05 ANOVA

One and Two-way Analysis of Variance

Anders Stockmarr

Course developers: Anders Stockmarr, Elisabeth Wreford Andersen

DTU Department of Applied Mathematics and Computer Science Section for Statistics and Data Analysis Technical University of Denmark anst@dtu.dk

January 8th, 2025



Department of Applied Mathematics and Computer Science



Monday Statistical inference, and the t-test
Tuesday Simple and Multiple regression
Wednesday ANOVA, ANCOVA, and linear models
Thursday Categorical data, statistical report writing,
logistic Regression
Friday Introduction to repeated measures, Principal

Component Analysis

Plan for today

05_ANOVA

Lecture on one-way analysis of variance

Exercise

Recap of exercise and discussion

Lecture on two-way analysis of variance and the general linear model

Exercise

Recap of exercises and discussion

Outline, 1 and 2 Way ANQYANOVA

- 🕦 One-way Analysis of Variance
 - Descriptive
 - ANOVA Model
 - Data Example
 - Model Control
 - Multiple Testing
 - Exercise
- Two-way Analysis of Variance
 - Interaction
 - Interaction estimates
 - Hypotheses
 - Exercise

Overview

05_ANOVA

- Done-way Analysis of Variance
 - Descriptive
 - ANOVA Model
 - Data Example
 - Model Control
 - Multiple Testing
 - Exercise
- Two-way Analysis of Variance
 - Interaction
 - Interaction estimates
 - Hypotheses
 - Exercise

One-way Analysis of Variange ANOVA

- Comparing the means in more than two groups.
- One-way ANOVA gives joint test for any differences between the groups.
- One-way because only one grouping.
- Analysis of variance because we are comparing the variance between groups to the variance within groups.

Example: Birth Weight (Hosmer, Lemeshow data 2000)

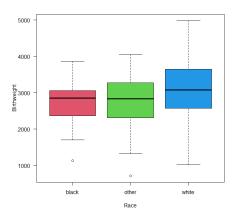
- * Birth weight of 189 babies.
- * Race is grouped in three groups: white, black and other.
- * We want to compare the birth weight for these three groups.
- * The first step is to describe the data using simple tables and plots.

Example: Birth Weight 05 ANOVA

```
lbw <- read.delim("lbw.txt")</pre>
by(lbw$bwt, lbw$race, summary)
## lbw$race: black
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 1140 2370 2850 2720 3060 3860
## lbw$race: other
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 709 2310 2840 2800 3270 4050
## lbw$race: white
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 1020 2580 3080 3100 3650
                                      4990
```

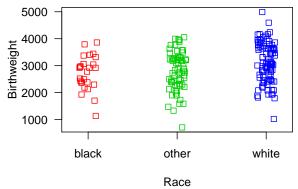
Example: Birth Weight, Boxplotva

```
boxplot(bwt ~ race, data = lbw, xlab = 'Race', ylab = 'Birthweight'
las = 1, col = 2:4)
```



Example: Birth Weight, Stripphorat

```
stripchart(bwt ~ race, data = lbw, vertical = TRUE, xlab = "Race",
ylab = "Birthweight", method = "jitter", las = 1, col = 2:4)
```



One-way Analysis of Varianse Madel

Let Y_{gi} be birth weight for child i in group $g, i \in \{1, \dots, n_g\}$ and $g \in \{1, \dots, k\}$.

We assume:

$$Y_{gi}$$
 = μ_g + ϵ_{gi}

Observation i in group g Mean in group g Individual variation

The observations are assumed to follow a normal distribution within each group, with a common variance σ^2

$$\epsilon_{gi} \sim N(0, \sigma^2), \quad Y_{gi} \sim N(\mu_g, \sigma^2)$$

These model assumptions need to be checked before interpreting the results.

