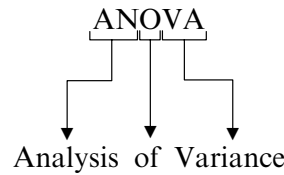


study unit seventeen

hypothesis tests applied
to means: one-way
analysis of variance

You were introduced to the t -test in study units 14 and 15. Now you will be studying a specific kind of F -test: one-way variance analysis, commonly known as



What is analysis of variance (ANOVA)?

When you read the heading of this study unit you probably recognised the word “variance” (from study unit 8, where you learned how to compute it). Now you must be wondering: what is analysis of variance?



To answer this question, read the introduction to Howell’s chapter 16 and study section 16.1.

Mention three assumptions on which ANOVA is based:

- 1 _____
- 2 _____
- 3 _____

If you know the three assumptions, you will understand the rationale of ANOVA.

- 1 Scores are distributed normally around the population mean.
- 2 There is homogeneity of variance, that is the variances of the populations’ scores are the same.
- 3 The observations are independent.

In ANOVA the **dependent variable** is always some kind of measurement — on an interval scale (do you remember this concept from study unit 4?) for parametric tests, and on an ordinal scale (see study unit 4) for nonparametric tests. One-way ANOVA is a parametric test (ie based on a normal distribution); we shall mention a nonparametric (distribution-free) equivalent at the end of this study unit. The researcher has no control over the dependent variable: it is the outcome of manipulation of the independent variable.

The **independent variable** is the one manipulated by the researcher. He or she decides on the number of levels of this variable and makes sure that members of each group are participated to only one of these levels.

The level of the independent variable can be divided into two types:

- 1 Quantitative (interval measurement)
for example: number of hours of study per week (6, 8, 10, 12 hours)
- 2 Qualitative (nominal measurement)
for example: teaching method (lecture, independent study, video, computer-based)

Go back to study unit 4 if you are not sure what dependent and independent variables are.



Create a few different experiments, each with its own independent and dependent variable. Remember that the independent variable must consist of three or more groups or levels to qualify for an ANOVA design. Thus, a variable such as *gender* as the only independent variable is incorrect.

Here are just a few examples. You can combine most of the independent variables (on the left) with most of the dependent variables (on the right), as long as they represent a meaningful study.

Did you also come up with these or similar examples? You may have unique examples of your own. Just compare them with ours to see if your reasoning was correct.

Independent variable

Dependent variable

Departments

Motivation of staff

- Personnel
- Production
- Marketing

(measured by a questionnaire)

Length of service

Number of accidents

- Short : 0 to 2 years
- Medium : 3 to 9 years
- Long : 10 years and longer

on a production line

Levels of management	<i>Quality of worklife</i>
● President	measure
● Vice-president	
● Chief executive officer	
● General manager	
Different shifts	Productivity
● Morning	(measured in terms of
● Afternoon	units manufactured)
● Evening	
● Night	
Training techniques	Number of errors on
● Lecture	a practical task
● Video	
● Self-study	
● Computer-assisted	
● Workshop	

Logic of analysis of variance

Howell explains the principles of ANOVA clearly in section 16.2. You need not be able to do any computations here — we shall get to the computations that you have to be able to do below.

The paragraph headed “Summary of the logic of the analysis of variance” is very useful. Make sure that you understand it.

To clarify the logic of ANOVA even further, let us look at the explanation of some other authors (Fox, Levin & Harkins 1993; Porter & Hamm 1986).

The total variability of a set of scores has two components:

- *Variability within groups.* This measures the normal variability of random errors in the data.
This is the **error effect**, that is, variability resulting from random sampling and chance events.
- *Variability between groups.* This measures the difference between group means.
This is the **treatment effect**, that is, the measure of the effect that the experimental treatment has on participants.

The F ratio is the ratio of variability between groups to variability within groups. It is represented by the following equation:

$$F = \frac{\text{variability between groups}}{\text{variability within groups}}$$

$$= \frac{\text{treatment effect}}{\text{error effect}}$$

When the treatment effect exceeds the error effect by a certain critical value, the F -test is significant. We conclude that there is a significant difference between group means and we reject the null hypothesis.

