

Analysis of Variance

44.1	One-Way Analysis of Variance	2
44.2	Two-Way Analysis of Variance	15
44.3	Experimental Design	40

Learning Outcomes

In this Workbook you will learn the basics of this very important branch of Statistics and how to do the calculations which enable you to draw conclusions about variance found in data sets. You will also be introduced to the design of experiments which has great importance in science and engineering.

One-Way Analysis of Variance

44.1

Introduction

Problems in engineering often involve the exploration of the relationships between values taken by a variable under different conditions. HELM 41 introduced hypothesis testing which enables us to compare two population means using hypotheses of the general form

$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu_1 \neq \mu_2$$

or, in the case of more than two populations,

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

$$H_1 : H_0 \text{ is not true}$$

If we are comparing more than two population means, using the type of hypothesis testing referred to above gets very clumsy and very time consuming. As you will see, the statistical technique called Analysis of Variance (ANOVA) enables us to compare several populations **simultaneously**. We might, for example need to compare the shear strengths of five different adhesives or the surface toughness of six samples of steel which have received different surface hardening treatments.



Prerequisites

Before starting this Section you should ...

- be familiar with the general techniques of hypothesis testing
- be familiar with the F -distribution



Learning Outcomes

On completion you should be able to ...

- describe what is meant by the term one-way ANOVA.
- perform one-way ANOVA calculations.
- interpret the results of one-way ANOVA calculations

1. One-way ANOVA

In this Workbook we deal with one-way analysis of variance (one-way ANOVA) and two-way analysis of variance (two-way ANOVA). One-way ANOVA enables us to compare several means simultaneously by using the F -test and enables us to draw conclusions about the variance present in the set of samples we wish to compare.

Multiple (greater than two) samples may be investigated using the techniques of two-population hypothesis testing. As an example, it is possible to do a comparison looking for variation in the surface hardness present in (say) three samples of steel which have received different surface hardening treatments by using hypothesis tests of the form

$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu_1 \neq \mu_2$$

We would have to compare all possible pairs of samples before reaching a conclusion. If we are dealing with three samples we would need to perform a total of

$${}^3C_2 = \frac{3!}{1!2!} = 3$$

hypothesis tests. From a practical point of view this is not an efficient way of dealing with the problem, especially since the number of tests required rises rapidly with the number of samples involved. For example, an investigation involving ten samples would require

$${}^{10}C_2 = \frac{10!}{8!2!} = 45$$

separate hypothesis tests.

There is also another crucially important reason why techniques involving such batteries of tests are unacceptable. In the case of 10 samples mentioned above, if the probability of correctly accepting a given null hypothesis is 0.95, then the probability of correctly accepting the null hypothesis

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_{10}$$

is $(0.95)^{45} \approx 0.10$ and we have only a 10% chance of correctly accepting the null hypothesis for all 45 tests. Clearly, such a low success rate is unacceptable. These problems may be avoided by simultaneously testing the significance of the difference between a set of more than two population means by using techniques known as the analysis of variance.

Essentially, we look at the variance *between* samples and the variance *within* samples and draw conclusions from the results. Note that the variation between samples is due to assignable (or controlled) causes often referred in general as **treatments** while the variation within samples is due to **chance**. In the example above concerning the surface hardness present in three samples of steel which have received different surface hardening treatments, the following diagrams illustrate the differences which may occur when between sample and within sample variation is considered.

Case 1

In this case the variation within samples is roughly on a par with that occurring between samples.

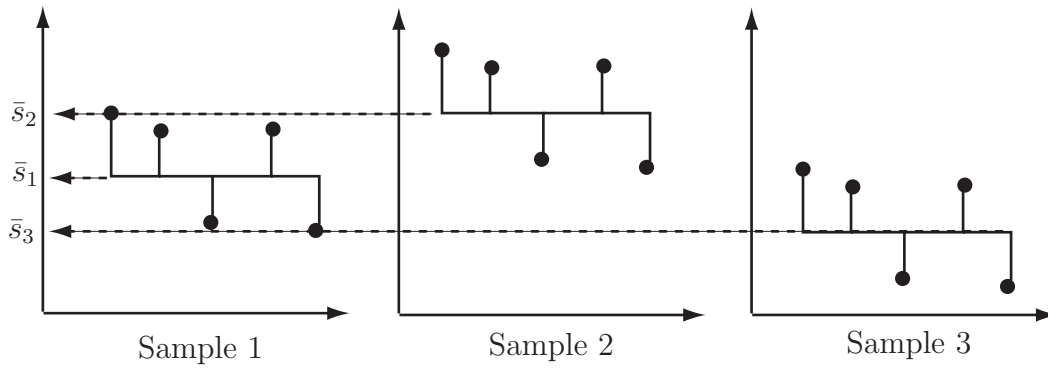


Figure 1

Case 2

In this case the variation within samples is considerably less than that occurring between samples.

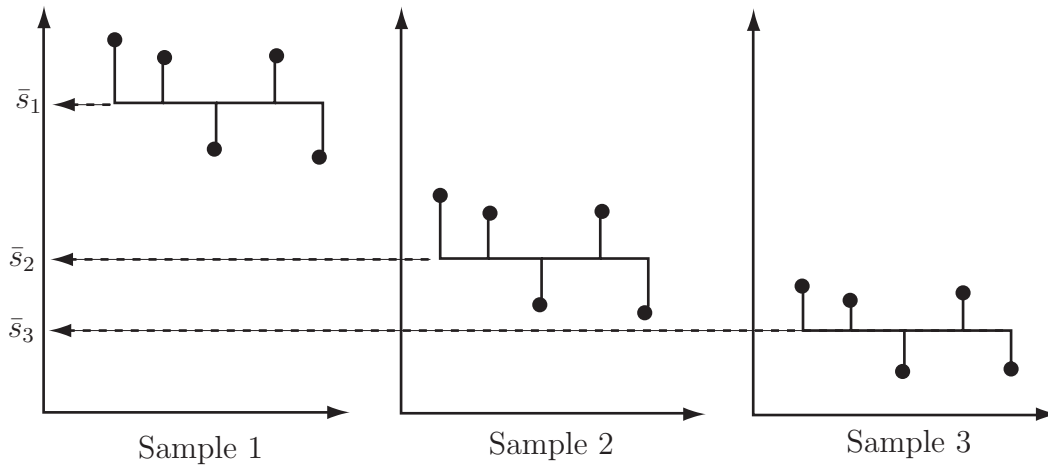


Figure 2

We argue that the greater the variation present *between* samples in comparison with the variation present *within* samples the more likely it is that there are ‘real’ differences between the population means, say μ_1, μ_2 and μ_3 . If such ‘real’ differences are shown to exist at a sufficiently high level of significance, we may conclude that there is sufficient evidence to enable us to reject the null hypothesis $H_0 : \mu_1 = \mu_2 = \mu_3$.

Example of variance in data

This example looks at variance in data. Four machines are set up to produce alloy spacers for use in the assembly of microlight aircraft. The spaces are supposed to be identical but the four machines give rise to the following varied lengths in mm.

Machine A	Machine B	Machine C	Machine D
46	56	55	49
54	55	51	53
48	56	50	57
46	60	51	60
56	53	53	51

Since the machines are set up to produce identical alloy spacers it is reasonable to ask if the evidence we have suggests that the machine outputs are the same or different in some way. We are really asking whether the sample means, say $\bar{X}_A, \bar{X}_B, \bar{X}_C$ and \bar{X}_D , are different because of differences in the respective population means, say μ_A, μ_B, μ_C and μ_D , or whether the differences in $\bar{X}_A, \bar{X}_B, \bar{X}_C$ and \bar{X}_D may be attributed to chance variation. Stated in terms of a hypothesis test, we would write

$$H_0 : \mu_A = \mu_B = \mu_C = \mu_D$$

H_1 : At least one mean is different from the others

In order to decide between the hypotheses, we calculate the mean of each sample and overall mean (the mean of the means) and use these quantities to calculate the variation present *between* the samples. We then calculate the variation present *within* samples. The following tables illustrate the calculations.

$$H_0 : \mu_A = \mu_B = \mu_C = \mu_D$$

H_1 : At least one mean is different from the others

Machine A	Machine B	Machine C	Machine D
46	56	55	49
54	55	51	53
48	56	50	57
46	60	51	60
56	53	53	51
$\bar{X}_A = 50$	$\bar{X}_B = 56$	$\bar{X}_C = 52$	$\bar{X}_D = 54$

The mean of the means is clearly

$$\bar{\bar{X}} = \frac{50 + 56 + 52 + 54}{4} = 53$$

so the variation present between samples may be calculated as

$$\begin{aligned} S_{Tr}^2 &= \frac{1}{n-1} \sum_{i=A}^D (\bar{X}_i - \bar{\bar{X}})^2 \\ &= \frac{1}{4-1} ((50-53)^2 + (56-53)^2 + (52-53)^2 + (54-53)^2) \\ &= \frac{20}{3} = 6.67 \quad \text{to 2 d.p.} \end{aligned}$$

Note that the notation S_{Tr}^2 reflects the general use of the word 'treatment' to describe assignable causes of variation between samples. This notation is not universal but it is fairly common.

Variation within samples

We now calculate the variation due to chance errors present *within* the samples and use the results to obtain a pooled estimate of the variance, say S_E^2 , present *within* the samples. After this calculation we will be able to compare the two variances and draw conclusions. The variance present within the samples may be calculated as follows.

Sample A

$$\sum (X - \bar{X}_A)^2 = (46 - 50)^2 + (54 - 50)^2 + (48 - 50)^2 + (46 - 50)^2 + (56 - 50)^2 = 88$$

Sample B

$$\sum (X - \bar{X}_B)^2 = (56 - 56)^2 + (55 - 56)^2 + (56 - 56)^2 + (60 - 56)^2 + (53 - 56)^2 = 26$$

Sample C

$$\sum (X - \bar{X}_C)^2 = (55 - 52)^2 + (51 - 52)^2 + (50 - 52)^2 + (51 - 52)^2 + (53 - 52)^2 = 16$$

Sample D

$$\sum (X - \bar{X}_D)^2 = (49 - 54)^2 + (53 - 54)^2 + (57 - 54)^2 + (60 - 54)^2 + (51 - 54)^2 = 80$$

An obvious extension of the formula for a pooled variance gives

$$S_E^2 = \frac{\sum (X - \bar{X}_A)^2 + \sum (X - \bar{X}_B)^2 + \sum (X - \bar{X}_C)^2 + \sum (X - \bar{X}_D)^2}{(n_A - 1) + (n_B - 1) + (n_C - 1) + (n_D - 1)}$$

where n_A, n_B, n_C and n_D represent the number of members (5 in each case here) in each sample. Note that the quantities comprising the denominator $n_A - 1, \dots, n_D - 1$ are the number of degrees of freedom present in each of the four samples. Hence our pooled estimate of the variance present within the samples is given by

$$S_E^2 = \frac{88 + 26 + 16 + 80}{4 + 4 + 4 + 4} = 13.13$$

We are now in a position to ask whether the variation *between* samples S_{Tr}^2 is large in comparison with the variation *within* samples S_E^2 . The answer to this question enables us to decide whether the difference in the calculated variations is sufficiently large to conclude that there is a difference in the population means. That is, do we have sufficient evidence to reject H_0 ?

Using the F -test

At first sight it seems reasonable to use the ratio

$$F = \frac{S_{Tr}^2}{S_E^2}$$

but in fact the ratio

$$F = \frac{nS_{Tr}^2}{S_E^2},$$

where n is the sample size, is used since it can be shown that if H_0 is true this ratio will have a value of approximately unity while if H_0 is not true the ratio will have a value greater than unity. This is because the variance of a sample mean is σ^2/n .

The test procedure (three steps) for the data used here is as follows.

- (a) Find the value of F ;
- (b) Find the number of degrees of freedom for both the numerator and denominator of the ratio;

- (c) Accept or reject depending on the value of F compared with the appropriate tabulated value.