Test of Hypothesis

05_ANOVA

In a one-way analysis of variance there is only one hypothesis: Are the group means the same?

$$H_0: \mu_1 = \mu_2 = \ldots = \mu_k$$

If the hypothesis is accepted it means that we can assume the simpler model:

$$Y_{gi} = \mu + \epsilon_{gi}, i \in \{1, \dots, n_g\}, g \in \{1, \dots, k\}$$

ANOVA math: Sums of Squaresva

Decomposition of deviation from grand mean:

$$y_{gi} - \bar{y}_{.} = (y_{gi} - \bar{y}_{g.}) + (\bar{y}_{g.} - \bar{y}_{.})$$

Decomposition of variation (sum of squares, SS).

$$\underbrace{\sum_{gi} (y_{gi} - \bar{y}.)^2}_{\text{Total variation}} = \underbrace{\sum_{gi} (y_{gi} - \bar{y}_{g.})^2}_{\text{Within groups}} + \underbrace{\sum_{gi} (\bar{y}_{g.} - \bar{y}.)^2}_{\text{Between groups}}$$

- y_{qi} i'th observation in group g.
- \bar{y}_g . Average in group g.
 - \bar{y} Overall average or "grand mean".

Decomposition of Variation of Nanova

Total = Between + Within

$$SS_{Total} = SS_{Between} + SS_{Within}$$

 $N-1 = (k-1) + (N-k)$

F-test statistic

$$F = \frac{MS_{Between}}{MS_{Within}} = \frac{SS_{Between}/(k-1)}{SS_{Within}/(N-k)}$$

Decomposition of Variation ANOVA

Total = Between + Within

$$SS_{Total} = SS_{Between} + SS_{Within}$$

 $N-1 = (k-1) + (N-k)$

F-test statistic

$$F = \frac{MS_{Between}}{MS_{Within}} = \frac{SS_{Between}/(k-1)}{SS_{Within}/(N-k)}$$

Hypothesis test

Reject null hypothesis if F large, i.e. if the variation *between* groups is too large compared to the variation *within* groups.

Decomposition of Variation 5_ANOVA

Total = Between + Within

$$SS_{Total} = SS_{Between} + SS_{Within}$$

 $N-1 = (k-1) + (N-k)$

F-test statistic

$$F = \frac{MS_{Between}}{MS_{Within}} = \frac{SS_{Between}/(k-1)}{SS_{Within}/(N-k)}$$

Hypothesis test

Reject null hypothesis if F large, i.e. if the variation *between* groups is too large compared to the variation *within* groups.

F-test statistic

The F-test statistic follows F-dist. with (k-1) and (N-k) df. Reject if F is large in F-distribution.

Test of Hypothesis in Examplenova

We want to test whether the birth weight of the children are the same irrespective of race. This corresponds to the hypothesis:

$$H_0: \mu_1 = \mu_2 = \mu_3$$

If the hypothesis is accepted it means that we can assume the simpler model:

$$Y_{qi} = \mu + \epsilon_{qi}, i \in \{1, \dots, n_q\}, g \in \{1, 2, 3\}$$

ANOVA IN R (Analysis of Mariance Table)

05_ANOVA

The test statistic F is calculated as:

$$F = \frac{Mean\ Sq\ race}{Mean\ Sq\ residual} = \frac{2524181}{510037} = 4.949$$

- Nominator (Mean Sq race) is the variation between races.
- Denominator (Mean Sq residual) is the residual variation (or within group).

We reject the null hypothesis of equal means if F is large, i.e. when large variation between groups compared to variation within groups.

05_ANOVA

The test statistic F is calculated as:

$$F = \frac{Mean\ Sq\ race}{Mean\ Sq\ residual} = \frac{2524181}{510037} = 4.949$$

- Nominator (Mean Sq race) is the variation between races.
- Denominator (Mean Sq residual) is the residual variation (or within group).

We reject the null hypothesis of equal means if F is large, i.e. when large variation between groups compared to variation within groups.