Practical example

Here is an example of an experiment (Fox, Levin & Harkins 1993, pp 218–219) which provides an excellent graphic illustration of our topic. Don't be put off by the four groups — you will soon see that this does not make the explanation any more complicated than with three groups. (Remember, ANOVA is used when you have three or more groups.)

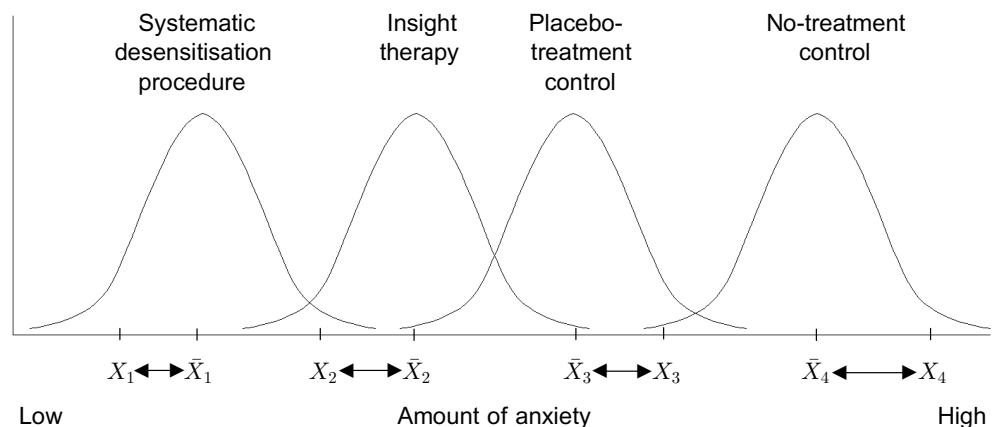
People who become very anxious when they have to speak in public were randomly assigned to four treatment groups for a therapeutic procedure:

- 1 Systematic desensitisation procedure
- 2 Insight therapy
- 3 Placebo therapy — control participants talk about their problem
- 4 No therapy — control no intervention

After exposure to one of the therapies participants had to make a public speech, during which their anxiety level was measured.

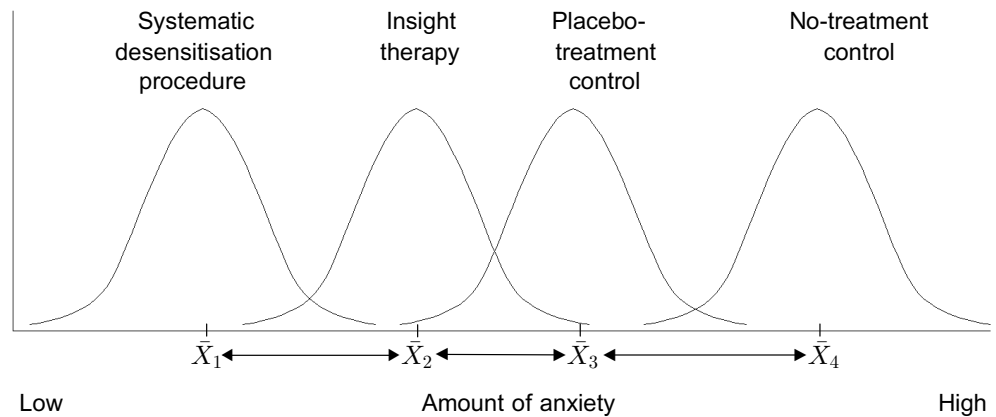
The results of the experiment were as follows:

- **Variability within groups**



Note the variability in any group mean ($\bar{X}_1, \bar{X}_2, \bar{X}_3, \bar{X}_4$) and any individual score within that group (X_1, X_2, X_3, X_4).

