31 Decomposition to Components with a Unit Degree of Freedom

31.1. Introduction

In previous chapters, the total sum of the squares was decomposed into the sum of the variations of a correction factor, the main effect of a factor, and the error. This decomposition can be extended by looking at what the researcher feels could be causing the variations; or the total sum of the squares can be decomposed into the sum of the variations of causes with a unit degree of freedom. The latter method is described in this chapter. This chapter is based on Genichi Taguchi et al., *Design of Experiments*. Tokyo: Japanese Standards Association, 1973.

31.2. Comparison and Its Variation

The analysis of variance for the example in Section 30.2 showed that A is significant. This meant that the roundness differed by the order of the various processes used in making the pinholes. Sometimes we would like to know whether the difference was caused by the difference between A_1 and A_2, or between A_2 and the mean of A_1 and A_2.

Everyone knows to compare A_1 and A_2 by

$$L_1 = \frac{A_1}{10} - \frac{A_2}{10} \tag{31.1}$$

However, A_3 and the mean of A_1 and A_1 would be compared by using the following linear equation:

$$L_2 = \frac{A_1 + A_2}{20} - \frac{A_3}{10} \tag{31.2}$$

 Taguchi's Quality Engineering Handbook. Genichi Taguchi, Subir Chowdhury and Yuin Wu
Copyright © 2005 Genichi Taguchi, Subir Chowdhury, Yuin Wu.

L_1 and L_2 are linear equations; their sum of the coefficients for A_1, A_2, and A_3 is equal to zero either in L_1 or in L_2.

$$L_1\colon \quad \frac{1}{10} - \frac{1}{10} = 0 \tag{31.3}$$

$$L_2\colon \quad \frac{1}{20} + \frac{1}{20} - \frac{1}{10} = 0 \tag{31.4}$$

Assume that the sums $A_1, A_2, \dots , A_a$ each has the same number of data (b) making up their respective total.

In a linear equation with constant coefficients $A_1, A_2, \dots , A_a$,

$$L = c_1A_1 + c_2A_2 + \cdots + c_aA_a \tag{31.5}$$

when the sum of the coefficients is equal to zero,

$$c_1 + c_2 + \cdots + c_a = 0 \tag{31.6}$$

then L is called either *contrast* or *comparison*.

As described previously, in a linear equation with constant coefficients,

$$L = c_1A_1 + c_2A_2 + \cdots + c_aA_a \tag{31.7}$$

the variation of L, or S_L, is given by

$$S_L \frac{(c_1A_1 + \cdots + c_aA_a)^2}{(c_1^2 + c_1^2 + \cdots + c_a^2)b} \tag{31.8}$$

where S_L has one degree of freedom. Calculation and estimation of a contrast are made in exactly the same way.

In two contrasts,

$$L_1 = c_1A_1 + \cdots + c_aA_a \tag{31.9}$$

$$L_{21} = c_1'A_1 + \cdots + c_a'A_a \tag{31.10}$$

when their sum of products is equal to zero,

$$c_1c_1' + c_2c_2' + \cdots + c_ac_a' = 0 \tag{31.11}$$

L_1 and L_2 are *orthogonal*. When L_1 and L_2 are orthogonal, each of

$$S_{L_1} = \frac{L_1^2}{(c_1^2 + \cdots + c_a^2)b} \tag{31.12}$$

$$S_{L_2} = \frac{L_2^2}{(c_1'2 + \cdots + c_a'2)b} \tag{31.13}$$

is variation having one degree of freedom; each variation consists of one of the components in S_A. Therefore, when $(a - 1)$ comparisons, $L_1, L_2, \dots, L_{(a-1)}$, are orthogonal to each other, the following equation is used:

$$S_A = S_{L_1} + S_{L_2} + \cdots + S_{L(a-1)} \tag{31.14}$$

❒ Example 1

In an example of pinhole processing,

$$L_1 = \frac{A_1}{10} - \frac{A_2}{10} \tag{31.15}$$

$$L_2 = \frac{A_1 + A_2}{20} - \frac{A_3}{10} \tag{31.16}$$

The orthogonality between L_1 and L_2 is proven by

$$\left(\frac{1}{10}\right)\left(\frac{1}{20}\right) + \left(-\frac{1}{10}\right)\left(\frac{1}{20}\right) + 0\left(-\frac{1}{10}\right) = 0 \tag{31.17}$$