Step 1

The value of F is given by

$$F = \frac{nS_{Tr}^2}{S_E^2} = \frac{5 \times 6.67}{13.13} = 2.54$$

Step 2

The number of degrees of freedom for S_{Tr}^2 (the numerator) is

$$\text{Number of samples} - 1 = 3$$

The number of degrees of freedom for S_E^2 (the denominator) is

$$\text{Number of samples} \times (\text{sample size} - 1) = 4 \times (5 - 1) = 16$$

Step 3

The critical value (5% level of significance) from the F -tables (Table 1 at the end of this Workbook) is $F_{(3,16)} = 3.24$ and since $2.54 < 3.224$ we see that we cannot reject H_0 on the basis of the evidence available and conclude that in this case the variation present is due to chance. Note that the test used is one-tailed.

ANOVA tables

It is usual to summarize the calculations we have seen so far in the form of an ANOVA table. Essentially, the table gives us a method of recording the calculations leading to both the numerator and the denominator of the expression

$$F = \frac{nS_{Tr}^2}{S_E^2}$$

In addition, and importantly, ANOVA tables provide us with a useful means of checking the accuracy of our calculations. A general ANOVA table is presented below with explanatory notes.

Define a = number of treatments, n = number of observations per sample.

Source of Variation	Sum of Squares SS	Degrees of Freedom	Mean Square MS	Value of F Ratio
<i>Between samples</i> (due to treatments) Differences between means \bar{X}_i and \bar{X}	$SS_{Tr} = n \sum_{i=1}^a (\bar{X}_i - \bar{X})^2$	$(a - 1)$	$MS_{Tr} = \frac{SS_{Tr}}{(a - 1)}$ $= nS_{\bar{X}}^2$	$F = \frac{MS_{Tr}}{MS_E}$ $= \frac{nS_{Tr}^2}{S_E^2}$
<i>Within samples</i> (due to chance errors) Differences between individual observations X_{ij} and means \bar{X}_i	$SS_E = \sum_{i=1}^a \sum_{j=1}^n (X_{ij} - \bar{X}_i)^2$	$a(n - 1)$	$MS_E = \frac{SS_E}{a(n - 1)}$ $= S_E^2$	
TOTALS	$SS_T = \sum_{i=1}^a \sum_{j=1}^n (X_{ij} - \bar{X})^2$	$(an - 1)$		

In order to demonstrate this table for the example above we need to calculate

$$SS_T = \sum_{i=1}^a \sum_{j=1}^n (X_{ij} - \bar{X})^2$$

a measure of the total variation present in the data. Such calculations are easily done using a computer (Microsoft Excel was used here), the result being

$$SS_T = \sum_{i=1}^a \sum_{j=1}^n (X_{ij} - \bar{X})^2 = 310$$

The ANOVA table becomes

Source of Variation	Sum of Squares SS	Degrees of Freedom	Mean Square MS	Value of F Ratio
<i>Between samples</i> <i>(due to treatments)</i> Differences between means \bar{X}_i and \bar{X}	100	3	$MS_{Tr} = \frac{SS_{Tr}}{(a-1)}$ $= \frac{100}{3}$ $= 33.33$	$F = \frac{MS_{Tr}}{MS_E}$ $= 2.54$
<i>Within samples</i> <i>(due to chance errors)</i> Differences between individual observations X_{ij} and means \bar{X}_i	210	16	$MS_E = \frac{SS_E}{a(n-1)}$ $= \frac{210}{16}$ $= 13.13$	
TOTALS	310	19		

It is possible to show theoretically that

$$SS_T = SS_{Tr} + SS_E$$

that is

$$\sum_{i=1}^a \sum_{j=1}^n (X_{ij} - \bar{X})^2 = n \sum_{i=1}^a (\bar{X}_i - \bar{X})^2 + \sum_{i=1}^a \sum_{j=1}^n (X_{ij} - \bar{X}_i)^2$$

As you can see from the table, SS_{Tr} and SS_E do indeed sum to give SS_T even though we can calculate them separately. The same is true of the degrees of freedom.

Note that calculating these quantities separately does offer a check on the arithmetic but that using the relationship can speed up the calculations by obviating the need to calculate (say) SS_T . As you might expect, it is recommended that you check your calculations! However, you should note that it is usual to calculate SS_T and SS_{Tr} and then find SS_E by subtraction. This saves a lot of unnecessary calculation but does not offer a check on the arithmetic. This shorter method will be used throughout much of this Workbook.

Unequal sample sizes

So far we have assumed that the number of observations in each sample is the same. This is not a necessary condition for the one-way ANOVA.



Key Point 1

Suppose that the number of samples is a and the numbers of observations are n_1, n_2, \dots, n_a . Then the between-samples sum of squares can be calculated using

$$SS_{Tr} = \sum_{i=1}^a \frac{T_i^2}{n_i} - \frac{G^2}{N}$$

where T_i is the total for sample i , $G = \sum_{i=1}^a T_i$ is the overall total and $N = \sum_{i=1}^a n_i$.

It has $a - 1$ degrees of freedom.

The total sum of squares can be calculated as before, or using

$$SS_T = \sum_{i=1}^a \sum_{j=1}^{n_i} X_{ij}^2 - \frac{G^2}{N}$$

It has $N - 1$ degrees of freedom.

The within-samples sum of squares can be found by subtraction:

$$SS_E = SS_T - SS_{Tr}$$

It has $(N - 1) - (a - 1) = N - a$ degrees of freedom.



Three fuel injection systems are tested for efficiency and the following coded data are obtained.

System 1	System 2	System 3
48	60	57
56	56	55
46	53	52
45	60	50
50	51	51

Do the data support the hypothesis that the systems offer equivalent levels of efficiency?

Your solution

Answer

Appropriate hypotheses are

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

H_1 : At least one mean is different to the others

Variation between samples

System 1	System 2	System 3
48	60	57
56	56	55
46	53	52
45	60	50
50	51	51
$\bar{X}_1 = 49$	$\bar{X}_2 = 56$	$\bar{X}_3 = 53$

The mean of the means is $\bar{\bar{X}} = \frac{49 + 56 + 53}{3} = 52.67$ and the variation present *between* samples is

$$S_{Tr}^2 = \frac{1}{n-1} \sum_{i=1}^3 (\bar{X}_i - \bar{\bar{X}})^2 = \frac{1}{3-1} ((49 - 52.67)^2 + (56 - 52.67)^2 + (53 - 52.67)^2) = 12.33$$

Variation within samples**System 1**

$$\sum (X - \bar{X}_1)^2 = (48 - 49)^2 + (56 - 49)^2 + (46 - 49)^2 + (45 - 49)^2 + (51 - 49)^2 = 76$$

System 2

$$\sum (X - \bar{X}_2)^2 = (60 - 56)^2 + (56 - 56)^2 + (53 - 56)^2 + (60 - 56)^2 + (51 - 56)^2 = 66$$

System 3

$$\sum (X - \bar{X}_3)^2 = (57 - 53)^2 + (55 - 53)^2 + (52 - 53)^2 + (50 - 53)^2 + (51 - 53)^2 = 34$$

Hence

$$S_E^2 = \frac{\sum (X - \bar{X}_1)^2 + \sum (X - \bar{X}_2)^2 + \sum (X - \bar{X}_3)^2}{(n_1 - 1) + (n_2 - 1) + (n_3 - 1)} = \frac{76 + 66 + 34}{4 + 4 + 4} = 14.67$$

The value of F is given by $F = \frac{nS_{Tr}^2}{S_E^2} = \frac{5 \times 12.33}{14.67} = 4.20$

The number of degrees of freedom for S_{Tr}^2 is No. of samples $- 1 = 2$

The number of degrees of freedom for S_E^2 is No. of samples \times (sample size $- 1$) $= 12$

The critical value (5% level of significance) from the F -tables (Table 1 at the end of this Workbook) is $F_{(2,12)} = 3.89$ and since $4.20 > 3.89$ we conclude that we have sufficient evidence to reject H_0 so that the injection systems are not of equivalent efficiency.

Exercises

1. The yield of a chemical process, expressed in percentage of the theoretical maximum, is measured with each of two catalysts, A, B, and with no catalyst (Control: C). Five observations are made under each condition. Making the usual assumptions for an analysis of variance, test the hypothesis that there is no difference in mean yield between the three conditions. Use the 5% level of significance.

Catalyst A	Catalyst B	Control C
79.2	81.5	74.8
80.1	80.7	76.5
77.4	80.5	74.7
77.6	81.7	74.8
77.8	80.6	74.9

2. Four large trucks, A, B, C, D, are used to move stone in a quarry. On a number of days, the amount of fuel, in litres, used per tonne of stone moved is calculated for each truck. On some days a particular truck might not be used. The data are as follows. Making the usual assumptions for an analysis of variance, test the hypothesis that the mean amount of fuel used per tonne of stone moved is the same for each truck. Use the 5% level of significance.

Truck	Observations										
A	0.21	0.21	0.21	0.21	0.20	0.19	0.18	0.21	0.22	0.21	
B	0.22	0.22	0.25	0.21	0.21	0.22	0.20	0.23			
C	0.21	0.18	0.18	0.19	0.20	0.18	0.19	0.19	0.20	0.20	0.20
D	0.20	0.20	0.21	0.21	0.21	0.19	0.20	0.20	0.21		

Answers

1. We calculate the treatment totals for A: 392.1, B: 405.0 and C: 375.7. The overall total is 1172.8 and $\sum \sum y^2 = 91792.68$.

The total sum of squares is

$$91792.68 - \frac{1172.8^2}{15} = 95.357$$

on $15 - 1 = 14$ degrees of freedom.

The between treatments sum of squares is

$$\frac{1}{5}(392.1^2 + 405.0^2 + 375.7^2) - \frac{1172.8^2}{15} = 86.257$$

on $3 - 1 = 2$ degrees of freedom.

By subtraction, the residual sum of squares is

$$95.357 - 86.257 = 9.100$$

on $14 - 2 = 12$ degrees of freedom.

The analysis of variance table is as follows:

Source of variation	Sum of squares	Degrees of freedom	Mean square	Variance ratio
Treatment	86.257	2	43.129	56.873
Residual	9.100	12	0.758	
Total	95.357	14		

The upper 5% point of the $F_{2,12}$ distribution is 3.89. The observed variance ratio is greater than this so we conclude that the result is significant at the 5% level and we reject the null hypothesis at this level. The evidence suggests that there are differences in the mean yields between the three treatments.

Answer

2. We can summarise the data as follows.

Truck	$\sum y$	$\sum y^2$	n
A	2.05	0.4215	10
B	1.76	0.3888	8
C	2.12	0.4096	11
D	1.83	0.3725	9
Total	7.76	1.5924	38

The total sum of squares is

$$1.5924 - \frac{7.76^2}{38} = 7.7263 \times 10^{-3}$$

on $38 - 1 = 37$ degrees of freedom.

The between trucks sum of squares is

$$\frac{2.05^2}{10} + \frac{1.76^2}{8} + \frac{2.12^2}{11} + \frac{1.83^2}{9} - \frac{7.76^2}{38} = 3.4581 \times 10^{-3}$$

on $4 - 1 = 3$ degrees of freedom.

By subtraction, the residual sum of squares is

$$7.7263 \times 10^{-3} - 3.4581 \times 10^{-3} = 4.2682 \times 10^{-3}$$

on $37 - 3 = 34$ degrees of freedom.

The analysis of variance table is as follows:

Source of variation	Sum of squares	Degrees of freedom	Mean square	Variance ratio
Trucks	3.4581×10^{-3}	3	1.1527×10^{-3}	9.1824
Residual	4.2682×10^{-3}	34	0.1255×10^{-3}	
Total	7.7263×10^{-3}	37		

The upper 5% point of the $F_{3,34}$ distribution is approximately 2.9. The observed variance ratio is greater than this so we conclude that the result is significant at the 5% level and we reject the null hypothesis at this level. The evidence suggests that there are differences in the mean fuel consumption per tonne moved between the four trucks.

Two-Way Analysis of Variance

44.2



Introduction

In the one-way analysis of variance (Section 44.1) we consider the effect of one factor on the values taken by a variable. Very often, in engineering investigations, the effects of two or more factors are considered simultaneously.