● **Variability between groups**



Note that the amount of anxiety differs between groups ($\bar{X}_1, \bar{X}_2, \bar{X}_3, \bar{X}_4$) as a function of the treatment (therapeutic technique), from low anxiety in group 1 to high anxiety in group 4.



Suppose you are an industrial psychologist in Organisation XYZ and you want to investigate the quality of working life in three departments — Personnel, Production and Marketing.

Design your own experiment and represent the variability within and between groups graphically.

Naturally no feedback can be given. Compare your graphs with the ones above — in your experiment you will substitute the three departments for the four types of therapy.

Computations for ANOVA

Time to get out your pocket calculator! The computations you will be doing now are the only ones that you have to be able to do when you are given a data set.

Study section 16.3 in Howell and make sure that you understand all the computations — they are summarised in table 16.4. With regard to the formula for SS_{total} in table 16.4 and on page 310, always leave out the “definition” part, that is $\sum(X_{ij} - \bar{X}_{..})^2$ when you are asked to do computations. The paragraph headed “Conclusions” at the end of this section is very illuminating.



- 1 Remember always to report ANOVA in a *summary table* (like table 16.4(c) in Howell) after you have completed the computations, even when we do not specifically ask you to do so in an assignment or examination question.
- 2 Remember that *SS*, *MS* and *F* can never have a negative value — all *SS* and *MS* values are squares. So if you obtain *negative values* for any of these statistics in your computations you have definitely *miscalculated!* Find the error (also in the examination).
- 3 For a quick check of your interim computation, do the five computations below once you have completed your summary table.

Example of a summary table

Source	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>
Group	2	44,78	22,39	8,08
Error	15	41,50	2,77	
Total	17	86,28		

- 1 $df_{\text{group}} + df_{\text{error}} = df_{\text{total}} \quad : \quad 2 + 15 = 17$
- 2 $SS_{\text{group}} + SS_{\text{error}} = SS_{\text{total}} \quad : \quad 44,78 + 41,50 = 86,28$
- 3 $SS_{\text{group}} \div df_{\text{group}} = MS_{\text{group}} \quad : \quad 44,78/2 = 22,39$
- 4 $SS_{\text{error}} \div df_{\text{error}} = MS_{\text{error}} \quad : \quad 41,50/15 = 2,77$
- 5 $MS_{\text{group}} \div MS_{\text{error}} = F \quad : \quad 22,39/2,77 = 8,08$

These facts will also enable you to draw up a partly completed summary table.



- 1 Do this exercise we found in Fox, Levin & Harkins (1993, pp 259–260).
In a study of worker satisfaction, shift workers whose work hours had been on a phase-advance schedule (progressively earlier starting times e.g., midnight, 16:00, and 08:00) were changed to a phase-delay schedule (progressively after starting times). One-half of these subjects remained on a weekly rotation, while the others moved to a three-week period between changes. A control group of nonrotating workers was included for comparison. Test the null hypothesis that shifts in shifts had no effect on satisfaction.

Set $\alpha = 0,01$ and test the null hypothesis:

$$H_0 : \mu_{\text{one}} = \mu_{\text{three}} = \mu_{\text{control}}$$

SATISFACTION LEVEL

Phase-delay one-week	Phase-delay three-weeks	Control group
13	18	7
11	16	9
14	19	11
10	21	9
12	13	18
8	18	12

(21-point scale; the higher the score, the greater the worker satisfaction)

- 2 Enter the values for N and k in the table below. Look up the critical F values (tables E.3 and E.4 in Howell) and answer “yes” or “no” to the question “Reject H_0 ” on the basis of the computed F -value in the first column.