Therefore, the following relation can be made:

$$S_A = S_L + S_{L_2} \tag{31.18}$$

where S_{L1} and S_{L2} are calculated from equation (31.8) as

$$\begin{aligned} S_{L_1} &= \frac{(A_1/10 - A_2/10)^2}{[(1/10)^2 + (-1/10)^2](10)} = \frac{(1/10)^2\,(A_1 - A_2)^2}{(1/10)^2\,[1^2 + (-1)^2](10)} \\ &= \frac{(A_1 - A_2)^2}{20} \\ &= \frac{(87 - 85)^2}{20} \\ &= 0.2 \end{aligned} \tag{31.19}$$

$$\begin{aligned} S_{L_2} &= \frac{[(A_1 + A_2)/20 - A_3/10)]^2}{[(1/20)^2 + (1/20)^2 + (-1/10)^2](10)} \\ &= \frac{(1/20)^2\,(A_1 + A_2 - 2A_3)^2}{(1/20)^2(6)(10)} \\ &= 173.4 \end{aligned} \tag{31.20}$$

Thus, the magnitude of the variation in roundness, which is caused by A_1, A_2, and A_3, namely, S_A, is decomposed into the variation caused by the difference between A_1 and A_2, namely, S_{L_1}, and that variation caused by the difference between A_3 and the mean of A_1 and A_2, namely, S_{L_2}:

$$S_{L_1} + S_{L_2} = 0.2 + 173.4 = 174 = S_A \tag{31.21}$$

❐ Example 2

We have the following four types of products:

A_1: foreign products

A_2: our company's products

A_3: domestic: α company's products

A_4: domestic: β company's products

Two, ten, six, and six products were sampled from each type, respectively, and a 300-hour continuous deterioration test was made. The percent of deterioration (Table 31.1) was determined as follows:

$$y = \frac{\text{(value after the test)} - \text{(initial value)}}{\text{initial value}} \tag{31.22}$$

$$S_m = \frac{T^2}{n} = \frac{488^2}{24}$$

$$= 9923 \tag{31.23}$$

$$S_A = \frac{A_1^2}{2} + \frac{A_2^2}{10} + \frac{A_3^2}{6} + \frac{A_4^2}{6} - S_m$$

$$= \frac{26^2}{2} + \frac{175^2}{10} + \frac{147^2 + 140^2}{6} - 9923$$

$$= 346 \tag{31.24}$$

$$S_T = 12^2 + 14^2 + 20^2 + \cdots + 24^2$$

$$= 10{,}426 \tag{31.25}$$

$$S_e = S_T - S_m - S_A$$

$$= 10{,}426 - 9923 - 346$$

$$= 157 \tag{31.26}$$

The analysis of variance is shown in Table 31.2.

Instead of making an overall comparison among the four products, we usually want to make a more detailed comparison, such as:

L_1: difference between foreign and domestic products

L_2: difference between our company and the other domestic products

L_3: difference between the other domestic companies' products

Table 31.1
Percent of deterioration

Level	Data	Total
A_1	12, 14	26
A_2	20, 18, 19, 17, 15, 16, 13, 18, 22, 17	175
A_3	26, 19, 26, 28, 23, 25	147
A_4	24, 25, 18, 22, 27, 24	140
Total		488

The comparison above can be made by using the following linear equations:

$$L_1 = \frac{A_1}{2} - \frac{A_2 + A_3 + A_4}{22} = \frac{26}{2} - \frac{462}{22} = -8.0 \qquad (31.27)$$

$$L_2 = \frac{A_2}{10} - \frac{A_3 + A_4}{12} = \frac{175}{10} - \frac{287}{12} = -6.4 \qquad (31.28)$$

$$L_3 = \frac{A_3 - A_4}{6} = \frac{147 - 140}{6} = 1.2 \qquad (31.29)$$

Table 31.2
ANOVA table

Source	*f*	*S*	*V*	*S′*	ρ (%)
m	1	9,923	9,923	9,915	95.1
A	3	346	115.3	322	3.0
e	20	157	7.85	189	1.8
Total	24	10,426		10,426	100.0

These equations are orthogonal to each other, and also orthogonal with the general mean,

$$L_m = \frac{A_1 + A_2 + A_3 + A_4}{24} \tag{31.30}$$

For example, the orthogonality between L_m and L_1 is proven by

$$\left(\frac{1}{2}\right)\left(\frac{1}{24}\right)(2) + \left(-\frac{1}{22}\right)\left(\frac{1}{24}\right)(22) = 0 \tag{31.31}$$

where the sum of product of the coefficients is zero. The orthogonality between L_1 and L_2 is proven by

$$\left(\frac{1}{2}\right)(0)(2) + \left(-\frac{1}{22}\right)\left(\frac{1}{10}\right)(10) + \left(-\frac{1}{22}\right)\left(-\frac{1}{12}\right)(12) = 0 \tag{31.32}$$

