Homework 4: Analysis of variance

STAT218

2024-02-22

library(oibiostat)  
library(openintro)  
library(Sleuth3)  
library(tidyverse)  
library(pander)  
library(emmeans)

1. [L5] Gifted children. The dataset openintro::gifted contains data on 36 children identified as gifted in a large city. This problem will consider only the father’s and mother’s IQ: you’ll perform inference on the difference in IQ between the mother and father. Assume the population is gifted children in the metropolitan region where the data were collected.
   1. Conduct an appropriate test to determine whether mean IQ differs between the father and mother. Interpret the results in context.
   2. Use a nonparametric rank procedure to test for a difference in IQ center between the father and mother. Interpret the results in context.
   3. Explain how the hypotheses differ between the test in parts (a) and (b).
   4. Make the appropriate histogram(s) to check whether the assumptions for the rank test are reasonable. Do you think the rank test is appropriate in this case?

* ANSWER:
  1. The data provide sufficient evidence to reject the null hypothesis that parents of gifted children have the same mean IQ in favor of the alternative hypothesis that the mean IQ differs (t = -2.7279, df = 35, p-value = 0.009898).
  2. The data provide sufficient evidence to reject the null hypothesis that parents of gifted children have the same typical IQ in favor of the alternative hypothesis that the typical IQ differs (p-value = 0.009677).
  3. The hypothesis in (a) pertains to a difference in mean IQ; the hypothesis in (b) pertains to a difference in central or typical IQ.
  4. The differences don’t look especially symmetric; the rank test is probably not appropriate.

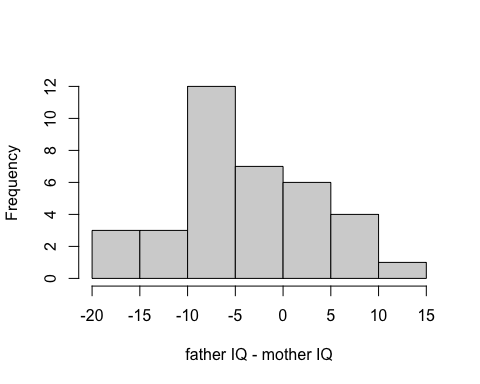
# part a  
t.test(Pair(fatheriq, motheriq) ~ 1, data = gifted)

Paired t-test  
  
data: Pair(fatheriq, motheriq)  
t = -2.7279, df = 35, p-value = 0.009898  
alternative hypothesis: true mean difference is not equal to 0  
95 percent confidence interval:  
 -5.910883 -0.866895  
sample estimates:  
mean difference   
 -3.388889

# part b  
wilcox.test(Pair(fatheriq, motheriq) ~ 1, data = gifted)

Wilcoxon signed rank test with continuity correction  
  
data: Pair(fatheriq, motheriq)  
V = 168, p-value = 0.009677  
alternative hypothesis: true location shift is not equal to 0

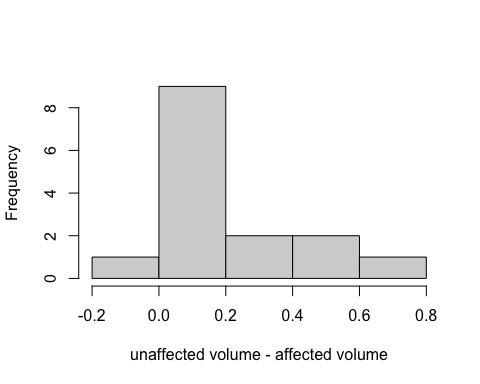
# part d  
(gifted$fatheriq - gifted$motheriq) |>   
 hist(breaks = 8, main = '', xlab = 'father IQ - mother IQ')



1. [L5] Physiological indicators of schizophrenia. The dataset Sleuth3::case0202 contains data on volumes of the left hippocampus in cubic centimeters for 15 pairs of monozygotic twins; one twin in each pair was schizophrenic and the other was not.
   1. Check the assumptions for a nonparametric test of whether hippocampus volume differs in center between the schizophrenic and non-schizophrenic twins. Show the graphic(s) you examined and comment on whether test assumptions seem plausible.
   2. Perform the test in (a) and interpret the results. Write a full narrative summary and include a 90% confidence interval for the difference in centers.

* ANSWERS:
  1. The plot is shown below; pairwise differences appear to be unimodal and not too asymmetrically distributed. Test assumptions seem plausible.
  2. Results suggest brain anatomy differs with schizophrenia. We used data on twins in which one twin was affected by schizophrenia and one was not to test the hypothesis that left hippocampal volume differed between the affected and unaffected twin. The data provide sufficiently strong evidence to reject the hypothesis that typical left hippocampal volume does not differ between affected and unaffected twins in favor of the alternative of a difference (p-value = 0.002014). With 90% confidence, the difference in central values of left hippocampal volume is estimated to be between 0.075 and 0.315 cubic centimeters.

