Automatic report for a Randomized Complete Block Design (RCBD)

AgroFIMS

January 22, 2021

Table of Contents

# 1. Model specification and data description

Data from 5 Treatments have been evaluated using a randomize complete block design with 2 blocks. The statistical model is

$$y\_{ij}=μ+τ\_{i}+β\_{j}+ϵ\_{ij}$$

where

* $y\_{ij}$ is the observed response with Treatment $i$ and block $j$.
* $μ$ is the mean response over all Treatments and blocks.
* $τ\_{i}$ is the effect for Treatment $i$.
* $β\_{j}$ is the effect for block $j$.
* $ϵ\_{ij}$ is the error term.

In this model we assume that the errors are independent and have a normal distribution with common variance, that is, $ϵ\_{ij}∼N(0,σ\_{ϵ}^{2})$.

# 2. Analysis for trait common\_bean\_grain\_fresh\_weight\_g

## 2.1. Exploratory analysis

It is always good to have some visualization of your data. Below a histogram and a boxplot are shown.

# Keep a copy of full data
mydata.full <- mydata
# Get means for subsampling
mydata <- st4gi::docomp("mean", trait, c(treatment, experimental.unit), rep, mydata)

par(mfrow = c(1, 2))
hist(mydata$trait)
boxplot(mydata$trait)



Since the number of Treatments in your experiment is not so large, we can plot the data for each Treatment:

st4gi::msdplot(trait, treatment, mydata, conf = 1, pch = 4)



## 2.2. ANOVA

You have fitted a linear model for a RCBD. The ANOVA table for your model is:

at <- aov.rcbd(trait, treatment, rep, mydata)
# Anova table
at

Analysis of Variance Table

Response: "common\_bean\_grain\_fresh\_weight\_g"
 Df Sum Sq Mean Sq F value Pr(>F)
treatment 4 148830 37208 3.0330 0.1540
block 1 12180 12180 0.9929 0.3754
Residuals 4 49070 12268

The coefficient of variation for this experiment is 38.85%. The p-value for Treatments is 0.154 which is not significant at the 5% level.

## 2.3. Assumptions

Don’t forget the assumptions of the model. It is supposed that the errors are independent with a normal distribution and with the same variance for all the Treatments. The following plots can help you evaluate this:

par(mfrow = c(1, 2))
plot(model, which = 1)
plot(model, which = 2)



Any trend in the residuals in the left plot would violate the assumption of independence while a trend in the variability of the residuals –for instance a funnel shape– suggests heterogeneity of variances. Departures from the theoretical normal line on the right plot are symptoms of lack of normality.

## 2.4. Treatment means

Because the effect of Treatments was not significant in the ANOVA, multiple comparison tests are not presented. The means of your Treatments are:

tapply(mydata$trait.est, mydata$treatment, mean, na.rm = TRUE)

G1\_/\_10 G1\_/\_30 G2\_/\_10 G2\_/\_30 G3\_/\_10
 150.0 325.0 200.5 250.0 500.0

## 2.5. Variance components

Below are the variance components for this model, under the assumption that Treatments and blocks are random. Here the model is fitted using REML and missing values are not estimated.

model <- lme4::lmer(trait ~ (1|treatment) + (1|rep), data = mydata)
vc <- data.frame(lme4::VarCorr(model))
vc[, c(1, 4, 5)]

boundary (singular) fit: see ?isSingular

 Variance Std.Dev.
treatment 1.247952e+04 1.117118e+02
block 2.549076e-05 5.048837e-03
Residual 1.224978e+04 1.106787e+02

# 3. Analysis for trait groundnut\_total\_biomass\_dry\_weight\_g

## 3.1. Exploratory analysis

It is always good to have some visualization of your data. Below a histogram and a boxplot are shown.

# Keep a copy of full data
mydata.full <- mydata
# Get means for subsampling
mydata <- st4gi::docomp("mean", trait, c(treatment, experimental.unit), rep, mydata)

par(mfrow = c(1, 2))
hist(mydata$trait)
boxplot(mydata$trait)



Since the number of Treatments in your experiment is not so large, we can plot the data for each Treatment:

st4gi::msdplot(trait, treatment, mydata, conf = 1, pch = 4)



## 3.2. ANOVA

You have fitted a linear model for a RCBD. The ANOVA table for your model is:

at <- aov.rcbd(trait, treatment, rep, mydata)
# Anova table
at

Analysis of Variance Table

Response: "groundnut\_total\_biomass\_dry\_weight\_g"
 Df Sum Sq Mean Sq F value Pr(>F)
treatment 4 3003503 750876 29.9444 0.003065 \*\*
block 1 24701 24701 0.9851 0.377133
Residuals 4 100303 25076
---
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

The coefficient of variation for this experiment is 4.872%. The p-value for Treatments is 0.003065 which is significant at the 5% level.

## 3.3. Assumptions

Don’t forget the assumptions of the model. It is supposed that the errors are independent with a normal distribution and with the same variance for all the Treatments. The following plots can help you evaluate this:

par(mfrow = c(1, 2))
plot(model, which = 1)
plot(model, which = 2)



Any trend in the residuals in the left plot would violate the assumption of independence while a trend in the variability of the residuals –for instance a funnel shape– suggests heterogeneity of variances. Departures from the theoretical normal line on the right plot are symptoms of lack of normality.