After calculating the variations of L_1, L_2, and L_3, we obtain the following decomposition:

$$S_A = S_{L_1} + S_{L_2} + S_{L_3} \tag{31.33}$$

$$\begin{aligned} S_{L_1} &= \frac{L_1^2}{\text{sum of the coefficients squared}} \\ &= \frac{[A_1/2 - (A_2 + A_3 + A_4)/22]^2}{(1/2)^2(2) + (-1/22)^2(22)} \\ &= \frac{[11A_1 - (A_2 + A_3 + A_4)]^2}{(11^2)(2) + (-1)^2(22)} \\ &= \frac{(286 - 462)^2}{264} \\ &= 117 \end{aligned} \tag{31.34}$$

$$\begin{aligned} S_{L_2} &= \frac{[A_2/10 - (A_3 + A_4)/12]^2}{(1/10)^2(10) + (-1/12)^2(12)} \\ &= \frac{[6A_2 - 5(A_3 + A_4)]^2}{(6^2)(10) + (-5)^2(12)} \\ &= \frac{(1050 - 1435)^2}{660} = 225 \end{aligned} \tag{31.35}$$

$$\begin{aligned} S_{L_2} &= \frac{(A_3/6 - A_4/6)^2}{(1/6)^2(6) + (-1/6)^2(6)} \\ &= \frac{(A_3 - A_4)^2}{12} \\ &= 4 \end{aligned} \tag{31.36}$$

An ANOVA table with the decomposition of A into three components with a unit degree of freedom is shown in Table 31.3.

From the analysis of variance, it has been determined that there is a significant difference between L_1 and L_2, but none between the other domestic companies.

When there is no significance, as in this case, it is better to pool the effect with the error to show the miscellaneous effect. Pooling the effect of L_3 with the error, the degrees of contribution would then be 1.9%.

Where the error variance, V_e, is calculated from the pooled error variation,

$$V_e = \frac{157 + 4}{21} = 7.67 \tag{31.37}$$

$$\begin{aligned}\text{number of units} &= \text{sum of coefficients squared} \\ &= \left(\frac{1}{2}\right)^2 (2) + \left(-\frac{1}{22}\right)^2 (22) \\ &= \frac{1}{2} + \frac{1}{22} \\ &= \frac{12}{22} \\ &= \frac{6}{11}\end{aligned} \tag{31.38}$$

The other confidence intervals were calculated in the same way. Since L_3 of equation (31.29) is not significant, it is generally not estimated.

31.3. Linear Regression Equation

The tensile strength of a product was measured at different temperatures, x_1, x_2, ... , x_n, to get y_1, y_2, ... , y_n.

Table 31.3
ANOVA table

Source	f	S	V	S'	ρ (%)
m	1	9,923	9,923	9,915	95.1
A					
L_1	1	117	117	109	1.0
L_2	1	225	225	217	2.0
L_3	1	4	4		
e	20	157	7.85	185	1.8
Total	24	10,426		10,426	100.0

The relationship of the tensile strength, y, to temperature, x, is usually expressed by a linear function:

$$a + bx = y \tag{31.39}$$

Then n observational values (x_1, y_1), (x_2, y_2), ... , (x_n, y_n), are put in equation (31.39):

$$a + bx_i = y_i \qquad (i = 1, 2, \ldots , n) \tag{31.40}$$

Equation (31.40) is called an *observational equation*. When the n pairs of observational values are put in the equation, there are n simultaneous equations with two unknowns, a and b.

When the number of equations exceeds the number of unknowns, a solution that perfectly satisfies both of these equations does not exist; however, a solution that minimizes the differences between both sides of the equations can be obtained. That is, to find a and b would minimize the residual sum of the squares or the differences between the two sides:

$$S_e = (y_1 - a - bx_1)^2 + (y_2 - a - bx_2)^2 + \cdots + (y_n - a - bx_n)^2 \tag{31.41}$$