The two-way ANOVA deals with the case where there are two factors. For example, we might compare the fuel consumptions of four car engines under three types of driving conditions (e.g. urban, rural, motorway). Sometimes we are interested in the effects of both factors. In other cases one of the factors is a 'nuisance factor' which is not of particular interest in itself but, if we allow for it in our analysis, we improve the power of our test for the other factor.

We can also allow for interaction effects between the two factors.



Prerequisites

Before starting this Section you should ...

- be familiar with the general techniques of hypothesis testing
- be familiar with the F -distribution
- be familiar with the one-way ANOVA calculations



Learning Outcomes

On completion you should be able to ...

- state the concepts and terminology of two-way ANOVA
- perform two-way ANOVA
- interpret the results of two-way ANOVA calculations

1. Two-way ANOVA without interaction

The previous Section considered a one-way classification analysis of variance, that is we looked at the variations induced by one set of values of a factor (or treatments as we called them) by partitioning the variation in the data into components representing 'between treatments' and 'within treatments.'

In this Section we will look at the analysis of variance involving two factors or, as we might say, two sets of treatments. In general terms, if we have two factors say A and B, there is no absolute reason to assume that there is no interaction between the factors. However, as an introduction to the two-way analysis of variance, we will consider the case occurring when there is no interaction between factors and an experiment is run only once. Note that some authors take the view that interaction may occur and that the residual sum of squares contains the effects of this interaction even though the analysis does not, at this stage, allow us to separate it out and check its possible effects on the experiment.

The following example builds on the previous example where we looked at the one-way analysis of variance.

Example of variance in data

In Section 44.1 we considered an example concerning four machines producing alloy spaces. This time we introduce an extra factor by considering both the machines producing the spacers and the performance of the operators working with the machines. In this experiment, the data appear as follows (spacer lengths in mm). Each operator made one spacer with each machine.

Operator	Machine 1	Machine 2	Machine 3	Machine 4
1	46	56	55	47
2	54	55	51	56
3	48	56	50	58
4	46	60	51	59
5	51	53	53	55

In a case such as this we are looking for discernible difference between the operators ('operator effects') on the one hand and the machines ('machine effects') on the other.

We suppose that the observation for operator i and machine j is taken from a normal distribution with mean

$$\mu_{ij} = \mu + \alpha_i + \beta_j$$

Here α_i is an operator effect and β_j is a machine effect. Our hypotheses may be stated as follows.

$$\begin{aligned} \text{Operator Effects} & \left\{ \begin{array}{l} H_0 : \mu_{1j} = \mu_{2j} = \mu_{3j} = \mu_{4j} = \mu_{5j} = \mu + \beta_j \\ \quad \text{That is } \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \alpha_5 = 0 \\ H_1 : \text{At least one of the operator effects is different to the others} \end{array} \right. \\ \text{Machine Effects} & \left\{ \begin{array}{l} H_0 : \mu_{i1} = \mu_{i2} = \mu_{i3} = \mu_{i4} = \mu + \alpha_i \\ \quad \text{That is } \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0 \\ H_1 : \text{At least one of the machine effects is different to the others} \end{array} \right. \end{aligned}$$

Note that the five operators and four machines give rise to data which has only one observation per 'cell.' For example, operator 2 using machine 3 produces a spacer 51 mm long, while operator 1 using machine 2 produces a spacer which is 56 mm long. Note also that in this example we have referred to the machines by number and not by letter. This is not particularly important but it will simplify

some of the notation used when we come to write out a general two-way ANOVA table shortly. We obtain one observation per cell and cannot measure variation within a cell. In this case we cannot check for interaction between the operator and the machine - the two factors used in this example. Running an experiment several times results in multiple observations per cell and in this case we should assume that there may be interaction between the factors and check for this. In the case considered here (no interaction between factors), the required sums of squares build easily on the relationship used in the one-way analysis of variance

$$SS_T = SS_{Tr} + SS_E$$

to become

$$SS_T = SS_A + SS_B + SS_E$$

where SS_A represent the sums of squares corresponding to factors A and B . In order to calculate the required sums of squares we lay out the table slightly more efficiently as follows.

Operator (j)	Machine (i)				Operator Means ($\bar{X}_{.j}$)	$(\bar{X}_{.j} - \bar{\bar{X}})$	Operator SS $(\bar{X}_{.j} - \bar{\bar{X}})^2$
	1	2	3	4			
1	46	56	55	47	51	-2	4
2	54	55	51	56	54	1	1
3	48	56	50	58	53	0	0
4	46	60	51	59	54	1	1
5	51	53	53	55	53	0	0
Machine Means ($\bar{X}_{i.}$)	49	56	52	55	$\bar{\bar{X}} = 53$	Sum = 0	$6 \times 4 = 24$

$(\bar{X}_{.j} - \bar{\bar{X}})$				
-4	3	-1	2	Sum = 0
Machine SS $(\bar{X}_{.j} - \bar{\bar{X}})^2$				
16	9	1	4	$30 \times 5 = 150$

Note 1

The \cdot notation means that summation takes place over that variable. For example, the five operator means $\bar{X}_{.j}$ are obtained as $\bar{X}_{.1} = \frac{46 + 56 + 55 + 47}{4} = 51$ and so on, while the four machine means $\bar{X}_{i.}$ are obtained as $\bar{X}_{1.} = \frac{46 + 54 + 48 + 46 + 51}{5} = 49$ and so on. Put more generally (and this is just an example)

$$\bar{X}_{.j} = \frac{\sum_{i=1}^m x_{ij}}{m}$$

Note 2

Multiplying factors were used in the calculation of the machine sum of squares (four in this case since there are four machines) and the operator sum of squares (five in this case since there are five operators).

Note 3

The two statements 'Sum = 0' are included purely as arithmetic checks.

We also know that $SS_O = 24$ and $SS_M = 150$.

Calculating the error sum of squares

Note that the total sum of squares is easy to obtain and that the error sum of squares is then obtained by straightforward subtraction.

The total sum of squares is given by summing the quantities $(X_{ij} - \bar{X})^2$ for the table of entries. Subtracting $\bar{X} = 53$ from each table member and squaring gives:

Operator (j)	Machine (i)			
	1	2	3	4
1	49	9	4	36
2	1	4	4	9
3	25	9	9	25
4	49	49	4	36
5	4	0	0	4

The total sum of squares is $SS_T = 330$.

The error sum of squares is given by the result

$$\begin{aligned}SS_E &= SS_T - SS_A - SS_B \\ &= 330 - 24 - 150 \\ &= 156\end{aligned}$$

At this stage we display the general two-way ANOVA table and then particularise the table for the example we are engaged in and draw conclusions by using the test as we have previously done with one-way ANOVA.

A General Two-Way ANOVA Table

Source of Variation	Sum of Squares SS	Degrees of Freedom	Mean Square MS	Value of F Ratio
<i>Between samples</i> (due to factor A) Differences between means \bar{X}_i and \bar{X}	$SS_A = b \sum_{i=1}^a (\bar{X}_i - \bar{X})^2$	$(a - 1)$	$MS_A = \frac{SS_A}{(a - 1)}$	$F = \frac{MS_A}{MS_E}$
<i>Between samples</i> (due to factor B) Differences between means \bar{X}_j and \bar{X}	$SS_B = a \sum_{j=1}^b (\bar{X}_j - \bar{X})^2$	$(b - 1)$	$MS_B = \frac{SS_B}{(b - 1)}$	$F = \frac{MS_B}{MS_E}$
<i>Within samples</i> (due to chance errors) Differences between individual observations and fitted values.	$SS_E = \sum_{i=1}^a \sum_{j=1}^b (X_{ij} - \bar{X}_i - \bar{X}_j + \bar{X})^2$	$(a - 1) \times (b - 1)$	$MS_E = \frac{SS_E}{(a - 1)(b - 1)}$	
Totals	$SS_T = \sum_{i=1}^a \sum_{j=1}^b (X_{ij} - \bar{X})^2$	$(ab - 1)$		

Hence the two-way ANOVA table for the example under consideration is

Source of Variation	Sum of Squares SS	Degrees of Freedom	Mean Square MS	Value of F Ratio
<i>Between samples</i> (due to factor A) Differences between means \bar{X}_i and \bar{X}	24	4	$\frac{24}{4} = 6$	$F = \frac{6}{13} = 0.46$
<i>Between samples</i> (due to factor B) Differences between means \bar{X}_j and \bar{X} height	150	3	$\frac{150}{3} = 50$	$F = \frac{50}{13} = 3.85$
<i>Within samples</i> (due to chance errors) Differences between individual observations and fitted values.	156	12	$\frac{156}{12} = 13$	
TOTALS	330	19		

From the F -tables (at the end of the Workbook) $F_{4,12} = 3.26$ and $F_{3,12} = 3.49$. Since $0.46 < 3.26$ we conclude that we **do not** have sufficient evidence to reject the null hypothesis that there is no difference between the operators. Since $3.85 > 3.49$ we conclude that we **do** have sufficient evidence at the 5% level of significance to reject the null hypothesis that there is no difference between the machines.



Key Point 2

If we have two factors, A and B , with a levels of factor A and b levels of factor B , and one observation per cell, we can calculate the sum of squares as follows.

The sum of squares for factor A is

$$SS_A = \frac{1}{b} \sum_{i=1}^a A_i^2 - \frac{G^2}{N} \quad \text{with } a - 1 \text{ degrees of freedom}$$

and the sum of squares for factor B is

$$SS_B = \frac{1}{a} \sum_{j=1}^b B_j^2 - \frac{G^2}{N} \quad \text{with } b - 1 \text{ degrees of freedom}$$

where

$$A_i = \sum_{j=1}^b X_{ij} \text{ is the total for level } i \text{ of factor } A,$$

$$B_j = \sum_{i=1}^a X_{ij} \text{ is the total for level } j \text{ of factor } B,$$

$$G = \sum_{i=1}^a \sum_{j=1}^b X_{ij} \text{ is the overall total of the data, and}$$

$$N = ab \text{ is the total number of observations.}$$

The total sum of squares is

$$SS_T = \sum_{i=1}^a \sum_{j=1}^b X_{ij}^2 - \frac{G^2}{N} \quad \text{with } N - 1 \text{ degrees of freedom}$$

The within-samples, or 'error', sum of squares can be found by subtraction. So

$$SS_E = SS_T - SS_A - SS_B$$

with

$$\begin{aligned} (N - 1) - (a - 1) - (b - 1) &= (ab - 1) - (a - 1) - (b - 1) \\ &= (a - 1)(b - 1) \text{ degrees of freedom} \end{aligned}$$



A vehicle manufacturer wishes to test the ability of three types of steel-alloy panels to resist corrosion when three different paint types are applied. Three panels with differing steel-alloy composition are coated with three types of paint. The following coded data represent the ability of the painted panels to resist weathering.

Paint Type	Steel-Alloy 1	Steel-Alloy 2	Steel-Alloy 3
1	40	51	56
2	54	55	50
3	47	56	50

Use a two-way ANOVA procedure to determine whether any difference in the ability of the panels to resist corrosion may be assigned to either the type of paint or the steel-alloy composition of the panels.

Your solution

Do your working on separate paper and enter the main conclusions here.

Answer

Our hypotheses may be stated as follows.

$$\text{Paint type } \begin{cases} H_0 : \mu_1 = \mu_2 = \mu_3 \\ H_1 : \text{At least one of the means is different from the others} \end{cases}$$

$$\text{Steel-Alloy } \begin{cases} H_0 : \mu_1 = \mu_2 = \mu_3 \\ H_1 : \text{At least one of the means is different from the others} \end{cases}$$

Following the methods of calculation outlined above we obtain:

Paint Type (j)	Steel-Alloy (i)			Paint Means (\bar{X}_j)	$(\bar{X}_j - \bar{\bar{X}})$	Paint SS ($\bar{X}_j - \bar{\bar{X}}$) ²
	1	2	3			
1	40	51	56	49	-2	4
2	54	55	50	53	2	4
3	47	54	52	51	0	0
Steel-Alloy Means (\bar{X}_i)	47	54	52	$\bar{\bar{X}} = 51$	Sum = 0	$8 \times 3 = 24$
$(\bar{X}_j - \bar{\bar{X}})$						
	-4	3	1	Sum = 0		
Steel-Alloy SS ($\bar{X}_j - \bar{\bar{X}}$)²						
	16	9	1	$26 \times 3 = 78$		

Hence $SS_{Pa} = 24$ and $SS_{St} = 78$. We now require SS_E . The calculations are as follows.

