Automatic report for a Completely Randomized Design (CRD)

International Potato Center

December 17, 2020

Table of Contents

# 1. Model specification and data description

Data from 85 genotypes have been evaluated using a completely randomized design. The statistical model is

where

* is the observed response with genotype and replication .
* is the mean response over all genotypes and replications.
* is the effect for genotype .
* is the error term.

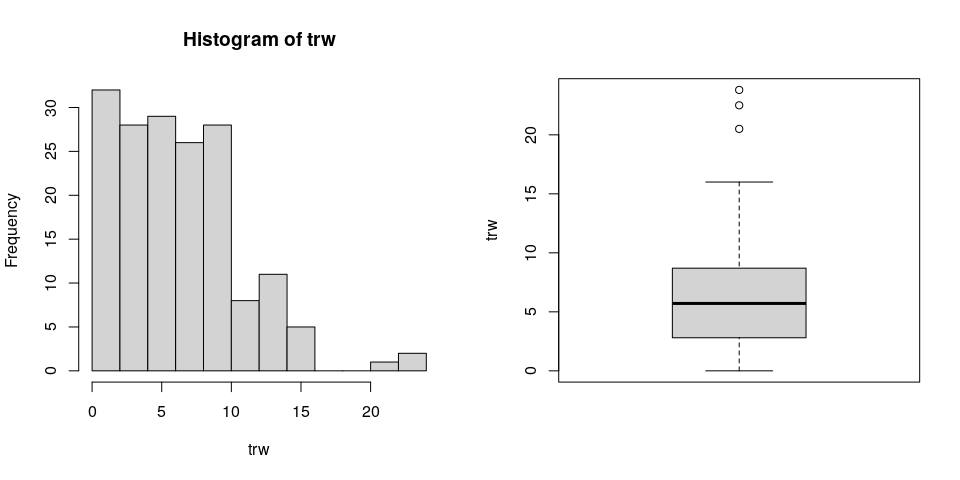
In this model we assume that the errors are independent and have a normal distribution with common variance, that is, .

# 2. Analysis for trait trw

## 2.1. Exploratory analysis

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

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



## 2.2. ANOVA

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

model <- aov(trait ~ treatment, data = mydata)  
# Anova table  
at <- anova(model)  
at

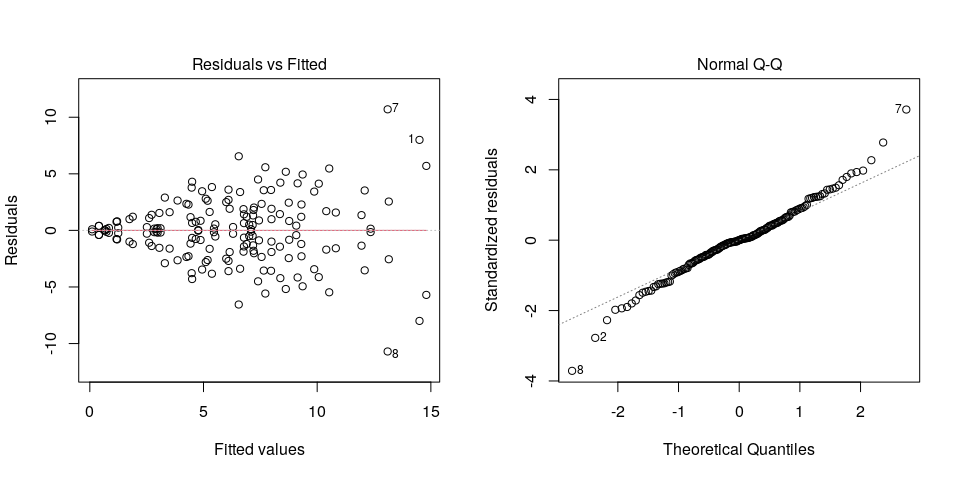
Analysis of Variance Table  
  
Response: "trw"  
 Df Sum Sq Mean Sq F value Pr(>F)   
geno 84 1962.2 23.360 1.4066 0.05932 .  
Residuals 85 1411.6 16.608   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

The coefficient of variation for this experiment is 65.39%. The p-value for genotypes is 0.05932 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 genotypes. The following residuals 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. Deviation from the theoretical normal line on the right plot is a sign of lack of normality.