Here $F\sim F(2,186)$ and the test probability is found to be $p=P(F\geq 4.495)=0.008<0.05$ so the null hypothesis is rejected

ANOVA IN R (estimates) 05_ANOVA

```
summary(model1)
##
## Call:
## lm(formula = bwt ~ race, data = lbw)
##
## Coefficients:
##
             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 2719.7 140.1 19.42 <2e-16 ***
## raceother 84.3 165.0 0.51 0.610
## racewhite 383.3 157.9 2.43 0.016 *
##
## Residual standard error: 714 on 186 degrees of freedom
## Multiple R-squared: 0.0505, Adjusted R-squared: 0.0403
## F-statistic: 4.95 on 2 and 186 DF, p-value: 0.00805
```

ANOVA IN R (estimates) 05_ANOVA

```
confint(model1)

## 2.5 % 97.5 %

## (Intercept) 2443.38 2996.00

## raceother -241.22 409.86

## racewhite 71.83 694.81
```

ANOVA IN R (estimates) 05_ANOVA

```
confint(model1)

## 2.5 % 97.5 %

## (Intercept) 2443.38 2996.00

## raceother -241.22 409.86

## racewhite 71.83 694.81
```

- Intercept is the level of the reference group. R has taken the first group alphabetically or numerically. In our example raceblack.
- racewhite this is the difference in birth weight between white and black babies.
- raceother this is the difference in birth weight between other and black babies.

05_ANOVA

There are some assumptions connected to the model that should be checked:

05_ANOVA

There are some assumptions connected to the model that should be checked:

• Independent observations.

05_ANOVA

There are some assumptions connected to the model that should be checked:

- Independent observations.
- Variance homogeneity.

05_ANOVA

There are some assumptions connected to the model that should be checked:

- Independent observations.
- Variance homogeneity.
- Normally distributed observations.

Model Control 1: Independent Space Servations

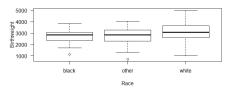
Based on knowledge about the sample.

- * No twins, siblings,...
- * Only one observation from each person.

Model Control 2: Same Variance in Groups

Also called variance homogeneity (just one σ^2). Can be checked:

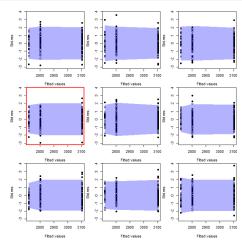
Boxplots



- Residuals versus fitted values, look for trumpet shape.
- Test the hypothesis of equal variances (Levene's test)

Model Control 2: Residual 65 ANOVA

library(MESS)
wallyplot(model1)



Model Control 2: Levene's of extorn

Test the hypothesis that

$$H_0: \sigma_1^2 = \sigma_2^2 = \sigma_3^2$$

```
library(car)
leveneTest(model1)

## Levene's Test for Homogeneity of Variance (center = median)
## Df F value Pr(>F)
## group 2 0.4666 0.6277
## 186
```

Model Control 2: Levene's Testova

Test the hypothesis that

$$H_0: \sigma_1^2 = \sigma_2^2 = \sigma_3^2$$

```
library(car)
leveneTest(model1)

## Levene's Test for Homogeneity of Variance (center = median)
## Df F value Pr(>F)
## group 2 0.4666 0.6277
## 186
```

As p = 0.6277 > 0.05 then H_0 is accepted.

Model Control 3: Normal Distribution

We have assumed that the observations follow a normal distribution within each group. This can be checked:

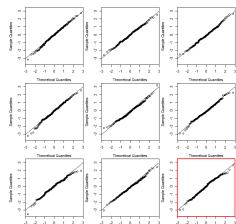
- Draw qq-plots for each group (only if many observations)
- Draw qq-plots for the residuals

$$r_{gi} = Y_{gi} - \hat{\mu}_g = Y_{gi} - \bar{Y}_g$$

- Not advisable to do formal tests for normality
 - If the data set is large then nearly always reject normality.
 - If the data set is small then nearly always accept.

