6510 Project ANOVA Models

Alisa Dmitrieva

2025-11-24

Split-Split-Plot ANOVA

The experiment was conducted using three blocks to control for field variability. Treatments were randomized within each block following a split–split plot design.

The structure of the experimental design was as follows:

- Main plot factor: Planting date (5 levels: July 16, August 1, August 16, September 1, September 16)
- Subplot factor: Rice genotype (3 varieties: Nehara, Bhasamanik, Bhasakalma)
- Sub-subplot factors: Plant spacing (6 in, 9 in, 12 in) and number of seedlings per hill (1, 2, and local method)

It was determined that a split-split-plot ANOVA model should be fitted to the data, as a standard ANOVA model might not be adequate. The *lmer* function from the *lme4* package will be used to fit the model.

```
# Loading the lme4 package
library(lmerTest)
```

```
## Loading required package: lme4
```

```
## Loading required package: Matrix
```

```
##
## Attaching package: 'lmerTest'
```

```
## The following object is masked from 'package:lme4':
##
##     lmer
```

```
## The following object is masked from 'package:stats':
##
##     step
```

```
## Fitting the mixed-effects model ##
  # Fixed effects include all factors and their interactions.
  # Random effects capture the appropriate
    # error structures for each plot size.
splitSplitAnova_fit <- lmer(yield ~ date * variety * seed * spacing +
```

```
                            (1 | block) +
                            (1 | block:date) +
                            (1 | block:date:variety),
                          data = df)
anova(splitSplitAnova_fit)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##                           Sum Sq Mean Sq NumDF DenDF  F value    Pr(>F)
## date                      426381  106595     4     8  78.4703 1.870e-06 ***
## variety                   326478  163239     2    20 120.1686 7.161e-12 ***
## seed                       65375   32687     2   240  24.0629 2.989e-10 ***
## spacing                   129705   64852     2   240  47.7413 < 2.2e-16 ***
## date:variety               54569    6821     8    20   5.0214 0.0016073 **
## date:seed                  19883    2485     8   240   1.8296 0.0722917 .
## variety:seed                4449    1112     4   240   0.8189 0.5141991
## date:spacing               42571    5321     8   240   3.9174 0.0002276 ***
## variety:spacing             5213    1303     4   240   0.9593 0.4305374
## seed:spacing                5574    1394     4   240   1.0259 0.3945266
## date:variety:seed          63885    3993    16   240   2.9393 0.0001857 ***
## date:variety:spacing       47251    2953    16   240   2.1740 0.0065104 **
## date:seed:spacing          23059    1441    16   240   1.0609 0.3937950
## variety:seed:spacing        9584    1198     8   240   0.8819 0.5323336
## date:variety:seed:spacing  30288     946    32   240   0.6968 0.8896726
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Updating the model formula to remove the highest-order
#   non-significant interaction term:
#     - date:variety:seed:spacing
splitSplitAnova2 <- update(splitSplitAnova_fit, .~.
                -date:variety:seed:spacing)
anova(splitSplitAnova2)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##                       Sum Sq Mean Sq NumDF DenDF  F value    Pr(>F)
## date                  411169  102792     4     8  78.4702 1.870e-06 ***
## variety               314831  157416     2    20 120.1688 7.160e-12 ***
## seed                   65375   32687     2   272  24.9531 1.122e-10 ***
## spacing               129705   64852     2   272  49.5074 < 2.2e-16 ***
## date:variety           52623    6578     8    20   5.0214 0.0016073 **
## date:seed              19883    2485     8   272   1.8973 0.0604901 .
## variety:seed            4449    1112     4   272   0.8492 0.4951088
## date:spacing           42571    5321     8   272   4.0623 0.0001386 ***
## variety:spacing         5213    1303     4   272   0.9948 0.4107872
## seed:spacing            5574    1394     4   272   1.0638 0.3747701
## date:variety:seed      63885    3993    16   272   3.0480 9.771e-05 ***
## date:variety:spacing   47251    2953    16   272   2.2544 0.0043273 **
## date:seed:spacing      23059    1441    16   272   1.1002 0.3545443
## variety:seed:spacing    9584    1198     8   272   0.9145 0.5047689
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# Updating the model formula to remove non-significant
#   3-way interaction terms:
#     - variety:seed:spacing
#     - date:seed:spacing
splitSplitAnova3 <- update(splitSplitAnova2, .~.
                -variety:seed:spacing
                -date:seed:spacing)
anova(splitSplitAnova3)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##                      Sum Sq Mean Sq NumDF DenDF  F value    Pr(>F)
## date                 412448  103112     4     8  78.4706 1.870e-06 ***
## variety              315808  157904     2    20 120.1688 7.160e-12 ***
## seed                  65375   32687     2   296  24.8759 1.033e-10 ***
## spacing              129705   64852     2   296  49.3542 < 2.2e-16 ***
## date:variety          52786    6598     8    20   5.0214 0.0016073 **
## date:seed             19883    2485     8   296   1.8914 0.0610042 .
## variety:seed           4449    1112     4   296   0.8465 0.4966250
## date:spacing          42571    5321     8   296   4.0497 0.0001375 ***
## variety:spacing        5213    1303     4   296   0.9917 0.4122963
## seed:spacing           5574    1394     4   296   1.0605 0.3762520
## date:variety:seed     63885    3993    16   296   3.0386 9.529e-05 ***
## date:variety:spacing  47251    2953    16   296   2.2474 0.0043292 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# Updating the model formula to remove non-significant
#   2-way interaction terms:
#     - seed:spacing
#     - variety:spacing
#     - variety:seed
splitSplitAnova4 <- update(splitSplitAnova3, .~.
                -seed:spacing
                -variety:spacing
                -variety:seed)
anova(splitSplitAnova4)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##                      Sum Sq Mean Sq NumDF  DenDF  F value    Pr(>F)
## date                 427283  106821     4  10.03  81.2276 1.339e-07 ***
## variety              316064  158032     2  20.00 120.1690 7.160e-12 ***
## seed                  65375   32687     2 300.00  24.8559 1.027e-10 ***
## spacing              129705   64852     2 300.00  49.3144 < 2.2e-16 ***
## date:variety          33277    4160     8  68.00   3.1630 0.0041865 **
## date:seed             18579    2322     8 300.00   1.7660 0.0832838 .
## date:spacing          31352    3919     8 300.00   2.9800 0.0031555 **
## date:variety:seed     68334    3417    20 300.00   2.5981 0.0002758 ***
## date:variety:spacing  52463    2623    20 300.00   1.9947 0.0075579 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The final reduced model includes the following terms:

- date
- variety
- seed
- spacing
- date:variety
- date:seed
- date:spacing
- date:variety:seed
- date:variety:spacing

Note that the p-value associated with the date:seed interaction is given by

`0.0832838.`

Despite being greater than $\alpha = 0.05$, I decided to keep this term in the model as it was very close.