In the table below, the predicted outputs are given in parentheses.

Paint Type (j)	Machine (i)			Paint Means (\bar{X}_j)	$(\bar{X}_j - \bar{\bar{X}})$
	1	2	3		
1	40 (45)	51 (52)	56 (50)	49	-2
2	54 (49)	55 (56)	50 (54)	53	2
3	47 (47)	56 (54)	50 (52)	51	0
Steel-Alloy Means (\bar{X}_i)	47	54	52	$\bar{\bar{X}} = 51$	Sum = 0
$(\bar{X}_j - \bar{\bar{X}})$					
	-4	3	1	Sum = 0	

Answers continued

A table of squared residuals is easily obtained as

Paint (<i>j</i>)	Steel (<i>i</i>)		
	1	2	3
1	25	1	36
2	25	1	16
3	0	4	4

Hence the residual sum of squares is $SS_E = 112$. The total sum of squares is given by subtracting $\bar{X} = 51$ from each table member and squaring to obtain

Paint (<i>j</i>)	Steel (<i>i</i>)		
	1	2	3
1	121	0	25
2	9	16	1
3	16	25	1

The total sum of squares is $SS_T = 214$. We should now check to see that $SS_T = SS_{Pa} + SS_{St} + SS_E$. Substitution gives $214 = 24 + 78 + 112$ which is correct.

The values of F are calculated as shown in the ANOVA table below.

Source of Variation	Sum of Squares SS	Degrees of Freedom	Mean Square MS	Value of F Ratio
<i>Between samples</i> (due to treatment A, say, paint)	24	2	$MS_A = \frac{24}{2} = 12$	$F = \frac{12}{28}$ $= 0.429$
<i>Between samples</i> (due to treatment B, say, Steel – Alloy)	78	2	$MS_B = \frac{78}{2} = 39$	$F = \frac{39}{28}$ $= 1.393$
<i>Within samples</i> (due to chance errors)	112	4	$MS_E = \frac{112}{4} = 28$	
Totals	214	8		

From the F -tables the critical values of $F_{2,4} = 6.94$ and since both of the calculated F values are less than 6.94 we conclude that we do not have sufficient evidence to reject either null hypothesis.

2. Two-way ANOVA with interaction

The previous subsection looked at two-way ANOVA under the assumption that there was no interaction between the factors A and B . We will now look at the developments of two-way ANOVA to take into account possible interaction between the factors under consideration. The following analysis allows us to test to see whether we have sufficient evidence to reject the null hypothesis that the amount of interaction is effectively zero.

To see how we might consider interaction between factors A and B taking place, look at the following table which represents observations involving a two-factor experiment.

	Factor B				
Factor A	1	2	3	4	5
1	3	5	1	9	12
2	4	6	2	10	13
3	6	8	4	12	15

A brief inspection of the numbers in the five columns reveals that there is a constant difference between any two rows as we move from column to column. Similarly there is a constant difference between any two columns as we move from row to row. While the data are clearly contrived, it does illustrate that in this case that no interaction arises from variations in the differences between either rows or columns. Real data do not exhibit such behaviour in general of course, and we expect differences to occur and so we must check to see if the differences are large enough to provide sufficient evidence to reject the null hypothesis that the amount of interaction is effectively zero.

Notation

Let a represent the number of 'levels' present for factor A , denoted $i = 1, \dots, a$.

Let b represent the number of 'levels' present for factor B , denoted $j = 1, \dots, b$.

Let n represent the number of observations per cell. We assume that it is the same for each cell.

In the table above, $a = 3$, $b = 5$, $n = 1$. In the examples we shall consider, n will be greater than 1 and we will be able to check for interaction between the factors.

We suppose that the observations at level i of factor A and level j of factor B are taken from a normal distribution with mean μ_{ij} . When we assumed that there was no interaction, we used the **additive model**

$$\mu_{ij} = \mu + \alpha_i + \beta_j$$

So, for example, the difference $\mu_{i1} - \mu_{i2}$ between the means at levels 1 and 2 of factor B is equal to $\beta_1 - \beta_2$ and does not depend upon the level of factor A . When we allow interaction, this is not necessarily true and we write

$$\mu_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij}$$

Here γ_{ij} is an **interaction effect**. Now $\mu_{i1} - \mu_{i2} = \beta_1 - \beta_2 + \gamma_{i1} - \gamma_{i2}$ so the difference between two levels of factor B depends on the level of factor A .

Fixed and random effects

Often the levels assigned to a factor will be chosen deliberately. In this case the factors are said to be *fixed* and we have a **fixed effects model**. If the levels are chosen at random from a population of all possible levels, the factors are said to be *random* and we have a **random effects model**. Sometimes one factor may be *fixed* while one may be *random*. In this case we have a **mixed effects model**. In effect, we are asking whether we are interested in certain particular levels of a factor (fixed effects) or whether we just regard the levels as a sample and are interested in the population in general (random effects).

Calculation method

The data you will be working with will be set out in a manner similar to that shown below.

The table assumes n observations per cell and is shown along with a variety of totals and means which will be used in the calculations of the various test statistics to follow.

	Factor B						
Factor A	Level 1	Level 2	...	Level j	...	Level b	Totals
Level 1	x_{111} \vdots x_{11n}	x_{121} \vdots x_{12n}	...	x_{1j1} \vdots x_{1jn}	...	x_{1b1} \vdots x_{1bn}	$T_{1..}$
Level 2	x_{211} \vdots x_{21n}	x_{221} \vdots x_{22n}	...	x_{2j1} \vdots x_{2jn}	...	x_{2b1} \vdots x_{2bn}	$T_{2..}$
\vdots	\vdots	\vdots	\vdots	\vdots	...	\vdots	\vdots
Level i	x_{i11} \vdots x_{i1n}	Sum of data in cell (i,j) is $T_{ij} = \sum_{k=1}^n x_{ijk}$		x_{ij1} \vdots x_{ijn}	...	x_{ib1} \vdots x_{ibn}	$T_{i..}$
\vdots	\vdots	\vdots	\vdots	\vdots	...	\vdots	\vdots
Level a	x_{a11} \vdots x_{a1n}	x_{a21} \vdots x_{a2n}	...	x_{aj1} \vdots x_{ajn}	...	x_{ab1} \vdots x_{abn}	$T_{a..}$
Totals	$T_{.1}$	$T_{.2}$...	$T_{.j}$...	$T_{.b}$	$T_{...}$

Notes

(a) $T_{...}$ represents the grand total of the data values so that

$$T_{...} = \sum_{j=1}^b T_{.j} = \sum_{i=1}^a T_{i..} = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n x_{ijk}$$

(b) $T_{i..}$ represents the total of the data in the i th row.

(c) $T_{.j}$ represents the total of the data in the j th column.

(d) The total number of data entries is given by $N = nab$.

Partitioning the variation

We are now in a position to consider the partition of the total sum of the squared deviations from the overall mean which we estimate as

$$\bar{\bar{x}} = \frac{T_{...}}{N}$$

The total sum of the squared deviations is

$$\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (x_{ijk} - \bar{\bar{x}})^2$$

and it can be shown that this quantity can be written as

$$SS_T = SS_A + SS_B + SS_{AB} + SS_E$$

where SS_T is the total sum of squares given by

$$SS_T = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n x_{ijk}^2 - \frac{T_{...}^2}{N};$$

SS_A is the sum of squares due to variations caused by factor A given by

$$SS_A = \sum_{i=1}^a \frac{T_{i..}^2}{bn} - \frac{T_{...}^2}{N}$$

SS_B is the sum of squares due to variations caused by factor B given by

$$SS_B = \sum_{j=1}^b \frac{T_{.j.}^2}{an} - \frac{T_{...}^2}{N}$$

Note that bn means $b \times n$ which is the number of observations at each level of A and an means $a \times n$ which is the number of observations at each level of B .

SS_{AB} is the sum of the squares due to variations caused by the interaction of factors A and B and is given by

$$SS_{AB} = \sum_{i=1}^a \sum_{j=1}^b \frac{T_{ij.}^2}{n} - \frac{T_{...}^2}{N} - SS_A - SS_B.$$

Note that the quantity $T_{ij.} = \sum_{k=1}^n x_{ijk}$ is the sum of the data in the $(i, j)^{\text{th}}$ cell and that the quantity

$$\sum_{i=1}^a \sum_{j=1}^b \frac{T_{ij.}^2}{n} - \frac{T_{...}^2}{N}$$

is the sum of the squares between cells.

SS_E is the sum of the squares due to chance or experimental error and is given by

$$SS_E = SS_T - SS_A - SS_B - SS_{AB}$$

The number of degrees of freedom ($N - 1$) is partitioned as follows:

SS_T	SS_A	SS_B	SS_{AB}	SS_E
$N - 1$	$a - 1$	$b - 1$	$(a - 1)(b - 1)$	$N - ab$

Note that there are $ab - 1$ degrees of freedom between cells and that the number of degrees of freedom for SS_{AB} is given by

$$ab - 1 - (a - 1) - (b - 1) = (a - 1)(b - 1)$$

This gives rise to the following two-way ANOVA tables.

Two-Way ANOVA Table - Fixed-Effects Model

Source of Variation	Sum of squares SS	Degrees of Freedom	Mean Square MS	Value of F Ratio
Factor A	SS_A	$(a - 1)$	$MS_A = \frac{SS_A}{(a - 1)}$	$F = \frac{MS_A}{MS_E}$
Factor B	SS_B	$(b - 1)$	$MS_B = \frac{SS_B}{(b - 1)}$	$F = \frac{MS_B}{MS_E}$
Interaction	SS_{AB}	$(a - 1) \times (b - 1)$	$MS_{AB} = \frac{SS_{AB}}{(a - 1)(b - 1)}$	$F = \frac{MS_{AB}}{MS_E}$
Residual Error	SS_E	$(N - ab)$	$MS_E = \frac{SS_E}{N - ab}$	
Totals	SS_T	$(N - 1)$		

Two-Way ANOVA Table - Random-Effects Model

Source of Variation	Sum of squares SS	Degrees of Freedom	Mean Square MS	Value of F Ratio
Factor A	SS_A	$(a - 1)$	$MS_A = \frac{SS_A}{(a - 1)}$	$F = \frac{MS_A}{MS_{AB}}$
Factor B	SS_B	$(b - 1)$	$MS_B = \frac{SS_B}{(b - 1)}$	$F = \frac{MS_B}{MS_{AB}}$
Interaction	SS_{AB}	$(a - 1) \times (b - 1)$	$MS_{AB} = \frac{SS_{AB}}{(a - 1)(b - 1)}$	$F = \frac{MS_{AB}}{MS_E}$
Residual Error	SS_E	$(N - ab)$	$MS_E = \frac{SS_E}{N - ab}$	
Totals	SS_T	$(N - 1)$		

Two-Way ANOVA Table - Mixed-Effects Model

Case (i) A fixed and B random.

Source of Variation	Sum of squares SS	Degrees of Freedom	Mean Square MS	Value of F Ratio
Factor A	SS_A	$(a - 1)$	$MS_A = \frac{SS_A}{(a - 1)}$	$F = \frac{MS_A}{MS_{AB}}$
Factor B	SS_B	$(b - 1)$	$MS_B = \frac{SS_B}{(b - 1)}$	$F = \frac{MS_B}{MS_E}$
Interaction	SS_{AB}	$(a - 1) \times (b - 1)$	$MS_{AB} = \frac{SS_{AB}}{(a - 1)(b - 1)}$	$F = \frac{MS_{AB}}{MS_E}$
Residual Error	SS_E	$(N - ab)$	$MS_E = \frac{SS_E}{N - ab}$	
Totals	SS_T	$(N - 1)$		

Case (ii) A random and B fixed.

