Introduction to Statistics C

Maria Anastasiadi

2025-04-25

Statistical Foundations Part3

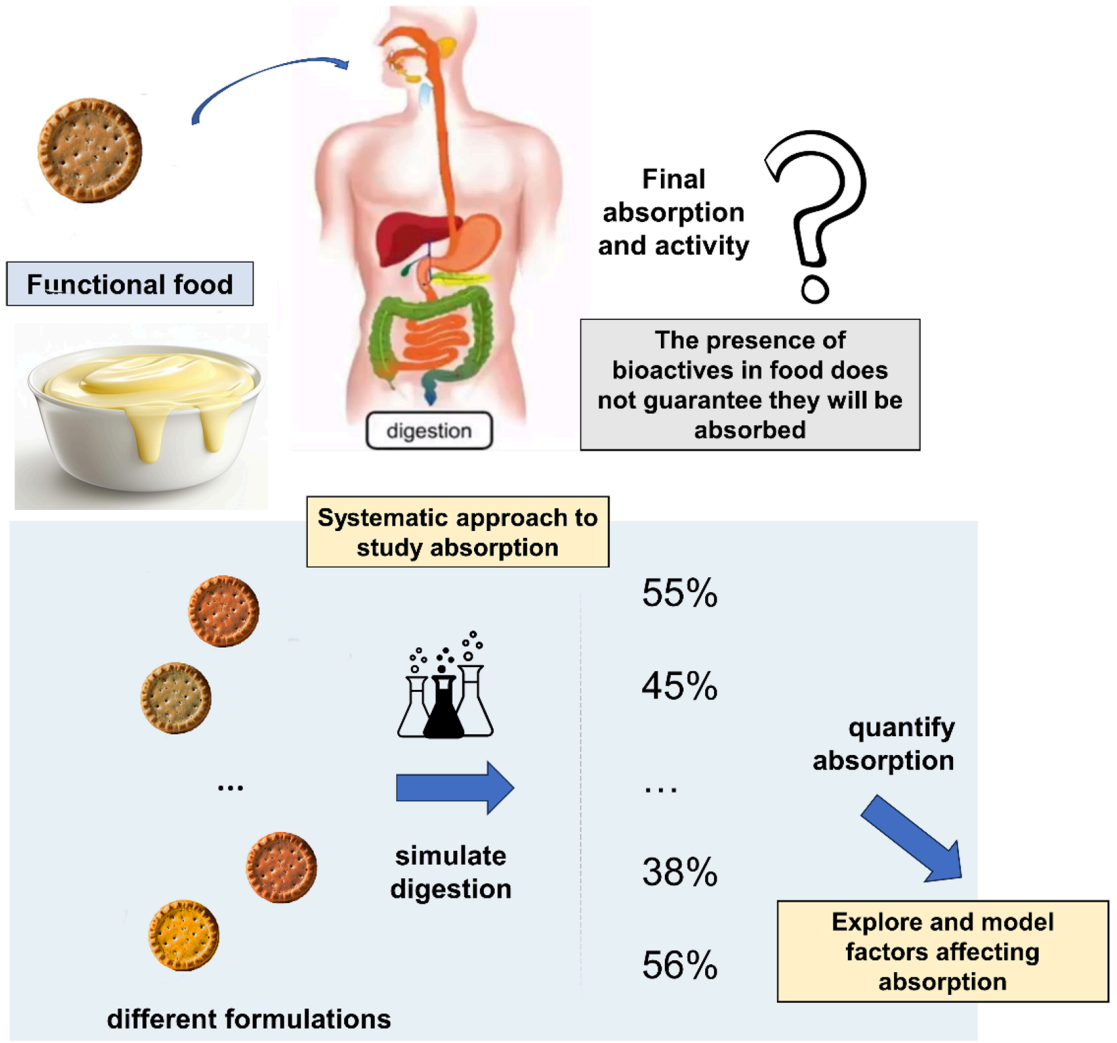
Other Types of ANOVA

Introduction

- The ANOVA examples we have considered so far in the course are one-way ANOVA since they involve only one factor, e.g. fibre type or wheat cultivar.

- If for the Bioaccessibility example we had considered two factors—e.g. fibre type AND food matrix type we would have to use something called a two-way ANOVA.