## 2.4. Genotype means

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

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

Beauregard Blesbok Brondal Cemsa Huambachero INA-100   
 14.500 4.430 6.610 13.100 6.900 1.745   
 Jewel Jonathan Kemb-27 Mohc Naveto Necsu1560   
 3.860 0.400 6.760 0.400 0.100 6.100   
 Ningshu1 PJ05.012 PJ05.018 PJ05.052 PJ05.064 PJ05.091   
 7.970 6.550 7.990 12.090 3.510 14.800   
 PJ05.108 PJ05.109 PJ05.114 PJ05.120 PJ05.124 PJ05.130   
 9.140 3.060 7.010 5.270 13.150 7.400   
 PJ05.171 PJ05.172 PJ05.180 PJ05.212 PJ05.213 PJ05.217   
 6.150 2.720 4.660 10.530 4.500 2.930   
 PJ05.219 PJ05.220 PJ05.224 PJ05.227 PJ05.233 PJ05.235   
 7.440 3.300 3.110 8.380 12.350 7.720   
 PJ05.236 PJ05.238 PJ05.239 PJ05.240 PJ05.243 PJ05.245   
 8.620 7.555 10.820 7.200 9.360 5.100   
 PJ05.247 PJ05.248 PJ05.251 PJ05.253 PJ05.254 PJ05.257   
 0.670 4.480 7.160 8.360 5.520 11.950   
 PJ05.258 PJ05.302 PJ05.303 PJ05.304 PJ05.306 PJ05.324   
 6.100 5.180 9.870 1.250 6.000 4.940   
 PJ05.347 PZ06.011 PZ06.026 PZ06.027 PZ06.048 PZ06.053   
 7.090 4.500 4.340 2.980 9.080 8.750   
 PZ06.062 PZ06.070 PZ06.072 PZ06.091 PZ06.099 PZ06.103   
 1.890 0.730 9.310 7.220 10.400 8.770   
 PZ06.115 PZ06.120 PZ06.124 PZ06.196 PZ06.235 PZ06.304   
 2.840 6.780 1.175 0.840 7.990 2.610   
 PZ06.348 PZ06.349 PZ06.353 PZ06.359 PZ06.385 PZ06.441   
 7.180 9.300 6.780 4.870 5.460 4.250   
 Resisto Santo\_Amaro Tanzania Xushu18 Yanshu1 Yurimaguas   
 7.650 6.300 1.200 5.370 10.070 2.510   
 Zapallo   
 4.770

## 2.5. Variance components

Below are the variance components for this model, under the assumption that genotypes are random. Here the model is fitted using REML.

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

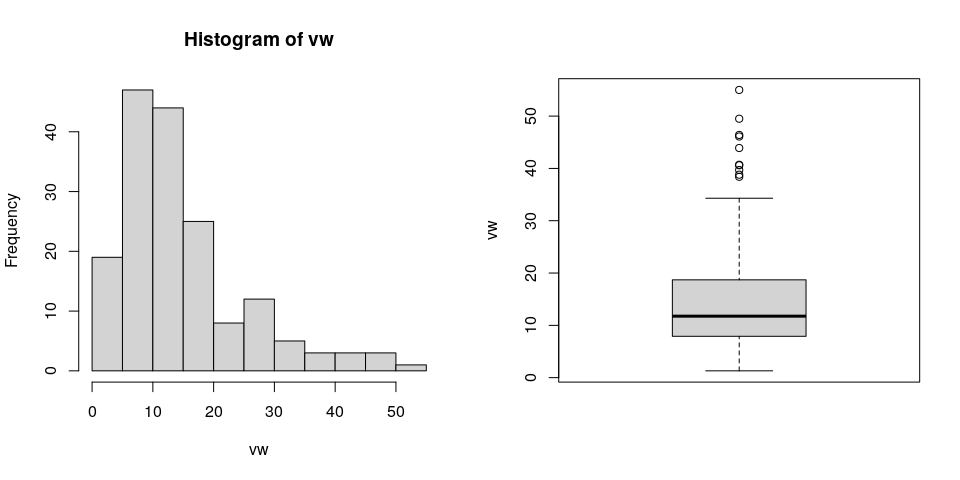
Variance Std.Dev.  
geno 3.376055 1.837404  
Residual 16.607569 4.075239

# 3. Analysis for trait vw

## 3.1. Exploratory analysis

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

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



## 3.2. ANOVA

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

model <- aov(trait ~ treatment, data = mydata)  
# Anova table  
at <- anova(model)  
at