Source of Variation	Sum of squares SS	Degrees of Freedom	Mean Square MS	Value of F Ratio
Factor A	SS_A	$(a - 1)$	$MS_A = \frac{SS_A}{(a - 1)}$	$F = \frac{MS_A}{MS_E}$
Factor B	SS_B	$(b - 1)$	$MS_B = \frac{SS_B}{(b - 1)}$	$F = \frac{MS_B}{MS_{AB}}$
Interaction	SS_{AB}	$(a - 1) \times (b - 1)$	$MS_{AB} = \frac{SS_{AB}}{(a - 1)(b - 1)}$	$F = \frac{MS_{AB}}{MS_E}$
Residual Error	SS_E	$(N - ab)$	$MS_E = \frac{SS_E}{N - ab}$	
Totals	SS_T	$(N - 1)$		



Example 1

In an experiment to compare the effects of weathering on paint of three different types, two identical surfaces coated with each type of paint were exposed in each of four environments. Measurements of the degree of deterioration were made as follows.

	Environment 1		Environment 2		Environment 3		Environment 4	
Paint A	10.89	10.74	9.94	11.25	9.88	10.13	14.11	12.84
Paint B	12.28	13.11	14.45	11.17	11.29	11.10	13.44	11.37
Paint C	10.68	10.30	10.89	10.97	10.61	11.00	12.22	11.32

Making the assumptions of normality, independence and equal variance, derive the appropriate ANOVA tables and state the conclusions which may be drawn at the 5% level of significance in the following cases.

- The types of paint and the environments are chosen deliberately because the interest is in these paints and these environments.
- The types of paint are chosen deliberately because the interest is in these paints but the environments are regarded as a sample of possible environments.
- The types of paint are regarded as a random sample of possible paints and the environments are regarded as a sample of possible environments.

Solution

We know that case (a) is described as a fixed-effects model, case (b) is described as a mixed-effects model (paint type fixed) and case (c) is described as a random-effects model. In all three cases the calculations necessary to find MS_P (paints), MS_N (environments), MS_P and MS_N are identical. Only the calculation and interpretation of the test statistics will be different. The calculations are shown below.

Subtracting 10 from each observation, the data become:

	Environment 1		Environment 2		Environment 3		Environment 4		Total
Paint A	0.89	0.74	-0.06	1.25	-0.12	0.13	4.11	2.84	9.78
	(total 1.63)		(total 1.19)		(total 0.01)		(total 6.95)		
Paint B	2.28	3.11	4.45	1.17	1.29	1.10	3.44	1.37	18.21
	(total 5.39)		(total 5.62)		(total 2.39)		(total 4.81)		
Paint C	0.68	0.30	0.89	0.97	0.61	1.00	2.22	1.32	7.99
	(total 0.98)		(total 1.86)		(total 1.61)		(total 3.54)		
Total	8.00		8.67		4.01		15.30		35.98

The total sum of squares is

$$SS_T = 0.89^2 + 0.74^2 + \dots + 1.32^2 - \frac{35.98^2}{24} = 36.910$$

We can simplify the calculation by finding the between samples sum of squares

$$SS_S = \frac{1}{2}(1.63^2 + 5.39^2 + \dots + 3.54^2) - \frac{35.98^2}{24} = 26.762$$

Solution (contd.)

Sum of squares for paints is

$$SS_P = \frac{1}{8}(9.78^2 + 18.15^2 + 7.99^2) - \frac{35.98^2}{24} = 7.447$$

Sum of squares for environments is

$$SS_N = \frac{1}{6}(8.00^2 + 8.67^2 + 3.98^2 + 15.30^2) - \frac{35.98^2}{24} = 10.950$$

So the interaction sum of squares is $SS_{PN} = SS_S - SS_P - SS_N = 8.365$ and

the residual sum of squares is $SS_E = SS_T - SS_S = 10.148$ The results are combined in the following ANOVA table

	Deg. of Freedom	Sum of Squares	Mean Square	Variance Ratio (fixed)	Variance Ratio (mixed)	Variance Ratio (random)
Paints	2	7.447	3.724	4.40 $F_{2,12} = 3.89$	2.67 $F_{2,6} = 5.14$	2.67 $F_{2,6} = 5.14$
Environments	3	10.950	3.650	4.31 $F_{3,12} = 3.49$	4.31 $F_{3,12} = 3.49$	2.61 $F_{3,6} = 4.76$
Interaction	6	8.365	1.394	1.65 $F_{6,12} = 3.00$	1.65 $F_{6,12} = 3.00$	1.65 $F_{6,12} = 3.00$
Treatment combinations	11	26.762	2.433			
Residual	12	10.148	0.846			
Total	23	36.910				

The following conclusions may be drawn. There is insufficient evidence to support the interaction hypothesis in any case. Therefore we can look at the tests for the main effects.

Case (a) Since $4.40 > 3.89$ we have sufficient evidence to conclude that paint type affects the degree of deterioration. Since $4.07 > 3.49$ we have sufficient evidence to conclude that environment affects the degree of deterioration.

Case (b) Since $2.67 < 5.14$ we do not have sufficient evidence to reject the hypothesis that paint type has no effect on the degree of deterioration. Since $4.07 > 3.49$ we have sufficient evidence to conclude that environment affects the degree of deterioration.

Case (c) Since $2.67 < 5.14$ we do not have sufficient evidence to reject the hypothesis that paint type has no effect on the degree of deterioration. Since $2.61 < 4.76$ we do not have sufficient evidence to reject the hypothesis that environment has no effect on the degree of deterioration.

If the test for interaction had given a significant result then we would have concluded that there was an interaction effect. Therefore the differences between the average degree of deterioration for different paint types would have depended on the environment and there might have been no overall 'best paint type'. We would have needed to compare combinations of paint types and environments. However the relative sizes of the mean squares would have helped to indicate which effects were most important.



A motor company wishes to check the influences of tyre type and shock absorber settings on the roadholding of one of its cars. Two types of tyre are selected from the tyre manufacturer who normally provides tyres for the company's new vehicles. A shock absorber with three possible settings is chosen from a range of shock absorbers deemed to be suitable for the car. An experiment is conducted by conducting roadholding tests using each tyre type and shock absorber setting. The (coded) data resulting from the experiment are given below.

Factor Tyre	Shock Absorber Setting		
	B1=Comfort	B2=Normal	B3=Sport
Type A1	5	8	6
	6	5	9
	8	3	12
Type A2	9	10	12
	7	9	10
	7	8	9

Decide whether an appropriate model has random-effects, mixed-effects or fixed-effects and derive the appropriate ANOVA table. State clearly any conclusions that may be drawn at the 5% level of significance.

Your solution

Do the calculations on separate paper and use the space here and on the following page for your summary and conclusions.

Answer

We know that both the tyres and the shock absorbers are not chosen at random from populations consisting of all possible tyre types and shock absorber types so that their influence is described by a fixed-effects model. The calculations necessary to find MS_A, MS_B, MS_{AB} and MS_E are shown below.

	B1	B2	B3	Totals
A1	5	8	6	
	6	5	9	
	8	3	12	
	$T_{11} = 19$	$T_{12} = 16$	$T_{13} = 27$	$T_{1..} = 62$
A2	9	10	12	
	7	9	10	
	7	8	9	
	$T_{21} = 23$	$T_{22} = 27$	$T_{23} = 31$	$T_{2..} = 81$
Totals	$T_{.1.} = 42$	$T_{.2.} = 43$	$T_{.3.} = 58$	$T_{...} = 143$

The sums of squares calculations are:

$$SS_T = \sum_{i=1}^2 \sum_{j=1}^3 \sum_{k=1}^3 x_{ijk}^2 - \frac{T_{...}^2}{N} = 5^2 + 6^2 + \dots + 10^2 + 9^2 - \frac{143^2}{18} = 1233 - \frac{143^2}{18} = 96.944$$

$$SS_A = \sum_{i=1}^2 \frac{T_{i..}^2}{bn} - \frac{T_{...}^2}{N} = \frac{62^2 + 81^2}{3 \times 3} - \frac{143^2}{18} = \frac{10405}{9} - \frac{143^2}{18} = 20.056$$

$$SS_B = \sum_{j=1}^3 \frac{T_{.j.}^2}{an} - \frac{T_{...}^2}{N} = \frac{42^2 + 43^2 + 58^2}{2 \times 3} - \frac{143^2}{18} = \frac{6977}{6} - \frac{143^2}{18} = 26.778$$

$$SS_{AB} = \sum_{i=1}^2 \sum_{j=1}^3 \frac{T_{ij.}^2}{n} - \frac{T_{...}^2}{N} - SS_A - SS_B = \frac{19^2 + \dots + 31^2}{3} - \frac{143^2}{18} - 20.056 - 26.778$$

$$= \frac{3565}{3} - \frac{143^2}{18} - 20.056 - 26.778 = 5.444$$

$$SS_E = SS_T - SS_A - SS_B - SS_{AB} = 96.944 - 20.056 - 26.778 - 5.444 = 44.666$$

The results are combined in the following ANOVA table.

Source	SS	DoF	MS	F (Fixed)	F (Fixed)
Factor A	20.056	1	20.056	$\frac{MS_A}{MS_E}$	5.39 $F_{1,12} = 4.75$
Factor B	26.778	2	13.389	$\frac{MS_B}{MS_E}$	3.60 $F_{2,12} = 3.89$
Interaction AB	5.444	2	2.722	$\frac{MS_{AB}}{MS_E}$	0.731 $F_{2,12} = 3.89$
Residual E	44.666	12	3.722		
Totals	96.944	17			

Answer

The following conclusions may be drawn:

Interaction: There is insufficient evidence to support the hypothesis that interaction takes place between the factors.

Factor *A*: Since $5.39 > 4.75$ we have sufficient evidence to reject the hypothesis that tyre type does not affect the roadholding of the car.

Factor *B*: Since $3.60 < 3.89$ we do not have sufficient evidence to reject the hypothesis that shock absorber settings do not affect the roadholding of the car.



The variability of a measured characteristic of an electronic assembly is a source of trouble for a manufacturer with global manufacturing and sales facilities. To investigate the possible influences of assembly machines and testing stations on the characteristic, an engineer chooses three testing stations and three assembly machines from the large number of stations and machines in the possession of the company. For each testing station - assembly machine combination, three observations of the characteristic are made.

The (coded) data resulting from the experiment are given below.

Factor Assembly Machine	Testing Station		
	B1	B2	B3
A1	2.3	3.7	3.1
	3.4	2.8	3.2
	3.5	3.7	3.5
A2	3.5	3.9	3.3
	2.6	3.9	3.4
	3.6	3.4	3.5
A3	2.4	3.5	2.6
	2.7	3.2	2.6
	2.8	3.5	2.5

Decide whether an appropriate model has random-effects, mixed-effects or fixed-effects and derive the appropriate ANOVA table.

State clearly any conclusions that may be drawn at the 5% level of significance.

Your solution

Do the calculations on separate paper and use the space here and on the following page for your summary and conclusions.

Your solution contd.

Answer

Both the machines and the testing stations are effectively chosen at random from populations consisting of all possible types so that their influence is described by a random-effects model. The calculations necessary to find MS_A , MS_B , MS_{AB} and MS_E are shown below.