This solution was named by K. F. Gauss and is known as the *least squares method*. It is obtained by solving the following normal equations:

$$na + \left(\sum x_i\right)b = \sum y_i \tag{31.42}$$

$$\left(\sum x_i\right) a + \left(\sum x_i^2\right) b = \sum x_i y_i \tag{31.43}$$

where

$$n = \text{sum of coefficients squared of } a \text{ in (31.40); the coefficients are equal to 1}$$

$$= 1^2 + 1^2 + \cdots + 1^2 \tag{31.44}$$

$$\left(\sum x_i\right) = \text{sum of products of the coefficients of } a \text{ and } b \text{ in (31.40)}$$

$$= 1x_1 + 1x_2 + \cdots + 1x_n \tag{31.45}$$

$$\left(\sum y_i\right) = \text{sum of products of the coefficients of } a \text{ and observational values of } y \text{ in (31.40)}$$

$$= 1y_1 + 1y_2 + \cdots + 1y_n \tag{31.46}$$

$$\left(\sum x_i^2\right) = \text{sum of the coefficients squared of } b \text{ in (31.40)}$$

$$= x_1^2 + x_2^2 + \cdots + x_n^2 \tag{31.47}$$

$$\sum x_i y_i = \text{sum of products of the coefficients of } b \text{ and observational values of } y \text{ in (31.40)}$$

$$= x_1 y_1 + x_2 y_2 + \cdots + x_n y_n \tag{31.48}$$

It is highly desirable that the reader be able to write equations (39.42) and (31.43) at any time. Their memorization should not be difficult.

To solve the simultaneous equations (31.42) and (31.43), x_i is multiplied by both sides of (31.42). Also, n has to be multiplied by both sides of (31.43). After subtracting the same sides of the two equations from each other, term a disappears and term b is left.

$$\left[\left(\sum x_i\right)^2 - n\sum\sum x_i\right)^2\right] b = \left(\sum x_i\right)\left(\sum y_i\right) - n\sum x_i y_i \tag{31.49}$$

From this,

$$b = \frac{n\sum x_i y_i - \left(\sum x_i\right)\left(\sum y_i\right)}{n\left(\sum x_i^2\right) - \left(\sum x_i\right)^2} \tag{31.50}$$

and then dividing the denominator and the numerator by n, respectively,

$$b = \frac{x_1 y_1 + \cdots + x_n y_n - [(x_1 + \cdots + x_n)(y_1 + \cdots + y_n)/n]}{x_1^2 + \cdots + x_n^2 - \dfrac{(x_1 + \cdots + x_n)^2}{n}}$$

$$= \frac{S_T(xy)}{S_T(xx)} \tag{31.51}$$

where $S_T(xx)$ is the total variation of the temperature, x:

$$S_T(xx) = x_1^2 + x_2^2 + \cdots + x_n^2 - \frac{(x_1 + x_2 + \cdots + x_n)^2}{n} \tag{31.52}$$

$S_T(xy)$ is called the *covariance* of x and y and is determined by

$$S_T(xy) = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n - \frac{(x_1 + x_2 + \cdots + x_n)(y_1 + y_2 + \cdots + y_n)}{n} \tag{31.53}$$

In the equation of covariance, the square term in the equation of variation is substituted for by the product of x and y.

Putting b of (31.50) into (31.42), a is obtained:

$$a = \frac{1}{n}\left[\sum y_i - \frac{S_T(xy)}{S_T(xx)}\left(\sum x_i\right)\right] \tag{31.54}$$

It is known from (31.54) that when

$$\sum x_i = 0 \tag{31.55}$$

then

$$a = \frac{y_1 + \cdots + y_n}{n} = \bar{y} \tag{31.56}$$

Thus, the estimation of the unknowns a and b becomes very simple. For this purpose, equation (31.39) may be expanded as follows:

$$m + b(x - \bar{x}) = y \tag{31.57}$$

where $\bar{x}$ is the mean of $x_1, x_2, \ldots, x_n$. Such an expansion is called an *orthogonal expansion*.

The orthogonal expansion has the following meaning: Either the general mean, m, or the linear coefficient, b, is estimated from the linear equation of y. Using equation (31.57), the two linear equations are orthogonal; therefore, the magnitudes of their influences are easily evaluated.