Model Control 3: QQ-plots_{05_ANOVA}

```
qqwrap <- function(x, y, ...) {qqnorm(y,main="",...)
   abline(a=0, b=1)}
wallyplot(model1, FUN=qqwrap)</pre>
```



Conclusion for the Birthweight Data

- The model check showed that the assumptions were fulfilled.
- The F-test showed that there is a significant difference in birthweight for different races.
- But not where the differences were to be found.

Conclusion for the Birthweight Data

- The model check showed that the assumptions were fulfilled.
- The F-test showed that there is a significant difference in birthweight for different races.
- But not where the differences were to be found.

Pairwise Comparisons

Give problems with multiple testing. We have m=k(k-1)/2 possible tests so true level of significance (probability of rejecting a true H_0) $1-(1-\alpha)^m$. For k=3 we get 0.14.

05_ANOVA

05_ANOVA

No completely satisfactory solution to the multiple testing problem:

• Try to avoid the problem (focus the project).

05_ANOVA

- Try to avoid the problem (focus the project).
- Select a small number of comparisons in the planning stage.

05_ANOVA

- Try to avoid the problem (focus the project).
- Select a small number of comparisons in the planning stage.
- ullet Make a graph of the averages $\pm 2 imes SEM$ and judge visually.

05_ANOVA

- Try to avoid the problem (focus the project).
- Select a small number of comparisons in the planning stage.
- ullet Make a graph of the averages $\pm 2 imes SEM$ and judge visually.
- Modify the t-test.
 - Bonferroni test at level α/m .
 - Conservative (accepts in general too often).

Pairwise Tests in R, no correction

```
pairwise.t.test(lbw$bwt, lbw$race, p.adj = "none")
##
##
    Pairwise comparisons using t tests with pooled SD
##
  data: lbw$bwt and lbw$race
##
##
        black other
## other 0.6100 -
## white 0.0161 0.0093
##
## P value adjustment method: none
```

Pairwise Tests in R, bonferge Riova

Exercise: One-way ANOVA_{05_ANOVA}

• Exercise 1 Birth Weight and Mother's Weight.

Overview

05_ANOVA

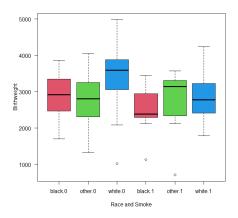
- One-way Analysis of Variance
 - Descriptive
 - ANOVA Model
 - Data Example
 - Model Control
 - Multiple Testing
 - Exercise
- Two-way Analysis of Variance
 - Interaction
 - Interaction estimates
 - Hypotheses
 - Exercise

Two-way analysis of variange ANOVA

- * Before we had one factor (race), now we have two (race and smoke).
- * We want to study how race and smoking during pregnancy are associated with birth weight.
- * Our outcome is still birth weight, factor 1: race (3 groups), factor 2: smoke (2 groups).

Example: Birth weight, Boxp Avova

```
boxplot(bwt ~ race*smoke, data=lbw, xlab = 'Race and Smoke',
ylab = 'Birthweight', las = 1, col = 2:4)
```



Two-way analysis of variance and ditive model

We have an effect of race (r) and smoke (s):

$$Y_{rsi} = \mu + \alpha_r + \beta_s + \epsilon_{rsi}, \quad \epsilon_{rsi} \sim N(0, \sigma^2)$$

and the effects are additive.

Here Y_{rsi} is the birth weight for baby i of race r and mother's smoking status s.

 α_r is the effect of race.

 eta_s is the effect of smoking.

Interaction

05_ANOVA

- * Or effect modification.
- * How do the factors work together?
- * Maybe the effect of smoking is not the same for all white, black and other.

Model with Interaction

05_ANOVA

$$Y_{rsi} = \mu + \alpha_r + \beta_s + \alpha \beta_{rs} + \epsilon_{rsi}, \quad \epsilon_{rsi} \sim N(0, \sigma^2)$$

Here Y_{rsi} is the birth weight for baby i of race r and mother's smoking status s.

lpha and eta are called main effects and lphaeta the interaction.