	B1	B2	B3	Totals
A1	2.3	3.7	3.1	
	3.4	2.8	3.2	
	3.5	3.7	3.5	
	$T_{11} = 9.2$	$T_{12} = 10.2$	$T_{13} = 9.8$	$T_{1..} = 29.2$
A2	3.5	3.9	3.3	
	2.6	3.9	3.4	
	3.6	3.4	3.5	
	$T_{21} = 9.7$	$T_{22} = 11.2$	$T_{23} = 10.2$	$T_{2..} = 31.1$
A3	2.4	3.5	2.6	
	2.7	3.2	2.6	
	2.8	3.5	2.5	
	$T_{31} = 7.9$	$T_{32} = 10.2$	$T_{33} = 7.7$	$T_{3..} = 25.8$
Totals	$T_{.1.} = 26.8$	$T_{.2.} = 31.6$	$T_{.3.} = 27.7$	$T_{...} = 86.1$

$a = 3$, $b = 3$, $n = 3$, $N = 27$ and the sums of squares calculations are:

$$SS_T = \sum_{i=1}^3 \sum_{j=1}^3 \sum_{k=1}^3 x_{ijk}^2 - \frac{T_{...}^2}{N} = 2.3^2 + 3.4^2 + \dots + 2.6^2 + 2.5^2 - \frac{86.1^2}{27} = 5.907$$

$$SS_A = \sum_{i=1}^3 \frac{T_{i..}^2}{bn} - \frac{T_{...}^2}{N} = \frac{29.2^2 + 31.1^2 + 25.8^2}{3 \times 3} - \frac{86.1^2}{27} = 1.602$$

$$SS_B = \sum_{j=1}^3 \frac{T_{.j.}^2}{an} - \frac{T_{...}^2}{N} = \frac{26.8^2 + 31.6^2 + 27.7^2}{3 \times 3} - \frac{86.1^2}{27} = 1.447$$

$$SS_{AB} = \sum_{i=1}^3 \sum_{j=1}^3 \frac{T_{ij.}^2}{n} - \frac{T_{...}^2}{N} - SS_A - SS_B$$

$$= \frac{9.2^2 + 10.2^2 + \dots + 10.2^2 + 7.7^2}{3} - \frac{86.1^2}{27} - 1.602 - 1.447 = 0.398$$

$$SS_E = SS_T - SS_A - SS_B - SS_{AB} = 5.907 - 1.602 - 1.447 - 0.398 = 2.46$$

Answer continued

The results are combined in the following ANOVA table

Source	<i>SS</i>	<i>DoF</i>	<i>MS</i>	<i>F</i> (Random)	<i>F</i> (Random)
Factor A (Machines)	1.602	2	0.801	$\frac{MS_A}{MS_{AB}}$	8.05 $F_{2,4} = 6.94$
Factor B (Stations)	1.447	2	0.724	$\frac{MS_B}{MS_{AB}}$	7.28 $F_{2,4} = 6.94$
Interaction AB	0.398	4	0.099(5)	$\frac{MS_{AB}}{MS_E}$	0.728 $F_{4,18} = 2.93$
Residual E	2.460	18	0.136		
Totals	5.907	26			

The following conclusions may be drawn.

Interaction: There is insufficient evidence to support the hypothesis that interaction takes place between the factors.

Factor *A*: Since $8.05 > 6.94$ we have sufficient evidence to reject the hypothesis that the assembly machines do not affect the assembly characteristic.

Factor *B*: Since $7.28 > 6.94$ we have sufficient evidence to reject the hypothesis that the choice of testing station does not affect the assembly characteristic.

3. Two-way ANOVA versus one-way ANOVA

You should note that a two-way ANOVA design is rather more efficient than a one-way design. In the last example, we could fix the testing station and look at the electronic assemblies produced by a variety of machines. We would have to replicate such an experiment for *every* testing station. It would be very difficult (impossible!) to exactly duplicate the same conditions for all of the experiments. This implies that the consequent experimental error could be very large. Remember also that in a one-way design we cannot check for interaction between the factors involved in the experiment. The three main advantages of a two-way ANOVA may be stated as follows:

- It is possible to simultaneously test the effects of two factors. This saves both time and money.
- It is possible to determine the level of interaction present between the factors involved.
- The effect of one factor can be investigated over a variety of levels of another and so any conclusions reached may be applicable over a range of situations rather than a single situation.

Exercises

- The temperatures, in Celsius, at three locations in the engine of a vehicle are measured after each of five test runs. The data are as follows. Making the usual assumptions for a two-way analysis of variance without replication, test the hypothesis that there is no systematic difference in temperatures between the three locations. Use the 5% level of significance.

Location	Run 1	Run 2	Run 3	Run 4	Run 5
A	72.8	77.3	82.9	69.4	74.6
B	71.5	72.4	80.7	67.0	74.0
C	70.8	74.0	79.1	69.0	75.4

- Waste cooling water from a large engineering works is filtered before being released into the environment. Three separate discharge pipes are used, each with its own filter. Five samples of water are taken on each of four days from each of the three discharge pipes and the concentrations of a pollutant, in parts per million, are measured. The data are given below. Analyse the data to test for differences between the discharge pipes. Allow for effects due to pipes and days and for an interaction effect. Treat the pipe effects as fixed and the day effects as random. Use the 5% level of significance.

Day	Pipe A				
1	160	181	163	173	178
2	175	170	219	166	171
3	169	186	179	178	183
4	230	206	216	195	250
Day	Pipe B				
1	172	164	186	185	172
2	177	170	156	140	155
3	193	194	189	156	181
4	212	235	195	206	209
Day	Pipe C				
1	214	196	207	219	200
2	186	184	181	189	179
3	209	220	199	185	228
4	254	293	283	262	259

Answers

1. We calculate totals as follows.

Run	Total	Location	Total
1	215.1	A	377.0
2	223.7	B	365.6
3	242.7	C	368.3
4	205.4	Total	1110.9
5	224.0		
Total	1110.9		

$$\sum \sum y_{ij}^2 = 82552.17$$

The total sum of squares is

$$82552.17 - \frac{1110.9^2}{15} = 278.916 \quad \text{on } 15 - 1 = 14 \text{ degrees of freedom.}$$

The between-runs sum of squares is

$$\frac{1}{3}(215.1^2 + 223.7^2 + 242.7^2 + 205.4^2 + 224.0^2) - \frac{1110.9^2}{15} = 252.796$$

on $5 - 1 = 4$ degrees of freedom.

The between-locations sum of squares is

$$\frac{1}{5}(377.0^2 + 365.6^2 + 368.3^2) - \frac{1110.9^2}{15} = 14.196 \quad \text{on } 3 - 1 = 2 \text{ degrees of freedom.}$$

By subtraction, the residual sum of squares is

$$278.916 - 252.796 - 14.196 = 11.924 \quad \text{on } 14 - 4 - 2 = 8 \text{ degrees of freedom.}$$

The analysis of variance table is as follows.

Source of variation	Sum of squares	Degrees of freedom	Mean square	Variance ratio
Runs	252.796	4	63.199	
Locations	14.196	2	7.098	4.762
Residual	11.924	8	1.491	
Total	278.916	14		

The upper 5% point of the $F_{2,8}$ distribution is 4.46. The observed variance ratio is greater than this so we conclude that the result is significant at the 5% level and reject the null hypothesis at this level. The evidence suggests that there are systematic differences between the temperatures at the three locations. Note that the Runs mean square is large compared to the Residual mean square showing that it was useful to allow for differences between runs.

Answers continued

2. We calculate totals as follows.

	Day 1	Day 2	Day 3	Day 4	Total
Pipe A	855	901	895	1097	3748
Pipe B	879	798	913	1057	3647
Pipe C	1036	919	1041	1351	4347
Total	2770	2618	2849	3505	11742

$$\sum \sum \sum y_{ijk}^2 = 2356870$$

The total number of observations is $N = 60$.

The total sum of squares is

$$2356870 - \frac{11742^2}{60} = 58960.6$$

on $60 - 1 = 59$ degrees of freedom.

The between-cells sum of squares is

$$\frac{1}{5}(855^2 + \dots + 1351^2) - \frac{11742^2}{60} = 58960.6$$

on $12 - 1 = 11$ degrees of freedom, where by "cell" we mean the combination of a pipe and a day.

By subtraction, the residual sum of squares is

$$58960.6 - 48943.0 = 10017.6$$

on $59 - 11 = 48$ degrees of freedom.

The between-days sum of squares is

$$\frac{1}{15}(2770^2 + 2618^2 + 2849^2 + 3505^2) - \frac{11742^2}{60} = 30667.3$$

on $4 - 1 = 3$ degrees of freedom.

The between-pipes sum of squares is

$$\frac{1}{20}(3748^2 + 3647^2 + 4347^2) - \frac{11742^2}{60} = 14316.7$$

on $3 - 1 = 2$ degrees of freedom.

By subtraction the interaction sum of squares is

$$48943.0 - 30667.3 - 14316.7 = 3959.0$$

on $11 - 3 - 2 = 6$ degrees of freedom.

Answers continued

The analysis of variance table is as follows.

Source of variation	Sum of squares	Degrees of freedom	Mean square	Variance ratio
Pipes	14316.7	2	7158.4	10.85
Days	30667.3	3	10222.4	48.98
Interaction	3959.0	6	659.8	3.16
Cells	48943.0	11	4449.4	21.32
Residual	10017.6	48	208.7	
Total	58960.6	59		

Notice that, because Days are treated as a random effect, we divide the Pipes mean square by the Interaction mean square rather than by the Residual mean square.

The upper 5% point of the $F_{6,48}$ distribution is approximately 2.3. Thus the Interaction variance ratio is significant at the 5% level and we reject the null hypothesis of no interaction. We must therefore conclude that there are differences between the means for pipes and for days and that the difference between one pipe and another varies from day to day. Looking at the mean squares, however, we see that both the Pipes and Days mean squares are much bigger than the Interaction mean square. Therefore it seems that the interaction effect is relatively small compared to the differences between days and between pipes.

Experimental Design

44.3

Introduction

In Sections 44.1 and 44.2 we have considered how to analyse data from experiments of certain types. Of course, before we can analyse any data we must conduct the experiment and before we can conduct an experiment we must design it. The work of applying statistical ideas to engineering experiments does not begin with the analysis of data. It begins with the design of the experiment. It is important to give proper consideration to experimental design to make sure that our experiment is efficient and that it will, in fact, give us the information we require. A badly designed experiment may give poor or misleading results or may turn out to be an expensive waste of time and money.

Prerequisites

Before starting this Section you should ...

- explain the concepts and terminology of the one-way and two-way ANOVA
- be familiar with the F -distribution
- understand the general techniques of hypothesis testing

Learning Outcomes

On completion you should be able to ...

- explain the basic concepts of experimental design
- apply randomized blocks and Latin Square designs
- analyse the results from randomized blocks and Latin Square designs

1. Experimental design

So far in this Workbook we have looked at some of the statistical methods used in the analysis and interpretation of experimental results. There are occasions when the planning of an experiment is not in the control of the statistician responsible for analysing the results. It is always preferable to have some idea of the likely variability of the data so that any experimental design can take this into account. For this reason, the **design of experiments** is of crucial importance if weight is to be given to the results obtained. Usually, the experimenter will have to take into account:

- (a) The definition of the problem to be investigated. This would usually include the selection of the response variable to be measured and the factors or treatments influencing the response. Remember that the factors may be quantitative (such as temperature, pressure or force), qualitative (such as days of the week, machine operators or machines themselves) and decisions must be taken as to whether these factors are *fixed* or *random* and at what levels they are to be used.
- (b) The sample size. Clearly the experimenter should determine the number of observations to be taken and the random order in which the experiments are to occur in order that the effects of uncontrollable or unforeseen variables are minimized.
- (c) Data collection. Decisions need to be taken as to how the data are to be collected and tabulated. The calculation of the test statistics needs to be taken into account as does the level of acceptance or rejection of any hypotheses used.

We have already looked at some introductory ANOVA situations and we now turn our attention to so-called **block designs** used in the conduct of experiments.

Block design

Block design, or more specifically, **randomized block design** enables an unbiased estimate of error to be obtained and ensures that the error obtained is a minimum.

As an illustration, imagine that we wish to compare four extrusion processes and measure their effect on the brittleness of copper wire produced. Assume further that the copper from which the wire is made is delivered in quantities to allow only four tests per batch. We will refer to such a batch as a **block**. If we replicate each treatment four times, we could organize the four members of block 1 all to receive treatment *A*, all four members of block 2 all to receive treatment *B* and so on. This situation could be represented as shown below.

Block	Treatment			
1	A	A	A	A
2	B	B	B	B
3	C	C	C	C
4	D	D	D	D

Giving every member of a given block the same treatment is not a very sensible thing to do. The reason for this is that any observed differences might well be due to differences between the blocks and not differences between the treatments. Remember that each block consists of a batch of material and so as engineers we would expect some variation in batches of materials delivered for processing. We can avoid this pitfall by ensuring that the blocks are distributed among the treatments. Ensuring that each block experiences every treatment could lead to the situation represented below.