F	N	k	df_{group}	df_{error}	$Alpha$	Critical F	Reject H_0 ?
5,45			2	18	0,05		
5,45			2	18	0,01		
4,95			3	12	0,01		
8,06			4	15	0,05		
6,22			5	24	0,05		

- 3 Complete the following summary table:

Source	df	SS	MS	F
Group		225		
Error	18		5	
Total	23			

- 1 X_A = Phase delay one-week
 X_B = Phase delay three-weeks
 X_C = Control group

	X_A	X_B	X_C	Total
\bar{X}	11,33	17,50	11,00	13,28
s	2,16	2,74	3,85	4,17

X_A	X_B	X_C	X_A^2	X_B^2	X_C^2
13	18	7	169	324	49
11	16	9	121	256	81
14	19	11	196	361	121
10	21	9	100	441	81
12	13	18	144	169	324
8	18	12	64	324	144
68	105	66	794	1 875	800

$$\Sigma X = 68 + 105 + 66 = 239$$

$$\Sigma X^2 = 794 + 1 875 + 800 = 3 469$$

Null hypothesis $H_0 : \mu_A = \mu_B = \mu_C$

One possible alternative hypothesis $H_1 : \mu_A \neq \mu_B \neq \mu_C$

Other possible alternative hypotheses $H_1 : \mu_A > \mu_B > \mu_C$

$H_1 : \mu_A < \mu_B < \mu_C$

etc.

$$SS_{\text{total}} = \Sigma X^2 - \frac{(\Sigma X)^2}{N}$$

$$= 3469 - \frac{(239)^2}{18}$$

$$= 3469 - \frac{57121}{18}$$

$$= 3469 - 3173,39$$

$$= 295,61$$

$$SS_{\text{group}} = n \Sigma (\bar{X}_j - \bar{X}_{..})^2$$

$$= 6[(11,33 - 13,28)^2 + (17,5 - 13,28)^2 + (11,0 - 13,28)^2]$$

$$= 6[(-1,95)^2 + (4,22)^2 + (-2,28)^2]$$

$$= 6[3,80 + 17,81 + 5,20] = 6(26,81)$$

$$= 160,86$$

$$SS_{\text{error}} = SS_{\text{total}} - SS_{\text{group}}$$

$$= 295,61 - 160,86$$

$$= 134,75$$

$$df_{\text{total}} = N - 1 \quad df_{\text{group}} = k - 1 \quad df_{\text{error}} = k(n - 1)$$

$$= 18 - 1 \quad = 3 - 1 \quad = 3(6 - 1)$$

$$= 17 \quad = 2 \quad = 15$$

$$MS_{\text{group}} = SS_{\text{group}} / df_{\text{error}}$$

$$= 160,86/2$$

$$= 80,43$$

$$MS_{\text{error}} = SS_{\text{error}} / df_{\text{error}}$$

$$= 134,75/15$$

$$= 8,98$$

$$\begin{aligned}
 F &= \frac{MS_{\text{group}}}{MS_{\text{error}}} \\
 &= \frac{80,43}{8,98} \\
 &= 8,96
 \end{aligned}$$

Source	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>
Group	2	160,86	80,43	8,96
Error	15	134,75	8,98	
Total	17	295,61		

$$F_{0,01} (2,15) = 6,36$$

From Table E.4

$$F_{\text{calc}} = 8,96$$

$$8,96 > 6,36$$

∴ Reject H_0

Interpretation

There is a significant difference in worker satisfaction among the three groups of shift workers. This can be said with 99% certainty.

2

<i>F</i>	<i>N</i>	<i>k</i>	<i>df_{group}</i>	<i>df_{error}</i>	<i>Alpha</i>	<i>Critical F</i>	<i>Reject H₀?</i>
5,45	21	3	2	18	0,05	3,55	Yes
5,45	21	3	2	18	0,01	6,01	No
4,95	16	4	3	12	0,01	5,95	No
8,06	20	5	4	15	0,05	3,06	Yes
6,22	30	6	5	24	0,05	2,62	Yes

3

Source	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>
Group	5	225	45	9
Error	18	90	5	
Total	23	315		

A clearer grasp of ANOVA

Before you resume your study of Howell we want to present you with a useful, concrete representation of ANOVA by David Johnson (1989), a psychologist teaching statistics. His technique introduces students to an intuitive understanding of the principles of between-group variance (MS_{group}) and within-group variance (MS_{error}), and their interrelationship. If you study it *carefully* you may find it useful.