In the observational equation

$$m + b(x_i - \bar{x}) = y_i \qquad (i = 1, 2, \ldots, n) \tag{31.58}$$

the sum of the products of the unknowns m and b,

$$\sum (x_i - \bar{x}) = 0$$

becomes zero. Accordingly, the normal equations become

$$nm + 0b = \sum y_i$$

$$0m + \left[\sum (x_i - \bar{x})^2\right] b = \sum (x_i - \bar{x}) y_i \tag{31.59}$$

Letting the estimates of m and b be $\hat{m}$ and $\hat{b}$,

$$\hat{m} = \frac{y_1 + y_2 + \cdots + y_n}{n} \tag{31.60}$$

$$\hat{b} = \frac{\sum (x_i - \bar{x}) y_i}{\sum (x_i - \bar{x})^2} = \frac{S_T(xy)}{S_T(xx)} \tag{31.61}$$

Not only $\hat{m}$, but also $\hat{b}$, is a linear equation of $y_1, y_2, \ldots, y_n$.

$$c_1 = \frac{x_1 - \bar{x}}{S_T(xx)}$$

$$c_2 = \frac{x_2 - \bar{x}}{S_T(xx)}$$

$$\vdots$$

$$c_n = \frac{x_n - \bar{x}}{S_T(xx)} \tag{31.62}$$

The number of units of the sum of the coefficients squared is

$$c_1^2 + c_2^2 + \cdots + c_n^2 = \left[\frac{x_1 - \bar{x}}{S_T(xx)}\right]^2 + \left[\frac{x_2 - \bar{x}}{S_T(xx)}\right]^2 + \cdots + \left[\frac{x_n - \bar{x}}{S_T(xx)}\right]^2$$

$$= \frac{1}{[S_T(xx)]^2}[(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \cdots (x_n - \bar{x})^2]$$

$$= \frac{S_T(xx)}{[S_T(xx)]^2}$$

$$= \frac{1}{S_T(xx)} \tag{31.63}$$

The variations of m, b, and error are

$$S_m = \text{CF} = \frac{(y_1 + \cdots + y_n)^2}{n}$$

$$S_b = \frac{(b)^2}{\text{no. of units}}$$

$$= \frac{[S_T(xy)/S_T(xx]^2}{1/S_T(xx)}$$

$$= \frac{S_T(xy)^2}{S_T(xx)} \tag{31.64}$$

$$S_e = y_1^2 + y_2^2 + \cdots + y_n^2 - S_m - S_b \tag{31.65}$$

The number of degrees of freedom is 1 for S_m or S_b, and $n - 2$ for S_e.

❒ Example

To observe the change of tensile strength of a product, given a change in temperature, the tensile strength of two test pieces was measured at four different temperatures, with the following results (kg/mm²):

A_1 (0°C): 84.0, 85.2

A_2 (20°C): 77.2, 76.8

A_3 (40°C): 67.4, 68.6

A_4 (60°C): 58.2, 60.4

If there is no objective value, such as a given specification, the degrees of freedom of the total variation is 7, where the degree of freedom for the correction factor is not included.

Subtracting a working mean, 70.0, the data in Table 31.4 are obtained.

$$\text{CF} = \frac{17.8^2}{8} = 39.60 \tag{31.66}$$

$$S_T = 14.0^2 + 15.2 + \cdots + (9.6)^2 - \text{CF} = 765.24 - 39.60$$
$$= 725.64 \tag{31.67}$$

Assuming that tensile strength changes in the same way as the linear function of temperature, A, the observational equations will become

$$y = m + b(A - \bar{A}) \tag{31.68}$$

where A signifies temperature and $\bar{A}$ represents the mean value of the various temperature changes:

$$\bar{A} = \tfrac{1}{8}[(2)(0) + (2)(20) + (2)(40) + (2)(60)] = 30°\text{C} \tag{31.69}$$

In the linear equation, m is a constant and b is a coefficient indicating how much the tensile strength decreases with a 1°C temperature change. The actual observational values are put into equation (31.68), as follows:

$$\begin{aligned} m + b(0 - 30) &= 84.0 \\ m + b(0 - 30) &= 85.2 \\ m + b(20 - 30) &= 77.2 \\ m + b(20 - 30) &= 76.8 \\ m + b(40 - 30) &= 67.4 \\ m + b(40 - 30) &= 68.6 \\ m + b(60 - 30) &= 58.2 \\ m + b(60 - 30) &= 60.4 \end{aligned} \tag{31.70}$$

Table 31.4
Data after subtracting a working mean

Level	Data	Total
A_1	14.0, 15.2	29.2
A_2	7.2, 6.8	14.0
A_3	−2.6, −1.4	−4.0
A_4	−11.8, −9.6	−21.4
Total		17.8

The unknowns are m and b, and there are eight equations. Therefore, the least squares method is used to find m and b.