Block	Treatment			
1	A	B	C	D
2	A	B	C	D
3	A	B	C	D
4	A	B	C	D

On the face of it, we appear to have a sensible way of organizing the treatments between blocks. However it is entirely possible that other variables might come into play which might change with time which would bias the observations made. For example, if the treatments are always applied in a particular order, say A, B, C and then D it could be that the state of the extrusion machines might change with time which would bias the results. For example, if treatment A is always applied first, it could be that the extrusion machine is not fully 'warmed up' and so non-typical results might occur.

In order to remove the bias referred to above, we could randomize the order in which the treatments are applied. This might result in the **randomized block design** represented below in which comparisons are made between sets of treatments applied to fairly homogeneous material.

Block	Treatment			
1	A	B	D	C
2	B	D	A	C
3	C	B	D	A
4	A	D	B	C

Note that this design may be taken to consist of two independent variables, namely blocks and treatments. The total sum of the squares (previously referred to as SS_T) may be partitioned as

$$SS_T = SS_{\text{Blocks}} + SS_{\text{Treatments}} + SS_{\text{Error}}$$

We illustrate this in Example 2.



Example 2

The compressive strength of concrete to be used in the construction of a dam may depend on the proportion of a particular component of the mix. In order to investigate this, the compressive strength of concrete containing a variety of percentages of the component, 2%, 4%,.....10% was measured. The concrete was made using four batches of cement. One part of each batch was used with each of the percentages of the component of interest. The resulting concrete samples were then subjected to tests to determine their compressive strength. The data obtained are given below.

Component %	Compressive Strength			
	Block (Batch)			
	1	2	3	4
2	201	185	182	179
4	200	195	220	199
6	257	240	224	225
8	252	228	275	250
10	280	275	277	260

What conclusions may be drawn about the effect of the mix component on the compressive strength of the resulting concrete? Is there evidence of differences between the batches of cement. Use the 5% level of significance.

Solution

Assuming that no interaction takes place between the blocks and the parts, we can calculate the appropriate sums of squares as done previously. The grand total of the observations is represented by T .

Block	Compressive Strength				Totals
	Part				
Component %	1	2	3	4	
2	201	185	182	179	747
4	200	195	220	199	814
6	257	240	224	225	946
8	252	228	275	250	1005
10	280	275	277	260	1092
Totals	1190	1123	1178	1113	$T = 4604$

The total sum of squares is given by

$$SS_T = 201^2 + \dots + 182^2 + \dots + 260^2 - \frac{4604^2}{5 \times 4} = 1082234 - 1059840.8 = 22393.2$$

The sum of squares for the treatments is given by

$$SS_{Tr} = \frac{1190^2 + \dots + 1113^2}{5} - \frac{4604^2}{5 \times 4} = 1060736.4 - 1059840.8 = 895.6$$

Solution (contd.)

The sum of squares for the blocks is given by

$$SS_{Bl} = \frac{747^2 + \dots + 1092^2}{4} - \frac{4604^2}{5 \times 4} = 107950.5 - 1059840.8 = 19661.7$$

The sum of squares for the errors is given by

$$SS_E = SS_T - SS_{Tr} - SS_{Bl} = 22393.2 - 895.6 - 19661.7 = 1835.9$$

These calculations give rise to the following ANOVA table

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Squares	Value of F
Blocks	895.6	3	298.53	$F_{Tr} = \frac{SS_{Tr}}{SS_E} = 1.95$
Treatments	19661.7	4	4915.43	$F_{Bl} = \frac{SS_{Bl}}{SS_E} = 32.13$
Error	1835.9	12	152.99	
Total	22393.2	19		

Conclusions

(a) **Blocks**

From the F -tables, $F_{3,12} = 3.49$ and since $1.95 < 3.49$ we have no evidence to reject the null hypothesis that there are no differences between the batches of cement.

(b) **Treatments**

From the F -tables, $F_{4,12} = 3.26$ and since $32.13 > 3.26$ we have sufficient evidence to reject the null hypothesis that the addition of the mix component has no effect on the compressive strength of the resulting concrete. Hence our conclusion (at the 5% level of significance) is that the addition of the mix component does affect the compressive strength of the resulting concrete.



The tensile strength of aluminium alloy tubing to be used in the construction of aircraft is known to depend on the extrusion process by which the tubing is produced. In order to investigate the tensile strength of the alloy made by four different extrusion processes, four large samples were made using each extrusion process. One sample from each process was sent to each of four laboratories for measurement of its tensile strength. The data obtained are given below.

	Block (lab)			
Process	1	2	3	4
1	202	186	182	180
2	200	195	224	199
3	259	243	225	223
4	252	227	276	251

What conclusions may be drawn about the effect of the extrusion processes on the tensile strength of the resulting tubing? Is there sufficient evidence to suggest that there are systematic differences between measurements from the different laboratories? Use the 5% level of significance.

Your solution

Answers

Assuming that no interaction takes place between the factors, we can calculate the appropriate sums of squares, representing the grand total by T .

Process	Blocks (lab)				Totals
	1	2	3	4	
1	202	186	182	180	750
2	200	195	224	199	818
3	259	243	225	223	950
4	252	227	276	251	1006
Totals	913	851	907	853	$T = 3524$

The total sum of squares is given by

$$SS_T = 202^2 + \dots + 182^2 + \dots + 251^2 - \frac{3524^2}{4 \times 4} = 789420 - 776161 = 13259$$

The sum of squares for the treatments is given by

$$SS_{Tr} = \frac{913^2 + \dots + 853^2}{5} - \frac{3524^2}{4 \times 4} = 3102808 - 776161 = 846$$

The sum of the squares for the blocks is given by

$$SS_{Bl} = \frac{750^2 + \dots + 1006^2}{4} - \frac{3524^2}{4 \times 4} = 786540 - 776161 = 10379$$

The sum of squares for the errors is given by

$$SS_E = SS_T - SS_{Tr} - SS_{Bl} = 13259 - 846 - 10379 = 2034$$

These calculations give rise to the following ANOVA table

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Squares	Value of F
Blocks (labs)	846	3	282	$F_{Tr} = \frac{SS_{Tr}}{SS_E} = 1.25$
Treatments	10379	3	3459.67	$F_{Bl} = \frac{SS_{Bl}}{SS_E} = 15.31$
Error	2034	9	226	
Total	13259	15		

Conclusions

(a) Blocks

From F -tables, $F_{3,9} = 3.86$ and since $1.25 < 3.86$ we have no evidence to reject the null hypothesis that there are no systematic differences between measurements from the different laboratories.

(b) Treatments

From the F -tables, $F_{3,9} = 3.86$ and since $15.31 > 3.86$ we have sufficient evidence to reject the null hypothesis that the extrusion process has no effect on the tensile strength of the tubing. Hence our conclusion (at the 5% level of significance) is that the extrusion process does affect the tensile strength of the aluminium tubing.

Latin Squares

We have previously looked at the elimination of a source of variation via the use of randomized blocks. If, in a given situation, there are two sources of variation to be controlled, a Latin Square design may provide the best possible analysis. Essentially this design groups the treatments involved into different blocks simultaneously. As an example, consider the problem of checking the quality of the output of four machines over four 6-hour shifts manned by the same four operators on each shift. A Latin Square design allocates the sixteen combinations to be used. The machines are labelled A, B, C and D .

A design is as follows.

		Operator			
		1	2	3	4
Shift	1	A	B	C	D
	2	B	C	D	A
	3	C	D	A	B
	4	D	A	B	C

Notice that each machine appears in each row and each column exactly once. Notice also that the number of observations must be equal to the square of the number of treatments (here $4^2 = 16$), or a multiple of this. In situations where a large number of treatments are used, a very substantial testing effort is implied. In return, a high reduction in errors is achieved since every row and every column is a complete replication. Experiments using Latin Squares should be designed so that the differences in rows and columns represent the major sources of variation to be considered.



Example 3

In an experiment designed to compare the tensile strengths of plastic tubes manufactured by different methods there are four different methods, A, B, C, D . It is also believed that there may be effects due to differences between batches of the plastic pellets which are used as raw material and between manufacturing plants. These factors are arranged in a **Latin Square** with four different manufacturing plants as the row factor and the pellet batch as the column factor.

Analyse the results (in coded units of tensile strength), given below, to determine whether or not there are significant differences between the methods. Use the 5% level of significance.

		Pellet batch			
		1	2	3	4
Plant	1	A 16.6	B 16.9	C 17.4	D 17.4
	2	D 17.1	C 16.8	B 19.2	A 16.6
	3	C 17.4	D 17.0	A 16.8	B 19.2
	4	B 18.6	A 17.4	D 17.4	C 19.2

Solution

The treatment totals are as follows:

Method	A	B	C	D
Total	67.4	73.9	70.8	68.9

The row totals are as follows:

Plant	1	2	3	4
Total	68.3	69.7	70.4	72.6

The column totals are as follows:

Batch	1	2	3	4
Total	69.7	68.1	70.8	72.4

The grand total is 281.0. The total sum of squares is

$$16.6^2 + \dots + 19.2^2 - \frac{281.0^2}{16} = 13.2375.$$

The treatments sum of squares is $\frac{1}{4}(67.4^2 + \dots + 68.9^2) - \frac{281.0^2}{16} = 5.8925.$

The batches sum of squares is $\frac{1}{4}(69.7^2 + \dots + 72.4^2) - \frac{281.0^2}{16} = 2.4625.$

The plants sum of squares is $\frac{1}{4}(68.3^2 + \dots + 72.6^2) - \frac{281.0^2}{16} = 2.4125.$

The analysis of variance table is as follows.

Source of variation	Degrees of freedom	Sum of squares	Mean square	Variance ratio
Treatments	3	5.8925	1.9642	4.77
Batches	3	2.4625	0.8208	1.99
Plants	3	2.4125	0.8042	1.95
Residual	6	2.4700	0.4117	
Total	15	13.2375		

The upper 5% point of the $F_{3,6}$ distribution is 4.76. Thus we can draw the following conclusions.

Treatments (manufacturing methods) The variance ratio is significant at the 5% level so we conclude that there is sufficient evidence to reject the null hypothesis that the treatments give equal mean tensile strengths.

Batches The variance ratio is not significant at the 5% level so we conclude that there is not sufficient evidence to reject the null hypothesis that the batches have no effect on mean tensile strengths.

Plants The variance ratio is not significant at the 5% level so we conclude that there is not sufficient evidence to reject the null hypothesis that the plants have no effect on mean tensile strengths.



The yields of a chemical process when using three different catalysts, A, B, C, are to be compared. It is also believed that there may be effects due to the reaction vessel used and the operator. A Latin Square design is used with three operators and three vessels. A batch is produced using each combination of catalyst, vessel and operator. The results (%) are as follows.

		Reaction vessel		
		1	2	3
Operator	1	A 81.4	B 63.9	C 59.6
	2	B 61.3	C 48.6	A 68.5
	3	C 58.3	A 70.2	B 72.5

Analyse the results to test for the effects of the factors. Use the 5% level of significance.

Your solution

Answer

The treatment totals are as follows:

Catalyst	A	B	C
Total	220.1	197.7	166.5

The row totals are as follows:

Operator	1	2	3
Total	204.9	178.4	201.0

The column totals are as follows:

Vessel	1	2	3
Total	201.0	182.7	200.6

The grand total is 584.3. The total sum of squares is

$$81.4^2 + \dots + 72.5^2 - \frac{584.3^2}{9} = 722.3556.$$

The treatments sum of squares is

$$\frac{1}{3}(220.1^2 + 197.7^2 + 166.5^2) - \frac{584.3^2}{9} = 483.1289.$$

The operators sum of squares is

$$\frac{1}{3}(204.9^2 + 178.4^2 + 201.0^2) - \frac{584.3^2}{9} = 136.4689.$$

The reaction vessels sum of squares is

$$\frac{1}{3}(201.0^2 + 182.7^2 + 200.6^2) - \frac{584.3^2}{9} = 72.8289.$$

The analysis of variance table is as follows.

Source of variation	Degrees of freedom	Sum of squares	Mean square	Variance ratio
Treatments (catalysts)	2	483.1289	241.56445	16.143
Operators	2	136.4689	68.23445	4.560
Reaction vessels	2	72.8289	36.41445	2.433
Residual	2	29.9289	14.96445	
Total	8	722.3556		

The upper 5% point of the $F_{2,2}$ distribution is 19.00.

