set.seed(9913)

n_per <- 16
vole <- tibble(
  Environment = rep(c("Isolated","Paired","Group","Crowded"), each = n_per),
  Aggression = c(
    round(rnorm(n_per, mean = 8, sd = 3)),
    round(rnorm(n_per, mean = 6, sd = 1.5)),
    round(rnorm(n_per, mean = 4, sd = 2)),
    round(rnorm(n_per, mean = 7, sd = 4))
  )
)
```

1. State the null and alternative hypotheses.
2. Are these data normal?

```
## # A tibble: 2 x 6
##   term           df sumsq meansq statistic p.value
##   <chr>       <int> <dbl>  <dbl>     <dbl>   <dbl>
## 1 Environment     3  24.4   8.12     0.776   0.512
## 2 Residuals      60 629.   10.5     NA      NA
```

3. Interpret the meaning of the assigned letters. How were they found, and what do they mean?
4. What else is unique about this data and might explain the results? *Hint*: Look at the final plot