The effect of smoking can depend on race and the effect of race can depend on smoking status.

Model with Interaction

05_ANOVA

$$Y_{rsi} = \mu + \alpha_r + \beta_s + \alpha \beta_{rs} + \epsilon_{rsi}, \quad \epsilon_{rsi} \sim N(0, \sigma^2)$$

Here Y_{rsi} is the birth weight for baby i of race r and mother's smoking status s.

 α and β are called main effects and $\alpha\beta$ the interaction.

The effect of smoking can depend on race and the effect of race can depend on smoking status.

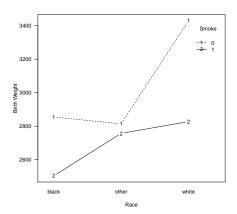
We would like to simplify the model to the additive $(\alpha \beta_{rs})$:

$$Y_{rsi} = \mu + \alpha_r + \beta_s + \epsilon_{rsi}$$

Here both smoking and race have an effect on birth weight. But the effect of smoking does not depend on race and vice versa.

Birth Weight Example with Notation

```
interaction.plot(lbw$race, lbw$smoke, lbw$bwt, type=c("b"),
ylab = "Birth Weight", las=1, trace.label = "Smoke", xlab = "Race")
```



Birth Weight Example with Intervaction

```
lbw$smoke <- as.factor(lbw$smoke)</pre>
model2 <- lm(bwt ~ race + smoke + race:smoke, data = lbw)
summary(model2)
##
## Coefficients:
##
                  Estimate Std. Error t value Pr(>|t|)
## (Intercept) 2854.50
                              170.84 16.708 <2e-16 ***
                              194.11 -0.207 0.83591
## raceother
              -40.26
                              199.50 2.878 0.00447 **
## racewhite 574.25
## smoke1
           -350.50 275.47 -1.272 0.20486
## raceother:smoke1 293.43
                              351.13 0.836 0.40443
## racewhite:smoke1 -250.87
                              309.00 -0.812 0.41792
##
## Residual standard error: 683.4 on 183 degrees of freedom
## Multiple R-squared: 0.1447, Adjusted R-squared: 0.1213
## F-statistic: 6.191 on 5 and 183 DF, p-value: 2.505e-05
```

Understanding the Estimates ANOVA

Interpretation of the estimates from the previous slide.

- Intercept = 2854.50
 The estimated birth weight for the reference group. Here race = "black"and smoke = 0.
- raceother = -40.26.
 The effect of Other vs. Black for the ref. group (i.e. non-smoker).
- racewhite = 574.25
 The effect of White vs. Black for the ref. group (i.e. non-smoker).
- smoke1 = -350.50
 The effect of smoker vs non-smoker for the ref. group (i.e. Black).

Understanding the Estimates (NOVAtd)

- * raceother:smoke1 = 293.43
 The extra effect of Other and smoker. The total effect of smoker vs nonsmoker for Other -350.50 + 293.43 = -57.07
- * racewhite:smoke1 = -250.87

 The extra effect of White and smoker. The total effect of smoker vs nonsmoker for White -350.50 250.87 = -601.37

Understanding the Estimates (NOVAtd)

The estimated birth weight for combinations of race and smoke.

| | | Race | |
|---------------|----------|-----------|-----------|
| ${\sf Smoke}$ | Black | Other | White |
| 0 | 2854.50 | 2854.50 | 2854.50 |
| | | - 40.26 | + 574.25 |
| | | = 2814.24 | = 3428.75 |
| 1 | 2854.50 | 2854.50 | 2854.50 |
| | - 350.50 | - 350.50 | - 350.50 |
| | | - 40.26 | + 574.25 |
| | | + 293.43 | - 250.87 |
| | = 2504 | = 2757.17 | = 2827.38 |