Answer continued

Thus we can draw the following conclusions.

(a) Treatments (catalysts)

The variance ratio is not significant at the 5% level so we conclude that there is not sufficient evidence to reject the null hypothesis that the treatments give equal mean yields.

(b) Operators

The variance ratio is not significant at the 5% level so we conclude that there is not sufficient evidence to reject the null hypothesis that the operators have no effect on mean yields.

(c) Reaction vessels

The variance ratio is not significant at the 5% level so we conclude that there is not sufficient evidence to reject the null hypothesis that the reaction vessels have no effect on mean yields.

Note that, with only two degrees of freedom in the residual, this design gives poor power to the tests.

Exercises

- Aluminium is produced industrially by electrolysis in reduction cells. The 'current efficiency' of a reduction cell is the yield of aluminium as a percentage of the yield predicted in an ideal cell by Faraday's law. Improvements in the computer control system could improve current efficiency. Current efficiency also varies from one cell to another. In an experiment, three control schemes were compared. Each control scheme was applied to each of ten cells in random order. The current efficiency was measured over one week's operation in each case. The data are as follows. Here the cells are 'blocks.'

Control Scheme	Cell (block)									
	1	2	3	4	5	6	7	8	9	10
A	80.27	79.44	81.59	79.78	80.39	81.92	82.87	82.04	83.41	84.52
B	84.31	83.33	86.57	84.49	84.15	85.45	85.05	83.62	85.96	85.62
C	83.59	80.36	84.55	80.03	81.59	80.75	82.61	85.20	84.29	81.60

What conclusions can be drawn about the effect of the control schemes on current efficiency? Use the 5% level of significance.

Exercises continued

2. Plastic tubing is made by an extrusion process, starting with the plastic in pellet form. The tensile strength of the tubes may depend on which of four extrusion processes is used. Tubes were made using each of the four methods and samples tested for tensile strength. Four different batches of plastic pellets were used and each method was used once with plastic from each batch. The data are given below.

Batch (Block)	Tensile Strength Method			
	1	2	3	4
1	202	186	182	180
2	200	195	224	199
3	259	243	225	223
4	252	227	276	251

What conclusions may be drawn about the effect of the extrusion process on the resulting tubing? Use the 5% level of significance.

3. Crash tests with dummies are used to investigate the effects of different car seat-belt mechanisms. The response variable of interest is the maximum acceleration, in units of g, of the dummy's head. Four different variations of the mechanism, A, B, C, D, are compared in a Latin Square design with four dummies and four different impact angles, but all at the same impact speed.

The data are as follows:

		Dummy			
		1	2	3	4
Impact Angle	1	<i>B</i> 3.84	<i>C</i> 3.75	<i>D</i> 4.26	<i>A</i> 3.97
	2	<i>A</i> 3.44	<i>B</i> 3.93	<i>C</i> 4.18	<i>D</i> 3.36
	3	<i>C</i> 3.82	<i>D</i> 3.04	<i>A</i> 3.77	<i>B</i> 3.34
	4	<i>D</i> 3.43	<i>A</i> 2.93	<i>B</i> 3.77	<i>C</i> 3.41

Use an analysis of variance to test for the effects of the three factors. Use the 5% level of significance. Treat all three factors as "fixed."

What assumptions have to be made to justify your analysis?

How might the design be altered to allow investigation of the possibility of interaction effects between the factors?

Answers

1. Subtracting 80 from every observation will have no effect on the analysis of variance. The data now become as follows.

Control Scheme	Cell (block)										Total
	1	2	3	4	5	6	7	8	9	10	
A	0.27	-0.56	1.59	-0.22	0.39	1.92	2.87	2.04	3.41	4.52	16.23
B	4.31	3.33	6.57	4.49	4.15	5.45	5.05	3.62	5.96	5.62	48.55
C	3.59	0.36	4.55	0.03	1.59	0.75	2.61	5.20	4.29	1.60	24.57
Total	8.17	3.13	12.71	4.30	6.13	8.12	10.53	10.86	13.66	11.74	89.35

The total sum of squares for the 30 observations is given by

$$SS_T = 0.27^2 + \dots + 1.60^2 - \frac{89.35^2}{30} = 388.5413 - \frac{89.35^2}{30} = 122.400$$

The sum of squares for the three control schemes is given by

$$SS_S = \frac{16.23^2 + 48.55^2 + 24.57^2}{10} - \frac{89.35^2}{30} = \frac{3224.2003}{10} - \frac{89.35^2}{30} = 56.306$$

The sum of squares for the ten cells (blocks) is given by

$$SS_B = \frac{8.17^2 + \dots + 11.74^2}{3} - \frac{89.35^2}{30} = \frac{913.3349}{3} - \frac{89.35^2}{30} = 38.331$$

The residual sum of squares is given by

$$SS_E = SS_T - SS_S - SS_B = 122.400 - 56.306 - 38.331 = 27.763$$

These calculations lead to the following ANOVA table.

Source of Variation	Deg. of Freedom	Sum of Squares	Mean Square	Variance Ratio
Control Schemes	2	56.306	28.153	18.25
Cells (blocks)	9	38.311	4.259	2.76
Residuals (error)	18	27.763	1.542	
Total	29	122.400		

The following conclusions may be drawn.

Control Schemes

From F -tables $F_{2,18} = 3.55$ and since $18.25 > 3.55$ we have sufficient evidence to reject the null hypothesis that there is no difference in effect between the control schemes. It appears that Control Scheme B gives the greatest current efficiency.

Cells (blocks)

From F -tables $F_{9,18} = 2.46$ and since $2.76 > 2.46$ we do not have sufficient evidence to reject the null hypothesis that there is no difference in mean current efficiency between the cells. Allowing for the cell effect may, nevertheless, have improved the sensitivity of the test for control scheme effects.

Answers continued

2. Since there is only one observation on each combination of batch and extrusion method, we must assume that there is no interaction between the factors. We can calculate the sums of squares as follows.

Batch (Block)	Tensile Strength Method				Total
	1	2	3	4	
1	202	186	182	180	750
2	200	195	224	199	818
3	259	243	225	223	950
4	252	227	276	251	1006
Total	913	851	907	853	3524

The total sum of squares is given by

$$SS_T = 202^2 + \dots + 251^2 - \frac{3524^2}{16} = 789420 - 776161 = 13259$$

The sum of squares for the extrusion methods is given by

$$SS_M = \frac{913^2 + \dots + 853^2}{4} - \frac{3524^2}{16} = 777007 - 776161 = 846$$

The sum of squares for the blocks (batches) is given by

$$SS_B = \frac{750^2 + \dots + 1006^2}{4} - \frac{3524^2}{16} = 786540 - 776161 = 10379$$

The residual sum of squares is given by

$$SS_E = SS_T - SS_M - SS_B = 13259 - 846 - 10379 = 2034$$

These calculations lead to the following ANOVA table.

Source of Variation	Deg. of Freedom	Sum of Squares	Mean Square	Variance Ratio
Extrusion Methods	3	846	282.00	1.25
Batches (blocks)	3	10379	3459.67	15.31
Residual (error)	9	2034	226.00	
Total	15	13259		

The following conclusions may be drawn.

Extrusion Methods

From F -tables $F_{3,9} = 3.86$ and since $1.25 < 3.86$ we do not have sufficient evidence to reject the null hypothesis that the choice of extrusion method has no effect on the tensile strength of the tubes.

Batches (blocks)

From F -tables $F_{3,9} = 3.86$ and since $15.31 > 3.86$ we have sufficient evidence to reject the null hypothesis that mean tensile strength is the same in all batches. We conclude that mean tensile strength does differ between batches.

Answers continued

3. The treatment totals are as follows:

Mechanism	A	B	C	D
Total	14.11	14.88	15.16	14.09

The row totals are as follows:

Angle	1	2	3	4
Total	15.82	14.91	13.97	13.54

The column totals are as follows:

Dummy	1	2	3	4
Total	14.53	13.65	15.98	14.08

The grand total is 58.24.

The total sum of squares is

$$3.84^2 + \dots + 3.41^2 - \frac{58.24^2}{16} = 2.1568.$$

The treatments sum of squares is

$$\frac{1}{4}(14.11^2 + \dots + 14.09^2) - \frac{58.24^2}{16} = 0.22145.$$

The angles sum of squares is

$$\frac{1}{4}(15.82^2 + \dots + 13.54^2) - \frac{58.24^2}{16} = 0.77465.$$

The dummies sum of squares is

$$\frac{1}{4}(14.53^2 + \dots + 14.08^2) - \frac{58.24^2}{16} = 0.76895.$$

The analysis of variance table is as follows.

Source of variation	Degrees of freedom	Sum of squares	Mean square	Variance ratio
Treatments (mechanisms)	3	0.22145	0.07382	1.131
Angles	3	0.77465	0.25822	3.955
Dummies	3	0.76895	0.25632	3.926
Residual	6	0.39175	0.06529	
Total	15	2.15680		

The upper 5% point of the $F_{3,6}$ distribution is 4.76. Thus we can draw the following conclusions.

Answers continued

Treatments (mechanisms)

The variance ratio is not significant at the 5% level so we conclude that there is not sufficient evidence to reject the null hypothesis that the mechanisms give equal mean maximum accelerations.

Angles

The variance ratio is not significant at the 5% level so we conclude that there is not sufficient evidence to reject the null hypothesis that the angles have no effect on mean maximum accelerations.

(b) Dummies

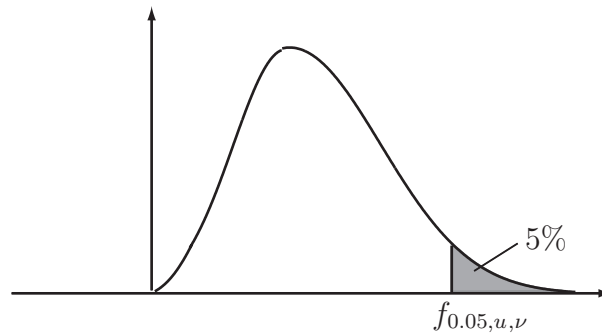
The variance ratio is not significant at the 5% level so we conclude that there is not sufficient evidence to reject the null hypothesis that the dummies have no effect on mean maximum accelerations.

The assumptions made are as follows.

- The observation in each cell is taken from a normal distributions.
- Each of these normal distributions has the same variance.
- The observations are independent.
- The effects of the factors are additive. That is, there are no interactions.

To allow investigation of the possibility of interactions we would need to make more than one observation in each cell.

Table 1: Percentage Points of the F Distribution (5% tail)



ν	Degrees of Freedom for the Numerator (u)														
	1	2	3	4	5	6	7	8	9	10	20	30	40	60	∞
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5	241.9	248.0	250.1	251.1	252.2	254.3
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.45	19.46	19.47	19.48	19.50
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.66	8.62	8.59	8.55	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.80	5.75	5.72	5.69	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.56	4.53	4.46	4.43	4.36
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	3.87	3.81	3.77	3.74	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.44	3.38	3.34	3.30	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.15	3.08	3.04	3.01	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	2.94	2.86	2.83	2.79	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.77	2.70	2.66	2.62	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.65	2.57	2.53	2.49	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.54	2.47	2.43	2.38	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.46	2.38	2.34	2.30	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.39	2.31	2.27	2.22	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.33	2.25	2.20	2.16	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.28	2.19	2.15	2.11	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.23	2.15	2.10	2.06	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.19	2.11	2.06	2.02	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.16	2.07	2.03	1.93	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.12	2.04	1.99	1.95	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.10	2.01	1.96	1.92	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.07	1.98	1.94	1.89	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.05	1.96	1.91	1.86	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.03	1.94	1.89	1.84	1.73
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.01	1.92	1.87	1.82	1.71
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	1.99	1.90	1.85	1.80	1.69
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	1.97	1.88	1.84	1.79	1.67
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	1.96	1.87	1.82	1.77	1.65
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	1.94	1.85	1.81	1.75	1.64
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	1.93	1.84	1.79	1.74	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	1.84	1.74	1.69	1.64	1.51
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.75	1.65	1.59	1.53	1.39
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.57	1.46	1.39	3.32	1.00