Checking the number of crops grown using Latin Square Design

Srikar

2021-09-22

Introduction:

ANOVA is a statistical test used to determine whether or not there is a significant difference between the means of treatments. The analysis of variance is the systematic algebraic procedure of decomposing (i.e. Partitioning) overall variation. If we find there is significant difference, we conduct a post-hoc test to check where the difference occurs. post hoc analysis consists of statistical analyses that were specified after the data were seen.

When the experimental material is divided into rows and columns and the treatments are llocated such that each treatment occurs only once in a row and once in a column, the design is known as Latin Square Design (LSD). Latin square designs allow for two blocking factors. In other words, these designs are used to simultaneously control (or eliminate) two sources of nuisance variability.

Objective:

i) To check whether the different fertilizers yield the same amount of crops
ii) To check whether the different tillage methods yield the same amount of crops
iii) To check whether the different kind of seeds yield the same amount of crops

Data Description

The dataset contains the productivity of 5 seeds grown using the different types of Fertilizers used (Fertilizer 1,2,3,4 and 5) and the different types of tillage methods which is the treatment on the seed. The different types of tillage methods are (A,B,C,D and E).Here tillage method is the treatment used on the seeds.The different types of seeds that are sowed are Seed (A,B,C,D and E).The frequency shows the number of crops that are grown using the combinations of different types of seeds, fertilizers and tilliage methods. The data totally contains 25 observations The frequency is measured in cwt/year.

~1 CWT {hundredweight} = 100pounds ~

Data Summary

```
library(readxl)
data<- read_excel("C:/Users/Srikar/Desktop/SS/R/Sem 5/Design of Exp/Practical 8/dataset.xlsx")
head(data,6)
```

Table 1: Dataset

	treatment	fertilizer	seed	freq
1	treatA	fertil1	A	42
2	treatA	fertil2	E	45
3	treatA	fertil3	C	41
4	treatA	fertil4	B	56
5	treatA	fertil5	D	47
6	treatB	fertil1	C	47

```
summary(data)
```

```
##   treatment          fertilizer            seed                freq      
##  Length:25          Length:25          Length:25          Min.   :41.00  
##  Class :character   Class :character   Class :character   1st Qu.:45.00  
##  Mode  :character   Mode  :character   Mode  :character   Median :48.00  
##                                                           Mean   :48.56  
##                                                           3rd Qu.:52.00  
##                                                           Max.   :57.00
```

The 4 variables are treatment, fertilizer, seed and frequency. We observe that the range of production is from 41 to 57 cwt/year. All variables are character types except frequency which is of integer type.

Hypothesis Statement:

1) Statement with respect to fertilizer

Null Hypothesis (H0): There is no significant difference in the growth by different fertilizers ($\mu1 = \mu2 = \mu3 = \mu4 = \mu5$) where (μ1, μ2, μ3, μ4, μ5) are mean crop growth

Alternative Hypothesis (H1): At least two fertilizers have significant difference.

$(\mu1 \neq \mu2 \neq \mu3 \neq \mu4 \neq \mu5)$

#2)Statement with respect to tillage

Null Hypothesis (H 0): There is no significant difference in the growth by different tilliage methods **(μ1 = μ2=μ3= μ4= μ5)** where (μ1, μ2, μ3, μ4, μ5) are mean crop growth

Alternative Hypothesis (H 1): At least two tillage methods have significant difference**(μ1 ≠ μ2≠μ3≠ μ4≠ μ5)**

Keeping the significance level as 5% or 0.05

Procedure:

1)Constructing the ANOVA model

```
mod=lm(data$freq~data$fertilizer+data$treatment+data$seed)
aov=anova(mod)
aov
```

Table 2:ANOVA table for crop growth

Source of Variation	Df	Sum Sq	Mean Sq	F value	P-Value
data$fertilizer	4	17.76	4.44	0.7967	0.549839
data$treatment	4	109.36	27.34	4.9055	0.014105 *
data$seed	4	286.16	71.54	12.8361	0.000271 ***
Residuals	12	66.88	5.573		

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

We observe that there is a significant difference between the different types of treatment which is the tillage method as the p-value is below 0.05 significance level. This means that different tilliage methods yeild to different amounts of crops grown.

Similarly we find that the seeds sown are also significantly different. The different seeds sown yeild to different amount of crops grown. Thus, we reject the null-hypothesis for treatment and seed. There is no difference in the fertilizers used as the p-value shows #no significance as its above 0.05. The different fertilizers have the same effect on the number of crops grown.

2) Performing our post-hoc test

We perform post-hoc test as we find there is a significant difference in treatment(tilliage methods) and seeds sown.

```
library(lsmeans)

x=lsmeans(mod,"treatment")
pairs(x)
```

Table 3: Test for significant difference of tillage methods

contrast	estimate	SE	df	t.ratio	p.value
treatA - treatB	-3.4	1.49	12	-2.277	0.2175
treatA - treatC	-5.4	1.49	12	-3.617	0.024
treatA - treatD	-3	1.49	12	-2.009	0.3182
treatA - treatE	0	1.49	12	0	1
treatB - treatC	-2	1.49	12	-1.339	0.6738
treatB - treatD	0.4	1.49	12	0.268	0.9987
treatB - treatE	3.4	1.49	12	2.277	0.2175
treatC - treatD	2.4	1.49	12	1.607	0.5201
treatC - treatE	5.4	1.49	12	3.617	0.024
treatD - treatE	3	1.49	12	2.009	0.3182

```
##
## Results are averaged over the levels of: fertilizer, seed
## P value adjustment: tukey method for comparing a family of 5 estimates
```

We observe that there is a difference between treatment A and treatment C.We can say that on an average, C yields more than A. There is also a difference between treatment C and treatment E.We can say that on an average C yeilds more corps than E.

```
y=lsmeans(mod,"seed")
pairs(y)
```

	Table 4: Test for significant difference of different seeds				
contrast	estimate	SE	df	t.ratio	p.value
A - B	-9.4	1.49	12	-6.296	0.0003
A - C	-3.2	1.49	12	-2.143	0.2642
A - D	-7.4	1.49	12	-4.956	0.0025
A - E	-2.8	1.49	12	-1.875	0.3793
B - C	6.2	1.49	12	4.152	0.0096
B - D	2	1.49	12	1.339	0.6738
B - E	6.6	1.49	12	4.42	0.0061
C - D	-4.2	1.49	12	-2.813	0.0936
C - E	0.4	1.49	12	0.268	0.9987
D - E	4.6	1.49	12	3.081	0.0599

```
## Results are averaged over the levels of: fertilizer, treatment
## P value adjustment: tukey method for comparing a family of 5 estimates
```

We observe that Seed A and seed B differ amongst each other. Seed A and Seed D are significantly different. Seed B and seed C are significantly different from each other. Seed B and Seed C are significantly different from each other and lastly. Seed B and Seed E are significantly different from each other as their p-values are below 0.05.

Conclusion:

1) All fertilizers have the same effect on crops

2) Tillage method A and C differ where C is better than A on an average. Tillaige method C is better than method E. Hence, there is a significant difference between the methods

3) The different types of seeds yeild different amount of crops. We find that on an average A yeilds the least and Seed B and D yield the most.