Bioaccesibility Experiment



Interaction effect

A two-way ANOVA on data from the Bioaccessibility experiment would tell us whether Bioaccessibility was affected by:

- Fibre Type –> **Main Effect**
- Food Type –> **Main Effect**
- Fibre/Food combination –> **The Interaction Effect**

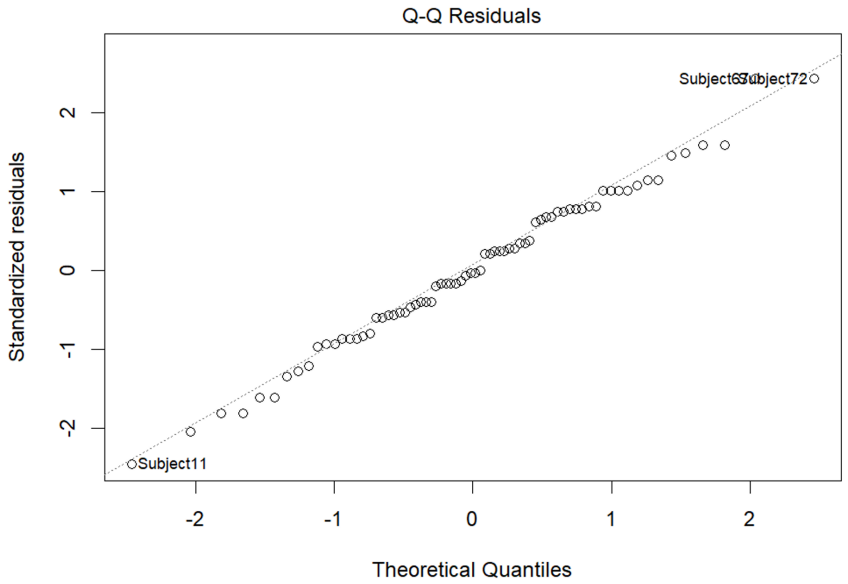
Interaction effect

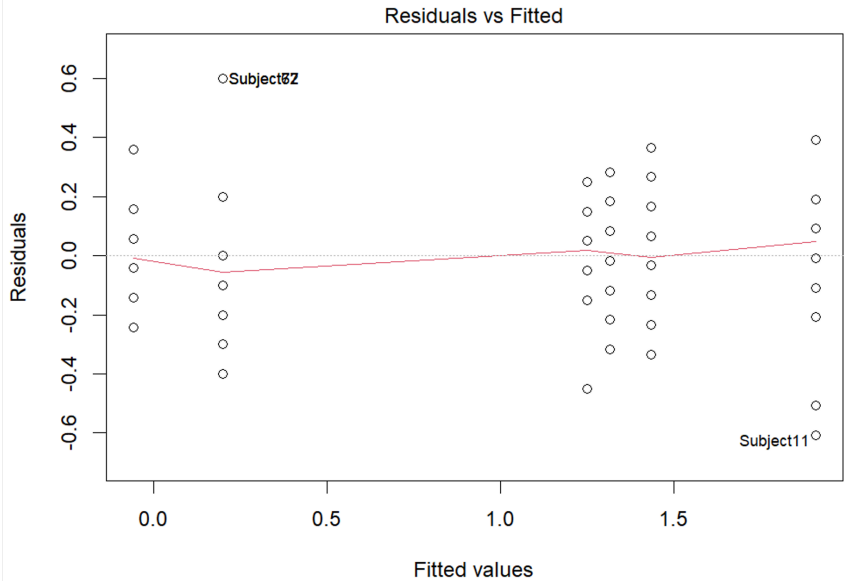
The interaction effect tells us whether the effect of Fibre depends on Food Type (and vice versa).

Two-Way ANOVA R syntax

```
bioaccess.anova<-aov(Bioaccessibility~Fibre*Food, data=bioaccess)
```