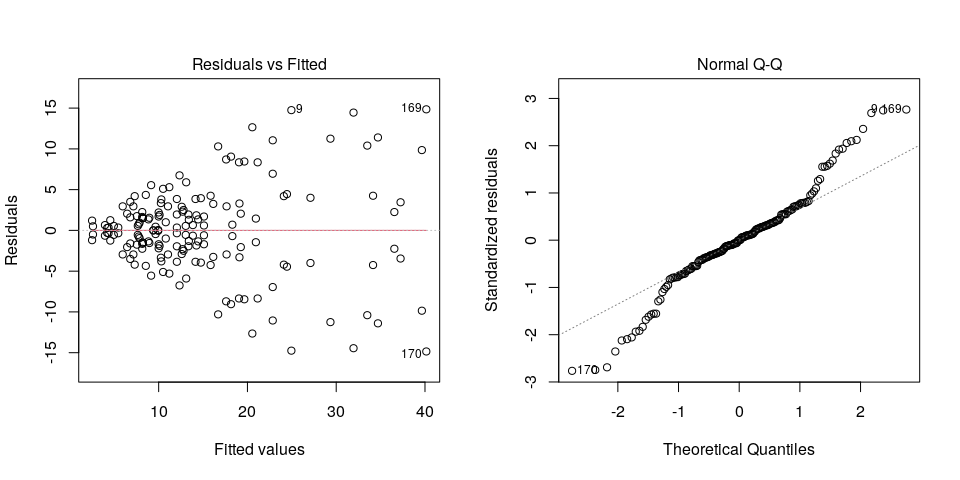
Analysis of Variance Table  
  
Response: "vw"  
 Df Sum Sq Mean Sq F value Pr(>F)   
geno 84 13610.1 162.024 2.8087 1.756e-06 \*\*\*  
Residuals 85 4903.4 57.687   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

The coefficient of variation for this experiment is 51.24%. The p-value for genotypes is 1.756e-06 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 genotypes. The following residuals 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. Deviation from the theoretical normal line on the right plot is a sign of lack of normality.

## 3.4. Genotype means

Below are the sorted means for each genotype 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

agricolae::LSD.test(mydata$trait, mydata$treatment, at[2, 1], at[2, 3])$groups

means groups  
Zapallo 40.15 a  
Xushu18 39.65 ab  
Tanzania 37.25 abc  
PZ06.124 36.55 abc  
PZ06.011 34.70 abcd  
INA-100 34.15 abcde  
Jonathan 33.50 abcdef  
Naveto 31.95 abcdefg  
Cemsa 29.35 abcdefgh  
PZ06.072 27.10 abcdefghi  
Huambachero 24.95 bcdefghij  
PZ06.196 24.45 cdefghijk  
Santo\_Amaro 24.10 cdefghijkl  
Jewel 22.85 cdefghijklm  
PJ05.257 22.85 cdefghijklm  
Necsu1560 21.15 defghijklmn  
PZ06.304 20.95 defghijklmno  
PJ05.108 20.55 defghijklmno  
Mohc 19.65 defghijklmnop  
PJ05.235 19.25 efghijklmnopq  
PZ06.062 19.10 efghijklmnopq  
PZ06.026 19.05 fghijklmnopq  
PZ06.027 18.30 ghijklmnopqr  
Kemb-27 18.15 ghijklmnopqr  
PZ06.115 17.65 ghijklmnopqrs  
PJ05.120 17.60 ghijklmnopqrs  
Beauregard 16.70 hijklmnopqrst  
Brondal 16.15 hijklmnopqrst  
PJ05.224 15.85 hijklmnopqrst  
PZ06.120 15.10 hijklmnopqrst  
PZ06.235 15.10 hijklmnopqrst  
PZ06.359 14.75 hijklmnopqrst  
Resisto 14.45 hijklmnopqrst  
PZ06.353 14.25 ijklmnopqrst  
PZ06.099 14.15 ijklmnopqrst  
PZ06.385 13.85 ijklmnopqrst  
PJ05.258 13.40 ijklmnopqrst  
PJ05.254 13.35 ijklmnopqrst  
Yanshu1 13.10 ijklmnopqrst  
PJ05.251 13.05 ijklmnopqrst  
PZ06.070 12.80 ijklmnopqrst  
PJ05.109 12.70 ijklmnopqrst  
Blesbok 12.60 ijklmnopqrst  
PJ05.212 12.35 ijklmnopqrst  
PJ05.213 12.05 ijklmnopqrst  
PZ06.349 12.05 ijklmnopqrst  
PZ06.441 12.05 ijklmnopqrst  
PJ05.247 11.20 jklmnopqrst  
PJ05.130 11.05 jklmnopqrst  
PJ05.227 10.80 jklmnopqrst  
PJ05.217 10.50 jklmnopqrst  
Ningshu1 10.30 jklmnopqrst  
PJ05.012 10.25 jklmnopqrst  
PJ05.091 10.10 jklmnopqrst  
PJ05.172 10.00 jklmnopqrst  
PJ05.324 10.00 jklmnopqrst  
PJ05.253 9.95 jklmnopqrst  
PZ06.053 9.65 klmnopqrst  
PZ06.091 9.40 klmnopqrst  
PJ05.248 9.15 lmnopqrst  
PJ05.240 8.95 mnopqrst  
PZ06.103 8.85 mnopqrst  
PJ05.052 8.55 mnopqrst  
PJ05.219 8.20 mnopqrst  
PJ05.304 8.20 mnopqrst  
PJ05.238 8.15 mnopqrst  
PZ06.048 8.15 mnopqrst  
PJ05.124 7.85 mnopqrst  
PJ05.347 7.75 nopqrst  
PJ05.114 7.65 nopqrst  
PZ06.348 7.55 nopqrst  
Yurimaguas 7.30 nopqrst  
PJ05.243 7.15 nopqrst  
PJ05.236 6.80 nopqrst  
PJ05.245 6.80 nopqrst  
PJ05.180 6.45 nopqrst  
PJ05.220 5.95 opqrst  
PJ05.233 5.45 pqrst  
PJ05.171 4.95 pqrst  
PJ05.303 4.55 qrst  
PJ05.018 4.35 qrst  
PJ05.239 4.20 qrst  
PJ05.302 3.95 rst  
PJ05.306 2.60 st  
PJ05.064 2.50 t