$$\begin{aligned} n &= \text{sum of the coefficients squared of } m \\ &= 8 \end{aligned} \tag{31.71}$$

$$\begin{aligned} \left(\sum x_i\right) &= \text{sum of the coefficients of } b \\ &= (-30) + (-30) + (-10) + (-10) + 10 + 10 + 30 + 30 \\ &= 0 \end{aligned} \tag{31.72}$$

$$\begin{aligned} \left(\sum y_i\right) &= \text{sum of products of } y \text{ and the coefficients of } m \\ &= 84.0 + 85.2 + \cdots + 60.4 \\ &= (70.0)(8) + 17.8 \\ &= 577.8 \end{aligned} \tag{31.73}$$

$$\begin{aligned} \left(\sum x_i^2\right) &= \text{sum of the coefficients squared of } b \\ &= (-30)^2 + (-30)^2 + (-10)^2 + (-10)^2 + 10^2 + 10^2 + 30^2 + 30^2 \\ &= 4000 \end{aligned} \tag{31.74}$$

$$\begin{aligned} \left(\sum x_i y_i\right) &= \text{sum of products of } y \text{ and the coefficients of } b \\ &= (-30)(84.0) + (-30)(85.2) + \cdots + (30)(58.2) + (30)(60.4) \\ &= (-30)(169.2) + (-10)(154.0) + (10)(136.0) + (30)(118.6) \\ &= -1698.0 \end{aligned} \tag{31.75}$$

The following simultaneous equations are then obtained:

$$\begin{aligned} 8m + 0b &= 577.8 \\ 0m + 4000b &= -1698.0 \end{aligned} \tag{31.76}$$

Solving these yields

$$\begin{aligned} \hat{m} &= \frac{577.8}{8} \\ &= 72.22 \end{aligned} \tag{31.77}$$

$$\hat{b} = \frac{-1698.0}{4000} = -0.4245 \tag{31.78}$$

The orthogonality of these equations is proved as follows: Let the eight observational data be $y_1, y_2, \ldots, y_8$.

$$\hat{m} = \frac{y_1 + y_2 + \cdots + y_8}{8} \tag{31.79}$$

$$\hat{b} = \frac{-30(y_1 + y_2) - 10(y_3 + y_4) + 10(y_5 + y_6) + 30(y_7 + y_8)}{4000} = \frac{-3(y_1 + y_2) - (y_3 + y_4) + (y_5 + y_6) + 3(y_7 + y_8)}{400} \tag{31.80}$$

The sum of products of the corresponding coefficients of $\hat{m}$ and $\hat{b}$ is

$$2\left[\left(\frac{1}{8}\right)\left(\frac{-3}{400}\right) + \left(\frac{1}{8}\right)\left(\frac{-1}{400}\right) + \left(\frac{1}{8}\right)\left(\frac{3}{400}\right)\right] = 0 \tag{31.81}$$

The variation of m, S_m, is identical to the correction factor.

$$S_m = \text{CF} = \frac{577.8^2}{8} = 41{,}731.60 \tag{31.82}$$

$$S_b = \frac{S_T(xy)^2}{S_T(xx)} = \frac{(-1698.0)^2}{4000} = 720.80 \tag{31.83}$$

The total sum of the observational values squared, S_T, is

$$S_T = 84.0^2 + 85.2^2 + \cdots + 60.4^2 = 42{,}457.24 \tag{31.84}$$

The error variation, S_e, is then

$$S_e = S_T - S_m - S_b = 42{,}457.24 - 41{,}731.60 - 720.80 = 4.84 \tag{31.85}$$

The error variance, with 6 degrees of freedom, is

$$V_e = \frac{S_e}{6} = \frac{4.84}{6} = 0.807 \qquad (31.86)$$

The pure variations of the general mean and the linear coefficient, b, are

$$S'_m = S_m - V_e = 41{,}731.60 - 0.807 = 41{,}730.793 \qquad (31.87)$$

$$S'_b = S_b - V_e = 720.80 - 0.807 = 719.993 \qquad (31.88)$$

Degrees of contributions are calculated as follows:

$$\rho_m = \frac{41{,}730.793}{42{,}457.24} = 98.289\% \qquad (31.89)$$

$$\rho_b = \frac{719.993}{42{,}457.24} = 1.696\% \qquad (31.90)$$

$$\rho_e = \frac{4.84 + (2)(0.807)}{42{,}457.24} = 0.015\% \qquad (31.91)$$

The analysis of variance is shown in Table 31.5.

