Beckham Carver

3/27/22

STAT 4025

Dr. Robinson

Homework 8

1. An experiment was conducted to determine the effects of four pesticides on the yield of fruit from three different varieties (B1,B2,B3) of a citrus tree. Eight trees from each variety were selected from an orchard. The four pesticides were then assigned to two trees of each variety and applications were made according to recommended levels. Yields of fruit (in bushels per tree) were obtained after the test period. The data are provided below:
2. If this is a completely randomized design, how would you set up this experiment? (3 pts)

I would randomly select and group eight trees from each variety, which makes this an interventional study. Each pesticide would be randomly assigned twice to each group. This fulfills randomized selection and RATG.

1. Write the effects model for this experiment (1 pt), explain the terms (1 pt), and compute the analysis of variance table for the data. Compute your ANOVA table by hand (8 pts) but then check it against what you get with R (2 pt). Supply your R code and only the relevant output. [12 pts total]

Yijk = μ+ αi+ βj+ (αβ)ij+ εijk

In this equation ‘a’ represents the main affect (variety), β represents the marginal affect (pesticide), aβ represents their interaction, and ε represents the experimental variance within the data.

**Variety Means:**

* Variety 1: (49+39+50+55+43+38+53+48) / 8 = 47
* Variety 2: (55+41+67+58+53+42+85+72) / 8 = 59
* Variety 3: (66+68+85+92+69+62+85+99) / 8 = 78
* Gran mean: (47+59+78) / 3 = 61
* Sum of Squares: 8(47-61)2 + 8(59-61)2 + 8(78-61)2 = 3912

**Pesticide Means:**

* Pesticide 1: (49+ 39+55+41+66+68) / 6 = 53
* Pesticide 2: (50+ 55+67+58+85+92) / 6 = 68
* Pesticide 3: (43+ 38+53+42+69+62) / 6 = 51
* Pesticide 4: (53+ 48+85+73+85+99) / 6 = 74
* Gran mean: (53 + 68 + 51+ 74) / 4 = 61
* Sum of Squares: 6(53-61)2 +6(68-61)2 +6(51-61)2 + 6(74-61)2 = 2292

**Sum of Squares:**

* V1P1: (49-44)2 + (39-44)2 = 50
* V2P1: (55-48)2+(41-48)2 = 98
* V3P1: (66-67)2 + (68-67)2 = 2
* V1P2: (50-52.5)2 + (55-52.5)2 = 25
* V2P2: (67-62.5)2 + (58-62.5)2 = 81
* V3P2: (85-88.5)2 + (92-88.5)2 = 49
* V1P3: (43-40.5)2 + (38-40.5)2 = 25
* V2P3: (53-47.5)2 + (42-47.5)2 = 22
* V3P3: (69-65.5)2 + (62-65.5)2 =14
* V1P4: (53-50.5)2 + (48-50.5)2 = 10
* V2P4: (85-79)2 + (73-79)2 = 24
* V3P4: (85-92)2 + (99- 92)2 = 28
* Sum of Squares within: (50+98+2+25+81+49+25+22+14+10+24+28) = 428
* Total sum of Squares: … 7193
* Sum or Squares Interaction: 7193 – 3912 – 2292 – 428 = 561

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Source of Variation | SS | df | MS | F | p-value |
| Variety | 3912 | 2 | 1956 | 105.16 | <.00001 |
| Pesticide | 2292 | 3 | 764 | 41.07 | <. 00001 |
| Interaction | 561 | 6 | 93.5 | 5.03 | .001998 |
| Within | 428 | 12 | 18.6 | .. | .. |
| Total | 7913 | 23 | .. | .. | .. |

(See call for code)

Text

Description automatically generated

1. Construct the interaction plot (1 pt). Do you feel that there is a significant interaction between Pesticide and Variety? Explain (2 pts). [3 pts]

Chart

Description automatically generated

There appears to be some interaction between the pesticides and variety. Pesticide 2 and 4’s effect on fruit yield changes from favoring pesticides 2, to greatly favoring 4 and then normalizes again, which shows an interaction between varieties and pesticides. Yield in general also has variance between varieties.

1. Write the indicator model that R fits to this data when using the lm() function – 2 pts. Using appropriate symbols, write the cell mean for Pesticide 2 and Variety 2 using the indicator model (i.e. the betas) – 2 pts. Write out the marginal mean for Variety 2 using the appropriate indicator symbols (i.e. the betas) – 3 pts. [7 pts total]

M = β0+ β1 + β2 + β12

μ ̂V2p2: 44 + 4 + 8.5 + 6

(See call for code)

A screenshot of a computer

Description automatically generated with medium confidence

1. Conduct the multi-factor ANOVA using R. What are your conclusions based on the p-values? What do you recommend as the next step of your analysis? [3 pts]

(See call for code)

Text

Description automatically generated

With p-values < 0.05 we can conclude that there is a statistically significant difference in the mean fruit yield by varieties. There is also a significant difference in fruit yield across the pesticides.

1. Conduct a residual analysis to check the normality assumption. Check the equal variance assumption. [2 pts]

The normality assumption is met by our Shapiro wilks test. Our qq-plot also supports equal variance.

(See call for code)

Text

Description automatically generated

Chart, scatter chart

Description automatically generated