## 3.4. Treatment means

Below are the sorted means for each Treatment using the Fisher’s Least Significant Difference method and the multiple comparisons method of Tukey, both at the 5% level. Letters indicate if there are significant differences.

### 3.3.1. LSD test

# If there are no missing values, get a copy for further use
mydata$trait.est <- mydata.trait
# If there are missing values, get estimates
mydata$trait.est <- mve.rcbd(trait, treatment, rep, mydata)[, trait.est]
# Run LSD test
agricolae::LSD.test(mydata$trait.est, mydata$treatment, at[3, 1], at[3, 3])$groups

 means groups
G3\_/\_10 4001.0 a
G1\_/\_30 3750.0 a
G1\_/\_10 3000.5 b
G2\_/\_30 3000.5 b
G2\_/\_10 2499.5 c

### 3.3.2. Tukey test

agricolae::HSD.test(mydata$trait.est, mydata$treatment, at[3, 1], at[3, 3])$groups

 means groups
G3\_/\_10 4001.0 a
G1\_/\_30 3750.0 a
G1\_/\_10 3000.5 b
G2\_/\_30 3000.5 b
G2\_/\_10 2499.5 b

### 3.3.3. Plot of means

## 3.5. Variance components

Below are the variance components for this model, under the assumption that Treatments and blocks are random. Here the model is fitted using REML and missing values are not estimated.

model <- lme4::lmer(trait ~ (1|treatment) + (1|rep), data = mydata)
vc <- data.frame(lme4::VarCorr(model))
vc[, c(1, 4, 5)]

boundary (singular) fit: see ?isSingular

 Variance Std.Dev.
treatment 3.629368e+05 6.024424e+02
block 8.619374e-06 2.935877e-03
Residual 2.500073e+04 1.581162e+02

# 4. Analysis for trait groundnut\_whole\_plant\_number\_of\_nodules

## 4.1. Exploratory analysis

It is always good to have some visualization of your data. Below a histogram and a boxplot are shown.

# Keep a copy of full data
mydata.full <- mydata
# Get means for subsampling
mydata <- st4gi::docomp("mean", trait, c(treatment, experimental.unit), rep, mydata)

par(mfrow = c(1, 2))
hist(mydata$trait)
boxplot(mydata$trait)



Since the number of Treatments in your experiment is not so large, we can plot the data for each Treatment:

st4gi::msdplot(trait, treatment, mydata, conf = 1, pch = 4)



## 4.2. ANOVA

You have fitted a linear model for a RCBD. The ANOVA table for your model is:

at <- aov.rcbd(trait, treatment, rep, mydata)
# Anova table
at

Analysis of Variance Table

Response: "groundnut\_whole\_plant\_number\_of\_nodules"
 Df Sum Sq Mean Sq F value Pr(>F)
treatment 4 440 110 11 0.01968 \*
block 1 10 10 1 0.37390
Residuals 4 40 10
---
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

The coefficient of variation for this experiment is 15.06%. The p-value for Treatments is 0.01968 which is significant at the 5% level.

## 4.3. Assumptions

Don’t forget the assumptions of the model. It is supposed that the errors are independent with a normal distribution and with the same variance for all the Treatments. The following plots can help you evaluate this:

par(mfrow = c(1, 2))
plot(model, which = 1)
plot(model, which = 2)



Any trend in the residuals in the left plot would violate the assumption of independence while a trend in the variability of the residuals –for instance a funnel shape– suggests heterogeneity of variances. Departures from the theoretical normal line on the right plot are symptoms of lack of normality.

## 4.4. Treatment means

Below are the sorted means for each Treatment using the Fisher’s Least Significant Difference method and the multiple comparisons method of Tukey, both at the 5% level. Letters indicate if there are significant differences.

### 4.3.1. LSD test

# If there are no missing values, get a copy for further use
mydata$trait.est <- mydata.trait
# If there are missing values, get estimates
mydata$trait.est <- mve.rcbd(trait, treatment, rep, mydata)[, trait.est]
# Run LSD test
agricolae::LSD.test(mydata$trait.est, mydata$treatment, at[3, 1], at[3, 3])$groups

 means groups
G3\_/\_10 30 a
G1\_/\_30 25 ab
G2\_/\_10 20 b
G2\_/\_30 20 b
G1\_/\_10 10 c

### 4.3.2. Tukey test

agricolae::HSD.test(mydata$trait.est, mydata$treatment, at[3, 1], at[3, 3])$groups

 means groups
G3\_/\_10 30 a
G1\_/\_30 25 a
G2\_/\_10 20 ab
G2\_/\_30 20 ab
G1\_/\_10 10 b

### 4.3.3. Plot of means

## 4.5. Variance components

Below are the variance components for this model, under the assumption that Treatments and blocks are random. Here the model is fitted using REML and missing values are not estimated.

model <- lme4::lmer(trait ~ (1|treatment) + (1|rep), data = mydata)
vc <- data.frame(lme4::VarCorr(model))
vc[, c(1, 4, 5)]

 Variance Std.Dev.
treatment 5.000082e+01 7.07112579
block 7.787455e-04 0.02790601
Residual 9.999163e+00 3.16214533