The tensile strength y at temperature x is estimated as

$$y = \hat{m} + \hat{b}(x - \bar{x}) = 72.22 - 0.4245(x - 30) \qquad (31.92)$$

31.4. Application of Orthogonal Polynomials

The type of contrast, or comparison, to cite is up to a researcher. The appropriate selection is crucial if the results achieved are to be based on practical and justifiable

Table 31.5
ANOVA table

Source	f	S	V	S′	ρ (%)
m	1	41,731.60	41,731.60	41,730.793	98.289
b	1	720.80	720.80	719.993	1.696
e	6	4.84	0.807	6.545	0.015
Total	8	42,457.24		42,457.240	100.000

comparisons rather than simply on the theoretical methodologies. Only a researcher or an engineer who is thoroughly knowledgeable of the products or conditions being investigated can know what comparison would be practical. However, the contrasts of orthogonal polynomials in linear, quadratic, ... , are often used. Let A_1, A_2, ... , A_a denote the values of first, second, ... , a level, respectively.

Assuming that the levels of A are of equal intervals, the levels are expressed as

$$\begin{aligned} A_1 &= A_1 \\ A_2 &= A_1 + h \\ A_3 &= A_1 + 2h \\ &\vdots \\ A_a &= A_1 + (a-1)h \end{aligned} \tag{31.93}$$

From each level of A_1, A_2, ... , A_a, r data are taken. The sum of each level is denoted by y_1, y_2, ... , y_a, respectively. When A has levels with the same interval, an orthogonal polynomial, which is called the *orthogonal function* of P. L. Chebyshev, is generally used.

The expanded equation is

$$y = b_0 + b_1(A - \overline{A}) + b_2\left[(A - \overline{A})^2 - \frac{a^2 - 1}{12}h^2\right] + \cdots \tag{31.94}$$

where $\overline{A}$ is the mean of the levels of A:

$$\overline{A} = A_1 + \frac{a-1}{2}h \tag{31.95}$$

The characteristics of the expansion above are that it attaches importance to the terms of the lower orders, such as a constant, a linear function, or even a quadratic function. First, the constant term b_0 is tried. If it does not fit well, a linear term is tried. If it still does not show linear tendency, the quadratic term is tried, and so on. When sums y_1, y_2, ... , y_a were obtained from r data of the levels A_1, A_2, ... , A_a, respectively, the values are b_0, b_1, ... Equation (31.94) will be obtained by solving the following observational equation with order a using the least squares method:

$$b_0 + b_1(A_1 - \overline{A}) + b_2\left[(A_1 - \overline{A})^2 - \frac{a^2 - 1}{12} h^2\right] + \cdots = \frac{y_1}{r}$$

$$\vdots \qquad (31.96)$$

$$b_0 = b_1(A_a - A) + b_2\left[(A_a - \overline{A})^2 - \frac{a^2 - 1}{12} h^2\right] + \cdots = \frac{y_a}{r}$$

The estimates of b_0, b_1, ... are denoted by b'_0, b'_0, b'_0, ... :

$$b'_0 = \frac{y_1 + \cdots + y_a}{ar} \qquad (31.97)$$

$$b'_1 = \frac{(A_1 - \overline{A})y_1 + \cdots + (A_a - \overline{A})y_a}{[(A_1 - \overline{A})^2 + \cdots + (A_a - \overline{A})^2]r} \qquad (31.98)$$

It is easy to calculate b'_0, since it is the mean of the total data. But it seems to be troublesome to obtain b'_1, b'_2, ... Actually, it is easy to estimate them by using Table 31.6, which shows the coefficients of orthogonal polynomials when $a = 3$ (three levels). b_1 and b_2 are linear and quadratic coefficients, respectively. These are given by

$$b'_1 = \frac{W_1A_1 + W_2A_2 + W_3A_3}{r(\lambda S)\,h} \qquad (31.99)$$

$$b'_2 = \frac{W_1A_1 + W_2A_2 + W_3A_3}{r(\lambda S)\,h^2} \qquad (31.100)$$

In the equations above, A_1, A_2, and A_3 are used instead of y_1, y_2, and y_3. For b_1, the values of W_1, W_2, and W_3 are -1, 0, and 1. Also, S is 2. b'_1 is then

$$b'_1 = \frac{-A_1 + A_3}{2rh} \qquad (31.101)$$

Similarly,

$$b'_2 = \frac{A_1 - 2A_2 + A_3}{2rh^2} \qquad (31.102)$$

Table 31.6
Coefficients of orthogonal polynomial for three levels

Coefficients	b_1	b_2
W_1	−1	1
W_2	0	−2
W_3	1	1
$\lambda^2 S$	2	6
λS	2	2
S	2	$\frac{2}{3}$

The variations of b_1' and b_2' are denoted by S_{A1} and S_{Aq}, respectively. These are given by the squares of equations (31.101) and (31.102) divided by their numbers of units, respectively.