```
# The final reduced split-split-plot ANOVA model
splitSplitAnova_reduced <- splitSplitAnova4
anova(splitSplitAnova_reduced)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##                       Sum Sq Mean Sq NumDF  DenDF  F value    Pr(>F)
## date                  427283  106821     4  10.03  81.2276 1.339e-07 ***
## variety               316064  158032     2  20.00 120.1690 7.160e-12 ***
## seed                   65375   32687     2 300.00  24.8559 1.027e-10 ***
## spacing               129705   64852     2 300.00  49.3144 < 2.2e-16 ***
## date:variety           33277    4160     8  68.00   3.1630 0.0041865 **
## date:seed              18579    2322     8 300.00   1.7660 0.0832838 .
## date:spacing           31352    3919     8 300.00   2.9800 0.0031555 **
## date:variety:seed      68334    3417    20 300.00   2.5981 0.0002758 ***
## date:variety:spacing   52463    2623    20 300.00   1.9947 0.0075579 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Standard ANOVA

A standard ANOVA model for a factorial experimental design with blocking will also be fitted to the data as an alternative strategy. In this case, the blocking factor will be treated as an independent nuisance factor.

A key assumption for this model is that there is no interaction between the blocks and the treatment factors.

```
# Fitting a standard factorial ANOVA model with blocking
factorialAnova_fit <- aov(yield ~ date * variety * seed * spacing + block, data=df)
summary(factorialAnova_fit)
```

```
##                              Df  Sum Sq Mean Sq F value   Pr(>F)
## date                          4 9559048 2389762 931.473  < 2e-16 ***
## variety                       2 1417021  708510 276.161  < 2e-16 ***
## seed                          2   65375   32687  12.741 5.18e-06 ***
## spacing                       2  129705   64852  25.278 8.77e-11 ***
## block                         2  288096  144048  56.147  < 2e-16 ***
## date:variety                  8  236848   29606  11.540 4.47e-14 ***
## date:seed                     8   19883    2485   0.969   0.4608
## variety:seed                  4    4449    1112   0.434   0.7843
## date:spacing                  8   42571    5321   2.074   0.0386 *
## variety:spacing               4    5213    1303   0.508   0.7299
## seed:spacing                  4    5574    1394   0.543   0.7042
## date:variety:seed            16   63885    3993   1.556   0.0806 .
## date:variety:spacing         16   47251    2953   1.151   0.3082
## date:seed:spacing            16   23059    1441   0.562   0.9105
## variety:seed:spacing          8    9584    1198   0.467   0.8789
## date:variety:seed:spacing    32   30288     946   0.369   0.9994
## Residuals                   268  687574    2566
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# Updating the model formula to remove the highest-order
#   non-significant interaction term:
#     - date:variety:seed:spacing
factorialAnova2 <- update(factorialAnova_fit, .~.
                -date:variety:seed:spacing)
summary(factorialAnova2)
```

```
##                          Df  Sum Sq Mean Sq F value   Pr(>F)
## date                      4 9559048 2389762 998.701  < 2e-16 ***
## variety                   2 1417021  708510 296.092  < 2e-16 ***
## seed                      2   65375   32687  13.660 2.10e-06 ***
## spacing                   2  129705   64852  27.102 1.51e-11 ***
## block                     2  288096  144048  60.199  < 2e-16 ***
## date:variety              8  236848   29606  12.373 2.53e-15 ***
## date:seed                 8   19883    2485   1.039   0.4068
## variety:seed              4    4449    1112   0.465   0.7615
## date:spacing              8   42571    5321   2.224   0.0257 *
## variety:spacing           4    5213    1303   0.545   0.7031
## seed:spacing              4    5574    1394   0.582   0.6756
## date:variety:seed        16   63885    3993   1.669   0.0517 .
## date:variety:spacing     16   47251    2953   1.234   0.2404
## date:seed:spacing        16   23059    1441   0.602   0.8816
## variety:seed:spacing      8    9584    1198   0.501   0.8555
## Residuals               300  717861    2393
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# Updating the model formula to remove all non-significant