Hypotheses

05_ANOVA

Model without interaction:

$$M_2: \mu_{rsi} = \mu + \alpha_r + \beta_s$$

$$H_{20}: (\alpha\beta)_{rs} = 0 \text{ for all } r, s$$

$$H_{2A}: (\alpha\beta)_{rs} \neq 0$$

No effect of race:

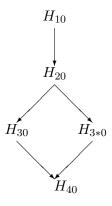
No effect of smoking:

$$M_3: \mu_{rsi} = \mu + \beta_s$$
 $M_{3*}: \mu_{rsi} = \mu + \alpha_r$ $H_{30}: \alpha_r = 0 \text{ for all } r$ $H_{3*0}: \beta_s = 0 \text{ for all } s$ $H_{3*1}: \beta_s \neq 0$

No effect of smoking or race:

$$M_{3*}: \mu_{rsi} = \mu$$

Hypotheses in Two-way Anglysisvef Variance



- If H_{10} or H_{20} is rejected then the analysis is finished.
- H_{30} and H_{3*0} are on the same level.
- If reject H_{30} then test H_{3*0} .

Hypotheses in Birth Weights Example

Note:

$$Mean\ Sq = \frac{Sum\ Sq}{Df}$$

Test for No Interaction, Bigthan Example

We will try to simplify the model by testing the hypothesis of no interaction:

$$F = \frac{Mean\ Sq\ race: smoke}{Mean\ Sq\ Residual} = \frac{1054321}{466993} = 2.26$$

Under the hypothesis of no interaction we have $F \sim F(2,183)$ and the test probability is found as:

$$p = P(F \ge f) = P(F \ge 2.26) = 0.107 > 0.05$$

Test for No Interaction, Birth Weight Example

We will try to simplify the model by testing the hypothesis of no interaction:

$$F = \frac{Mean\ Sq\ race: smoke}{Mean\ Sq\ Residual} = \frac{1054321}{466993} = 2.26$$

Under the hypothesis of no interaction we have $F \sim F(2,183)$ and the test probability is found as:

$$p = P(F \ge f) = P(F \ge 2.26) = 0.107 > 0.05$$

The hypothesis of no interaction is accepted.

Simplifying the model, Birth Weight Example

Simplifying the model, Birth Weight Example

Test for no effect of race H_{30} :

$$F = \frac{Mean\ Sq\ race}{Mean\ Sq\ Residual} = \frac{2524181}{473343} = 5.3327,\ p = 0.006 < 0.05$$

Simplifying the model, Birth Weight Example

```
model3 <- lm(bwt ~ race + smoke, data = lbw)
anova(model3)

## Analysis of Variance Table

##
## Response: bwt

## Df Sum Sq Mean Sq F value Pr(>F)

## race 2 5048361 2524181 5.3327 0.0056021 **

## smoke 1 7298537 7298537 15.4191 0.0001214 ***

## Residuals 185 87568401 473343
```

Test for no effect of race H_{30} :

$$F = \frac{Mean\ Sq\ race}{Mean\ Sq\ Residual} = \frac{2524181}{473343} = 5.3327,\ p = 0.006 < 0.05$$

Test for no effect of smoke H_{3*0} :

$$F = \frac{Mean\ Sq\ smoke}{Mean\ Sq\ Residual} = \frac{7298537}{473343} = 15.4191,\ p = 0.0001 < 0.05$$

Final model, Birth Weight Fxancule

```
confint(model3)
## 2.5 % 97.5 %
## (Intercept) 2605.5688 3163.0662
## raceother -320.3599 313.0773
## racewhite 148.5607 752.5194
## smoke1 -643.0746 -212.9761
```

Interpretation of estimates:

For two mothers of the same race where one is a smoker and the other nonsmoker, then the birth weight of the smoker's baby will be 428g less than for the nonsmoker. The difference could be as much as 643g or as little as 213g based on the 95% CI.

Exercise: Two-way ANOVA_{05_ANOVA}

• Exercise: Beer Tasting

Exercise: Fish