Comprehensive Statistical Analysis Report

ANCOVA • 3^3 Factorial CRD • Confounded 2■ Factorial in RBD

1. ANCOVA (Analysis of Covariance)

1.1 Purpose of ANCOVA

ANCOVA is used when we compare treatment means while adjusting for the effect of a covariate (x). In this dataset, potato varieties are categorical treatments and 'x' is a continuous concomitant variable.

1.2 Explanation of Code

Data Input and Factor Conversion

a$potatos = as.factor(a$potato) converts numerical potato codes (1–4) into categorical factor levels. ANCOVA requires treatments as factors.

ANOVA for Covariate x (Txx, Exx, Gxx)

Txx = Treatment sum of squares for x. Measures variation in x due to varieties.
Exx = Error sum of squares for x. Measures variation within varieties.
Gxx = Total SS = Txx + Exx.

ANOVA for Response y (Tyy, Eyy, Gyy)

Same logic as x: partition total variation in y into treatment and error components.

1.3 Regression Slopes

beta (adjusted slope) is obtained from lm(y ~ x + potato). This slope removes treatment effects.
beta2 (unadjusted slope) from lm(y ~ x) does not remove treatment variation.

1.4 Covariance Components

Exy = Exx * beta computes adjusted covariance under treatments.
Gxy = Gxx * beta2 computes total covariance without controlling for potato variety.
Txy = Gxy – Exy gives treatment-level covariance.

1.5 ANCOVA Table

ANCOVA: aov(y ~ x + potato) partitions variation due to x and then potato while adjusting for x.
This improves precision compared to ordinary ANOVA.

2. 3^3 Factorial Experiment in CRD

2.1 Purpose

A 3×3×3 factorial experiment studies three independent factors each with 3 levels: pressure, nozzle type, and speed. The goal is to estimate main effects and interactions.

2.2 Factor Conversion

mutate(across(!'yield', as.factor)) converts numeric treatment levels to factors. This is required because factorial ANOVA treats them as categorical groups.

2.3 Model

model_crd = aov(yield ~ nozzle.type * speed * pressure)
The * operator expands to main effects + all 2-way + 3-way interactions.

2.4 ANOVA Interpretation

Pressure and speed have large significant effects ($p < 0.001$).
Significant two-way interactions (speed×pressure, pressure×nozzle, speed×nozzle) indicate combined factor effects.
Three-way interaction is not significant.

2.5 TukeyHSD

Used for pairwise comparisons between treatment levels after significant ANOVA. Free from Type I error inflation unlike LSD.
Visual plots indicate which factor levels differ statistically.

2.6 Mean Plot

Plots of mean yield by speed show how response changes across factor levels. Used for interpretation and visual checking of factor effects.

3. Confounded 2■ Factorial in RBD

3.1 Purpose

Large factorials like 2■ require many plots. Confounding assigns one high-order interaction to blocks so that block variation does not affect important main effects.

3.2 Factor Conversion

mutate(across(!'yield', as.factor)) ensures block, n, p, k, d are treated as categorical.

3.3 Full Model

lm(yield ~ block + n + p + k + d + interactions...) fits:

- all 4 main effects
- important two-way interactions
- all three-way interactions
- the four-way interaction

3.4 Singularities (NA Coefficients)

Because confounding aliases certain interactions with blocks, some model terms become linearly dependent. R returns NA because those interactions cannot be estimated separately.

3.5 ANOVA Interpretation

Significant effects include:

- n (very strong)
- k
- p×d interaction
- n×k×d interaction

Block effect is not significant, showing successful blocking.

3.6 Final Notes on Confounding

Confounding helps reduce experimental size by sacrificing unimportant high-order interactions. The ANOVA shows main effects remain estimable while certain interactions are aliased.

Overall Summary

• ANCOVA increases precision by adjusting for covariate x.
• 3^3 CRD detects strong main effects of speed and pressure, plus moderate interactions.
• Confounded 2■ RBD successfully estimates main effects while reducing experiment size.

These analyses demonstrate strong experimental design understanding and proper statistical application.