Two-way ANOVA Assumptions



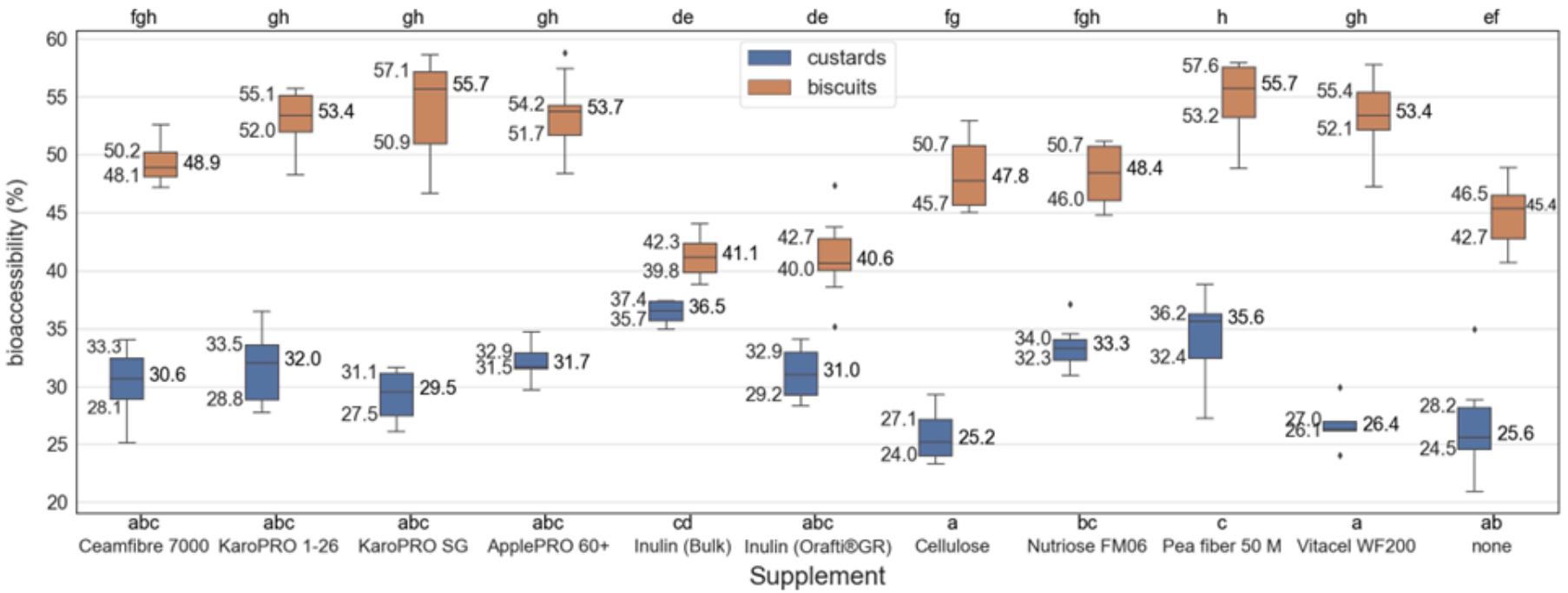


Post-Hoc Tests

If the interactions effect is significant we can do multiple comparison tests.

For example: `TukeyHSD(biaccess.anova)`

Results


fgh
gh
gh
gh
de
de
fg
fgh
h
gh
ef
custards
biscuits
bioaccessibility (%)
60
55
50
45
40
35
30
25
20
50.2
48.1
48.9
55.1
52.0
53.4
57.1
50.9
55.7
54.2
51.7
53.7
42.3
39.8
41.1
42.7
40.0
40.6
50.7
45.7
47.8
50.7
46.0
48.4
57.6
53.2
55.7
55.4
52.1
53.4
46.5
42.7
45.4
33.3
28.1
30.6
33.5
28.8
32.0
31.1
27.5
29.5
32.9
31.5
31.7
37.4
35.7
36.5
32.9
29.2
31.0
27.1
24.0
25.2
34.0
32.3
33.3
36.2
32.4
35.6
27.0
26.1
26.4
28.2
24.5
25.6
abc
Ceamfibre 7000
abc
KaroPRO 1-26
abc
KaroPRO SG
abc
ApplePRO 60+
cd
Inulin (Bulk)
abc
Inulin (Orafti®GR)
a
Cellulose
bc
Nutriose FM06
c
Pea fiber 50 M
a
Vitacel WF200
ab
none
Supplement

Nested ANOVA

▪ The two factors in the two-way ANOVA design are known as **factorial**:

▪ We have every combination of every level of each factor.