The experiment

The data derive from a **hypothetical experiment**. Three independent groups comprising five participants each are exposed to one of three levels of an independent variable. The spectrum of possible responses to the dependent variable ranges from 0 to 4.

The results

Below are **five possible outcomes** of the experiment. The data set and summary table are presented together with an explanation.

Outcome 1

Data

Group 1	Group 2	Group 3
1	1	1
1	1	1
1	1	1
1	1	1
1	1	1

Summary table

Source	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>
Group	2	0	0	—
Error	12	0	0	
Total	14	0		

You no doubt realise that this is a highly unlikely outcome. Note that there is **no variability** in the data at all between and within groups and that this leads to mean squares of 0 in the summary table.

Outcome 2

Data

Group 1	Group 2	Group 3
1	1	3
1	1	3
1	1	3
1	1	3
1	1	3

Summary table

Source	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>
Group	2	13,33	6,67	—
Error	12	0	0	
Total	14	13,33		

Between-groups variance is introduced by increasing all five values in group 3 to 3.

This indicates the possible effect of the independent variable which is manipulated. Group 3 appears to differ from the other two groups ($MS_{\text{group}} = 6,67$); but there is no **within-groups variance** ($MS_{\text{error}} = 0$).

Outcome 3

Data

Group 1	Group 2	Group 3
0	1	3
1	1	3
1	1	3
1	1	3
2	1	3

Summary Table

Source	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>
Group	2	13,33	6,67	39,24
Error	12	2	0,17	
Total	14	15,33		

A third data set introduces **within-groups variance** in group 1. This data set contains the same measure of between-groups variance as the previous one, that is, the means remain unchanged. Note how the introduction of dissimilar scores in group 1 increases the value of the **mean square** for within-groups variance ($MS_{\text{error}} = 0,17$). Note also the value of the *F* statistic ($F = 39,24$).

Outcome 4

Data

Group 1	Group 2	Group 3
0	0	3
1	1	3
1	1	3
1	1	3
2	2	3

Summary table

Source	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>
Group	2	13,33	6,67	20,21
Error	12	4	0,33	
Total	14	17,33		

This data set differs from the previous one in only one respect: **additional within-groups variance** is introduced by changing the values in group 2 to match those in group 1. Compared with the previous data set the between-groups mean square remains unchanged, but the **within-groups mean square increases** ($MS_{\text{error}} = 0,33$). As a result the *F* value is lower ($F = 20,21$).

Outcome 5

Data

Group 1	Group 2	Group 3
0	0	2
1	1	3
1	1	3
1	1	3
2	2	4

Summary table

Source	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>
Group	2	13,33	6,67	13,34
Error	12	6	0,5	
Total	14	19,33		

Additional within-groups variance is introduced by changing two values in group 3. Note the unchanged between-groups mean square and the increased **within-groups mean square** ($MS_{\text{error}} = 0,5$), compared with the previous data set. As a result the F statistic is **smaller** than before ($F = 13,34$).



So there you have it! We are sure you now have a better grasp of the concepts relating to ANOVA and the way in which the dynamic relationship of between-groups and within-groups variance determines the F statistic.



Your company, ABC MARKETING, has bought a new desktop publishing package called BestPublish. The training manager of the Computer Services Department, who is looking for an effective training method for employees, decides to run an experiment. She uses three different training methods to train the employees in BestPublish. All three methods utilise the tutorial of BestPublish. The different training techniques thus comprise the following:

- **Lecture group:** Tutorial + lectures by the training manager
- **Assistance group:** Tutorial + assistance available
(Training manager available for help when Lecture group does the practical exercises – five 15-minute periods during the day)
- **Self-study group:** Only the online tutorial help is available.

Thus, *training method* is the independent variable in this study.

She asks you, as the industrial psychologist of ABC MARKETING, to help with the design of the experiment, analysis and interpretation of the data. You randomly draw 27 employees from the total of 120 employees and randomly allocate nine people to each of the three training groups.