# part a  
(case0202$Unaffected - case0202$Affected) |>   
 hist(breaks = 5, main = '', xlab = 'unaffected volume - affected volume')



# part b  
wilcox.test(Pair(Unaffected, Affected) ~ 1, data = case0202,   
 conf.int = T, conf.level = 0.9)

Wilcoxon signed rank exact test  
  
data: Pair(Unaffected, Affected)  
V = 111, p-value = 0.002014  
alternative hypothesis: true location shift is not equal to 0  
90 percent confidence interval:  
 0.075 0.315  
sample estimates:  
(pseudo)median   
 0.1575

1. **Vu and Harrington exercise 5.48.** Determine if the following statements are true or false, and explain your reasoning for statements you identify as false. If the null hypothesis that the means of four groups are all the same is rejected using ANOVA at a 5% significance level, then…
   1. we can then conclude that all the means are different from one another.
   2. the standardized variability between groups is higher than the standardized variability within groups.
   3. a pairwise analysis will identify at least one pair of means that are significantly different.
2. [L9] Consider an experiment in which overweight subjects were randomly allocated to one of three diets: low-fat, low-carb, and Mediterranean. After two years, each subject’s weight loss was measured. The output below shows the result of an ANOVA model fitted to this data.
   1. Construct the ANOVA table.
   2. Test the hypothesis that mean weight loss does not differ among the three diets. Write a narrative summary of the result. (*Hint:* the -value from an F model can be computed as pf(fstat, num.df, denom.df, lower.tail = F).)

* ANSWERS:
  1. The table is shown below.
  2. The data provide sufficiently strong evidence to reject the null hypothesis of no effect of diet on weight loss (F = 3.236 on 2 and 269 df, p = 0.0409).

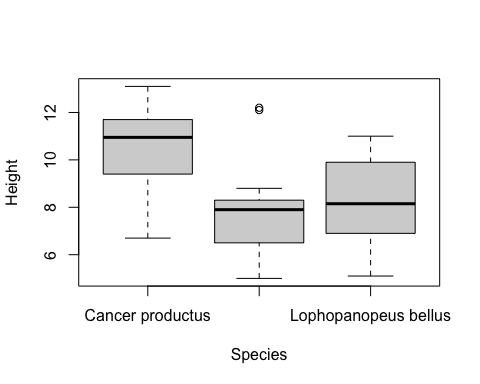
weightloss <- ex0623  
aov(WtLoss24 ~ Group, data = weightloss) |> summary()

Df Sum Sq Mean Sq F value Pr(>F)   
Group 2 217 108.43 3.236 0.0409 \*  
Residuals 269 9014 33.51   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

1. **Vu and Harrington exercise 5.40.** A study compared five different methods for teaching descriptive statistics. The five methods were traditional lecture and discussion, programmed textbook instruction, programmed text with lectures, computer instruction, and computer instruction with lectures. 45 students were randomly assigned, 9 to each method. After completing the course, students took a 1-hour exam.
   1. What are the hypotheses for evaluating if the average test scores are different for the different teaching methods?
   2. What are the degrees of freedom associated with the F-test for evaluating these hypotheses?
   3. Suppose the p-value for this test is 0.0168. What is the conclusion?
2. [L9] The dataset Sleuth3::ex0722 contains measurements of closing force and propodus height of the claws of crabs from three predatory species.
   1. Produce side-by-side boxplots of claw height by species, ordering species by increasing median claw height.
   2. Produce a table of grouped summaries of claw heights showing means, standard deviations, and sample sizes for each species.
   3. Comment on whether the assumptions for analysis of variance seem to be met.
   4. Test whether mean claw height differs among the three species. Write a short narrative summary of the results.
   5. Conduct post-hoc pairwise comparisons to determine which means differ.

* ANSWERS:
  1. Boxplots are shown below.
  2. Table is shown below.
  3. The standard deviations of each group are close enough; measurements are independent since they are taken on distinct individuals.
  4. The data provide sufficiently strong evidence to reject the null hypothesis of no difference in mean claw height among the three species of crab (F = 5.898 on 2 and 35 df, p = 0.0062).
  5. Post-hoc tests for pairwise comparisons are shown below. It is estimated that mean claw height differs between *cancer productus* and each of the other two species (*hemigraspus nudus* and *lophapanopeus bellus*), but not between *hemigraspus nudus* and *lophapanopeus bellus*. Student answers should include p-values if a table is not provided.

# part a  
boxplot(Height ~ Species, data = ex0722)



# part b  
ex0722 |>  
 group\_by(Species) |>  
 summarize(mean = mean(Height),  
 sd = sd(Height),  
 n = n()) |> pander()

| Species | mean | sd | n |
| --- | --- | --- | --- |
| Cancer productus | 10.43 | 1.987 | 12 |
| Hemigrapsus nudus | 7.921 | 2.074 | 14 |
| Lophopanopeus bellus | 8.242 | 1.853 | 12 |

# part d  
fit <- aov(Height ~ Species, data = ex0722)  
summary(fit) |> pander()

Analysis of Variance Model

|  | Df | Sum Sq | Mean Sq | F value | Pr(>F) |
| --- | --- | --- | --- | --- | --- |
| **Species** | 2 | 46.23 | 23.11 | 5.898 | 0.0062 |
| **Residuals** | 35 | 137.2 | 3.919 | NA | NA |

# part e  
emmeans(fit, 'Species') |> pairs(adjust = 'tukey') |> test() |> pander()

| contrast | estimate | SE | df | t.ratio | p.value |
| --- | --- | --- | --- | --- | --- |
| Cancer productus - Hemigrapsus nudus | 2.504 | 0.7788 | 35 | 3.215 | 0.007706 |
| Cancer productus - Lophopanopeus bellus | 2.183 | 0.8082 | 35 | 2.702 | 0.02782 |
| Hemigrapsus nudus - Lophopanopeus bellus | -0.3202 | 0.7788 | 35 | -0.4112 | 0.9112 |