### 3.3.2. Tukey test

agricolae::HSD.test(mydata$trait, mydata$treatment, at[2, 1], at[2, 3])$groups

means groups  
Zapallo 40.15 a  
Xushu18 39.65 ab  
Tanzania 37.25 abc  
PZ06.124 36.55 abc  
PZ06.011 34.70 abcd  
INA-100 34.15 abcd  
Jonathan 33.50 abcd  
Naveto 31.95 abcd  
Cemsa 29.35 abcd  
PZ06.072 27.10 abcd  
Huambachero 24.95 abcd  
PZ06.196 24.45 abcd  
Santo\_Amaro 24.10 abcd  
Jewel 22.85 abcd  
PJ05.257 22.85 abcd  
Necsu1560 21.15 abcd  
PZ06.304 20.95 abcd  
PJ05.108 20.55 abcd  
Mohc 19.65 abcd  
PJ05.235 19.25 abcd  
PZ06.062 19.10 abcd  
PZ06.026 19.05 abcd  
PZ06.027 18.30 abcd  
Kemb-27 18.15 abcd  
PZ06.115 17.65 abcd  
PJ05.120 17.60 abcd  
Beauregard 16.70 abcd  
Brondal 16.15 abcd  
PJ05.224 15.85 abcd  
PZ06.120 15.10 abcd  
PZ06.235 15.10 abcd  
PZ06.359 14.75 abcd  
Resisto 14.45 abcd  
PZ06.353 14.25 abcd  
PZ06.099 14.15 abcd  
PZ06.385 13.85 abcd  
PJ05.258 13.40 abcd  
PJ05.254 13.35 abcd  
Yanshu1 13.10 abcd  
PJ05.251 13.05 abcd  
PZ06.070 12.80 abcd  
PJ05.109 12.70 abcd  
Blesbok 12.60 abcd  
PJ05.212 12.35 abcd  
PJ05.213 12.05 abcd  
PZ06.349 12.05 abcd  
PZ06.441 12.05 abcd  
PJ05.247 11.20 abcd  
PJ05.130 11.05 abcd  
PJ05.227 10.80 abcd  
PJ05.217 10.50 abcd  
Ningshu1 10.30 abcd  
PJ05.012 10.25 abcd  
PJ05.091 10.10 abcd  
PJ05.172 10.00 abcd  
PJ05.324 10.00 abcd  
PJ05.253 9.95 abcd  
PZ06.053 9.65 abcd  
PZ06.091 9.40 abcd  
PJ05.248 9.15 abcd  
PJ05.240 8.95 abcd  
PZ06.103 8.85 abcd  
PJ05.052 8.55 abcd  
PJ05.219 8.20 abcd  
PJ05.304 8.20 abcd  
PJ05.238 8.15 abcd  
PZ06.048 8.15 abcd  
PJ05.124 7.85 abcd  
PJ05.347 7.75 abcd  
PJ05.114 7.65 abcd  
PZ06.348 7.55 abcd  
Yurimaguas 7.30 abcd  
PJ05.243 7.15 abcd  
PJ05.236 6.80 abcd  
PJ05.245 6.80 abcd  
PJ05.180 6.45 bcd  
PJ05.220 5.95 cd  
PJ05.233 5.45 cd  
PJ05.171 4.95 cd  
PJ05.303 4.55 cd  
PJ05.018 4.35 cd  
PJ05.239 4.20 cd  
PJ05.302 3.95 cd  
PJ05.306 2.60 d  
PJ05.064 2.50 d

## 3.5. Variance components

Below are the variance components for this model, under the assumption that genotypes are random. Here the model is fitted using REML.

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

Variance Std.Dev.  
geno 52.16888 7.222803  
Residual 57.68671 7.595176