These 27 employees all do the 4-day training course (7 hours per day), each group of nine according to one of the three methods described above. On the fifth day all of them do the same two-hour practical skills test. The data comprise a score out of 20.

Thus the *test score* is the dependent variable in this study.

The training manager gives you the data below and asks you to determine which method is the best for training employees in BestPublish.

Lecture	Assistance	Self-study
12	14	17
10	10	16
13	10	12
12	8	10
9	10	18
10	10	10
15	12	15
8	8	11
17	11	10

She also tells you that she has already computed the means and standard deviations out of curiosity. These are given next.

	Lecture	Assistance	Self-study	Total
\bar{X}	11,78	10,33	13,22	11,78
s	2,91	1,87	3,27	2,90

She remarks that *self-study* is probably the best method because the mean for this group is the highest. You tell her that no conclusion can be based upon the means only, and that you will give her your conclusion after you have tested the following null hypothesis:

$$\mu_{\text{Lecture}} = \mu_{\text{Assistance}} = \mu_{\text{Self-study}}$$

You decide to use $\alpha = 0,05$.

- 1 Do the appropriate statistical test, that is the F -test, and give your answer in a summary table.
 - 2 Draw your conclusion.
 - 2.1 What is the critical value?
 - 2.2 Do you reject H_0 ?
 - 2.3 What is your answer to the training manager? (Remember that she does not know what is meant by rejection or nonrejection of the null hypothesis.)
-

1

$X_L = \text{Lecture}$ $X_A = \text{Assistance}$ $X_S = \text{Self-study}$ **Total**
 $\bar{X}_L = 11,78$ $\bar{X}_A = 10,33$ $\bar{X}_S = 13,22$ $\bar{X}_{..} = 11,78$

X_L	X_A	X_S	X_L^2	X_A^2	X_S^2
12	14	17	144	196	289
10	10	16	100	100	256
13	10	12	169	100	144
12	8	10	144	64	100
9	10	18	81	100	324
10	10	10	100	100	100
15	12	15	225	144	225
8	8	11	64	64	121
17	11	10	289	121	100
106	93	119	1 316	989	1 659

$$\Sigma X = 106 + 93 + 119 = 318$$

$$\Sigma X^2 = 1\,316 + 989 + 1\,659 = 3\,964$$

$$\begin{aligned} SS_{\text{total}} &= \Sigma X^2 - \frac{(\Sigma X)^2}{N} & df_{\text{total}} &= N - 1 \\ &= 3964 - \frac{(318)^2}{27} & &= 27 - 1 \\ &= 3964 - \frac{101124}{27} & &= 26 \\ &= 3964 - 3745,33 \\ &= 218,67 \end{aligned}$$

$$\begin{aligned} SS_{\text{group}} &= n \Sigma (\bar{X}_j - \bar{X}_{..})^2 & df_{\text{group}} &= k - 1 \\ &= 9[(11,78 - 11,78)^2 + (10,33 - 11,78)^2 \\ &\quad + (13,22 - 11,78)^2] & &= 3 - 1 \\ &= 9[0^2 + (-1,45)^2 + (1,44)^2] & &= 2 \\ &= 9[0 + 2,10 + 2,07] = 9(4,17) \\ &= 37,53 \end{aligned}$$

$$\begin{aligned} SS_{\text{error}} &= SS_{\text{total}} - SS_{\text{group}} & df_{\text{error}} &= k(n - 1) \\ &= 218,67 - 37,53 & &= 3(9 - 1) \\ &= 181,14 & &= 24 \end{aligned}$$

$$\begin{aligned} MS_{\text{group}} &= SS_{\text{group}} / df_{\text{group}} \\ &= 37,53 / 2 \\ &= 18,77 \end{aligned}$$

$$\begin{aligned} MS_{\text{error}} &= SS_{\text{error}} / df_{\text{error}} \\ &= 181,14 / 24 \\ &= 7,55 \end{aligned}$$

$$\begin{aligned}
 F &= \frac{MS_{\text{group}}}{MS_{\text{error}}} \\
 &= \frac{18,77}{7,55} \\
 &= 2,49
 \end{aligned}$$