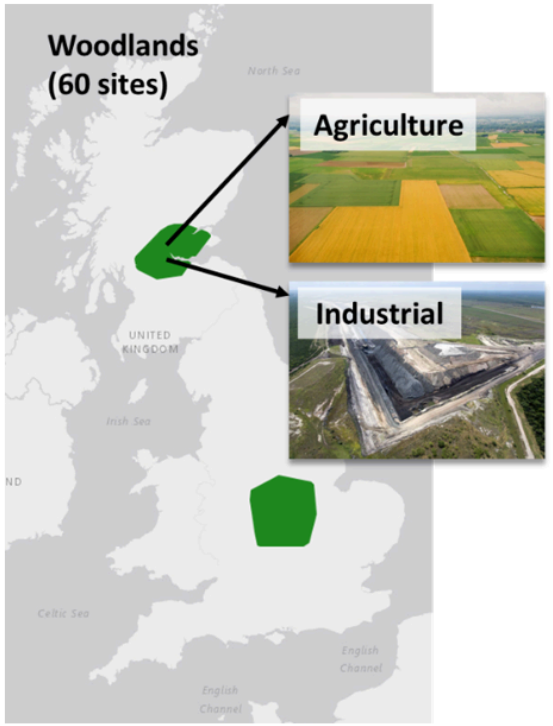
▪ If the levels of one of the factors appear in only one level of the other factors we have **nested** factors

Nested factors

- If levels of factor B appear in only one of the levels in factor A, factor B is termed **nested within factor A**.
- To symbolise that factor B is nested with A we write **B(A)**.
- Nested designs are quite common in ecology and environmental sciences.

Example of Nested Design

- We conduct an experiment to test the soil pH in Agricultural vs Industrial land.
- We have the factor Treatment with 2 levels (Agricultural and Industrial).
- We have 4 areas nested within each treatment and 5 soil samples in each area.



Nested ANOVA in R

▪ The previous design is nested because an area cannot belong to both the Agricultural and the Industrial group.

```
nested.anova <- aov(pH ~ Treatment + Area %in% Treatment, data = pH.soi
summary(nested.anova)
```

```
##                Df Sum Sq Mean Sq F value Pr(>F)
## Treatment       1  5.417   5.417   6.831 0.0143 *
## Treatment:Area  5  5.702   1.140   1.438 0.2416
## Residuals      28 22.204   0.793
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

p>0.05 for the nested factor –> the pH variance within areas in the same group is small.

Fundamental ANOVA assumptions

1. The population from which samples are drawn should be normally distributed.

2. Homogeneity of variance: Homogeneity means that the variance among the groups should be approximately equal.

3. Independence of cases: the sample cases should be **independent** of each other.

Repeated Measures ANOVA

▪ A study design with repeated measures data has multiple dependent variable observations collected at several time points.

▪ This design violates the samples independence assumption.

▪ For time course experiments we can use a **repeated measures ANOVA**.

▪ A repeated measures ANOVA accounts for the **correlation** within and between experimental groups along with the time of the measurements (time point 1, time point 2, …).

ANOVA and Repeated Measures ANOVA Summary

Number of factors	Name	Data organisation	Example
One factor	One-way ANOVA	Several groups. Each experimental unit can only be in one group	Plant yields measured at different nutrient levels: low, medium and high nutrient level). Each plant can only be in one nutrient level group
	(One-way) repeated measures ANOVA	Several related groups. Each experimental unit occurs in all groups. The related groups can be time, or different experimental conditions, the key is that all experimental units experience all conditions	Plant yield measured at different time points. The same plants are measured over time
Two factors (Key interest lies in assessing the interaction between the two factors)	Two-way ANOVA	Each experimental unit can only occur at one combination of the two factors	Plant yields measured at different nutrient levels (low, medium and high) and different light levels (low and bright light). Each plant can only be in one nutrient x light level combination
	Mixed (repeated measures) ANOVA / split plot ANOVA	One factor is fixed and each experimental unit can only occur at one level/in one group. The main effect of this factor is assessed between groups. Second factor contains related groups and each experimental unit occurs at all levels. The main effect of this factor is assessed within groups/individuals	Plant yields at different nutrient levels measured over time. Each plant can only be in one nutrient level group (fixed) but is measured at every time point eg, one week, one month, two months.
	Two-way repeated measures ANOVA	Both factors contain related groups	In a crossover study, each person receives both treatments and outcome is measured at multiple time points

Source: Jenny Freeman, The University of Sheffield