#   3-way interaction terms:
#     - variety:seed:spacing
#     - date:seed:spacing
#     - date:variety:spacing
```

```r
factorialAnova3 <- update(factorialAnova2, .~.
                 -variety:seed:spacing
                 -date:seed:spacing
                 -date:variety:spacing)
summary(factorialAnova3)
```

```
##                     Df  Sum Sq Mean Sq  F value   Pr(>F)
## date                 4 9559048 2389762 1018.507  < 2e-16 ***
## variety              2 1417021  708510  301.964  < 2e-16 ***
## seed                 2   65375   32687   13.931 1.53e-06 ***
## spacing              2  129705   64852   27.640 7.54e-12 ***
## block                2  288096  144048   61.393  < 2e-16 ***
## date:variety         8  236848   29606   12.618 6.91e-16 ***
## date:seed            8   19883    2485    1.059   0.3914
## variety:seed         4    4449    1112    0.474   0.7548
## date:spacing         8   42571    5321    2.268   0.0225 *
## variety:spacing      4    5213    1303    0.555   0.6952
## seed:spacing         4    5574    1394    0.594   0.6673
## date:variety:seed   16   63885    3993    1.702   0.0446 *
## Residuals          340  797755    2346
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# Updating the model formula to remove all non-significant
#   2-way interaction terms:
#     - seed:spacing
#     - variety:spacing
#     - variety:seed
#     - date:seed
factorialAnova4 <- update(factorialAnova3, .~.
                 -seed:spacing
                 -variety:spacing
                 -variety:seed
                 -date:seed)
summary(factorialAnova4)
```

```
##                     Df  Sum Sq Mean Sq  F value   Pr(>F)
## date                 4 9559048 2389762 1028.564  < 2e-16 ***
## variety              2 1417021  708510  304.946  < 2e-16 ***
## seed                 2   65375   32687   14.069 1.33e-06 ***
## spacing              2  129705   64852   27.913 5.72e-12 ***
## block                2  288096  144048   61.999  < 2e-16 ***
## date:variety         8  236848   29606   12.743 4.32e-16 ***
## date:spacing         8   42571    5321    2.290   0.0212 *
## date:variety:seed   28   88217    3151    1.356   0.1111
## Residuals          348  808542    2323
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The final reduced model includes the following terms:

- date

- variety
- seed
- spacing
- date:variety
- date:spacing
- date:variety:seed

Note that the p-value associated with the date:variety:seed interaction is given by

`0.1111.`

However, in the previous reduced model, the p-value associated with the date:variety:seed interaction was less than $\alpha = 0.05$. Therefore, I decided to keep this term in the final model out of an abundance of caution.

```
# The final reduced factorial ANOVA model with blocking
factorialAnova_reduced <- factorialAnova4
summary(factorialAnova_reduced)
```

```
##                     Df  Sum Sq Mean Sq  F value   Pr(>F)
## date                 4 9559048 2389762 1028.564  < 2e-16 ***
## variety              2 1417021  708510  304.946  < 2e-16 ***
## seed                 2   65375   32687   14.069 1.33e-06 ***
## spacing              2  129705   64852   27.913 5.72e-12 ***
## block                2  288096  144048   61.999  < 2e-16 ***
## date:variety         8  236848   29606   12.743 4.32e-16 ***
## date:spacing         8   42571    5321    2.290   0.0212 *
## date:variety:seed   28   88217    3151    1.356   0.1111
## Residuals          348  808542    2323
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```