Source	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>
Group	2	37,53	18,77	2,49
Error	24	181,14	7,55	
Total	26	218,67		

2 Your conclusion

2.1 $F_{0,05}(2, 24) = 3,40$

2.2 No

(Because $2,49 < 3,40$)

2.3 There is no significant difference between the three training methods. I can say this with 95% certainty.

Thus, none of the training methods is superior to any of the others, in the sense that it generates better skills test results. You could use any training method for the publishing package.

A nonparametric equivalent of ANOVA

The **Kruskal-Wallis** one-way ANOVA is the nonparametric or distribution-free equivalent of the standard one-way ANOVA or *F*-test. In this technique it is assumed that the dependent variable has an underlying continuous distribution of at least ordinal measurement. Study the first paragraph of section 20.3 in Howell.



We use the data from an exercise in Lehman (1991, p 376, ex 3) to illustrate a typical computer output using *SPSS for Windows*.

The data are the time (in minutes) that it took students to complete four tests of different length comprising 10, 15, 20 and 25 four-alternative multiple-choice items. The research problem was: Is there a significant difference in the time that it takes students to complete tests comprising varying numbers of multiple-choice items? The data are given in the table below.

Number of test items			
10	15	20	25
13	17	18	16
10	5	17	21
13	13	10	28
5	14	23	32
6	11	24	24
6	12	17	12
2	3	10	9
12	16	27	18
17	18	21	18
12	8	24	9
12	14	23	5
2	7	8	15
17	13	14	10
3	16	1	7
7	14	8	30
1	18	10	5

Go through the printout and note particularly how the results are presented.

*** ANALYSIS OF VARIANCE ***

N_ITEMS
by GROUP No of Test Items

UNIQUE sums of squares
All effects entered simultaneously

Source of Variation	Sum of Squares	DF	Mean Square	F	Sig of F
Main Effects	606,297	3	202,099	4,421	,007
GROUP	606,297	3	202,099	4,421	,007
Explained Group	606,297	3	202,099	4,421	,007
Residual Error	2743,062	60	45,718		
Total	3349,359	63	53,164		

This output makes provision for factorial ANOVA
If you ignore the grey coloured area, you will recognise the one-way ANOVA table

Do we reject *H*₀? **Yes**

Exercises 16.18, 16.19 and 16.21

Exercise 16.21 looks like this:

Summary of # of rear end collisions			
conditio	Mean	Std. Dev.	Freq.
active	9.9333333	6.005553	15
delayed	6.8	5.4405882	15
nonsmoke	2.3333333	2.2886885	15
Total	6.3555556	5.7014972	45

Descriptive statistics for each group

Analysis of Variance					
Source	SS	df	MS	F	Prob > F
Between groups	437.644444	2	218.822222	9.26	0.0005
Within groups	992.666667	42	23.6349206		
Total	1430.31111	44	32.5070707		

Bartlett's test for equal variances: $\chi^2(2) = 11.6570$ Prob>chi2 = 0.003

In some statistical packages the output swops the *df* and *SS* columns around. Keep to what we have taught you, ie first *df* and then *SS*

F value is significant: $p < 0.01$

Thus: Reject H_0

Exercise 16.18: $F = 0,05$ Prob > $F = 0,95$ ∴ Do not reject H_0

Exercise 16.19: $F = 4,74$ $p < 0,01$ ∴ Reject H_0

See *Student handbook* for computer output

Having completed study unit 17 you should be able to

- mention and briefly define the assumptions on which ANOVA is based
- make ANOVA computations and report them in a summary table
- interpret the *F* value
- give the name of the nonparametric equivalent of one-way ANOVA
- interpret a computer printout of ANOVA

Conclusion

We conclude your learning material for t - and F -tests with a schematic representation of hypothesis tests for means in order to put all of them in perspective.