$$S_{A_l} = \frac{(-A_1 + A_3)^2}{2r} \tag{31.103}$$

$$S_{A_q} = \frac{(A_1 - 2A_2 + A_3)^2}{6r} \tag{31.104}$$

The denominator of equations (31.103) and (31.104) is $r(\lambda^2 S)$ ($\lambda^2 S$ is the sum of squares of the coefficients of W). The effective number of replication, n_e, is given by

$$b_1': \quad n_e = rSh^2 \tag{31.105}$$

$$b_2': \quad n_e = rSh^4 \tag{31.106}$$

In general, the ith coefficient b_i and its variation on S_{A_i} are

$$b_1' = \frac{W_1A_1 + \cdots + W_aA_a}{r(\lambda S)\,h^i} \tag{31.107}$$

$$S_{A_i} = \frac{(W_1A_1 + \cdots + W_aA_a)^2}{(W_1^2 + \cdots + W_a^2)r} \tag{31.108}$$

❒ Example

To observe the change of the tensile strength of a synthetic resin due to changes in temperature, the tensile strength of five test pieces was measured at A_1 = 5°C, A_2 = 20°C, and A_3 = 35°C, respectively, to get the results in Table 31.7. The main effect A is decomposed into linear, quadratic, and cubic components to find the correct order of polynomial to be used. From Table 31.8, find the coefficients at the number of levels k = 4.

Table 31.7
Tensile strength (kg/mm²)

Level	Data	Total
A_1 (5°C)	43, 47, 45, 43, 45	233
A_2 (20°C)	43, 41, 45, 41, 39	209
A_3 (35°C)	37, 36, 39, 40, 38	190
A_4 (50°C)	34, 32, 36, 35, 35	172

Table 31.8
Orthogonal polynomials with equal intervals

	k = 2	k = 3		k = 4			k = 5			
Coeff.	b_1	b_1	b_2	b_1	b_2	b_3	b_1	b_2	b_3	b_4
W_1	−1	−1	1	−3	1	−1	−2	2	−1	1
W_2	1	0	−2	−1	−1	3	−1	−1	2	−4
W_3		1	1	1	−1	−3	0	−2	0	6
W_4				3	1	1	1	−1	−2	−4
W_5							2	2	1	1
$\lambda'S$	2	2	6	20	4	20	10	14	10	70
λS	1	2	2	10	4	6	10	14	12	24
S	1/2	2	2/3	5	4	9/5	10	14	72/5	288/85
λ	2	1	8	2	1	10/3	1	1	5/6	85/12

	k = 6					k = 7				
Coeff.	b_1	b_2	b_3	b_4	b_5	b_1	b_2	b_3	b_4	b_5
W_1	−5	5	−5	1	−1	−3	5	−1	3	−1
W_2	−3	−1	7	−3	5	−2	0	1	−7	4
W_3	−1	−4	4	2	−10	−1	−3	1	1	−5
W_4	1	−4	−4	2	10	0	−4	0	6	0
W_5	3	−1	−7	−3	−5	1	−3	−1	1	5
W_6	5	5	5	1	1	2	0	−1	−7	−4
W_7						3	5	1	3	1
$\lambda'S$	70	84	180	28	252	28	84	6	154	84
λS	35	56	108	48	120	28	84	36	264	240
S	35/2	112/3	324/5	576/7	400/7	28	84	216	3,168/7	4,800/7
λ	2	3/2	5/3	7/12	21/10	1	1	1/6	7/12	7/20

	k = 8					k = 9				
Coeff.	b_1	b_2	b_3	b_4	b_5	b_1	b_2	b_3	b_4	b_5
W_1	−7	7	−7	7	−7	−4	28	−14	14	−4
W_2	−5	1	5	−13	23	−3	7	7	−21	11
W_3	−3	−3	7	−3	−17	−2	−8	13	−11	−4
W_4	−1	−5	3	9	−15	−1	−17	9	9	−9
W_5	1	−5	−3	9	15	0	−20	0	18	0
W_6	3	−3	−7	−3	17	1	−17	−9	9	9
W_7	5	1	−5	−13	−23	2	−8	−13	−11	4
W_8	7	7	7	7	7	3	7	−7	−21	−11
W_9						4	28	14	14	4
$\lambda' S$	168	168	264	616	2184	60	2772	990	2,002	468
λS	84	168	396	1,056	3,102	60	924	1,188	3,432	3,120
S	42	168	594	12,672/7	31,200/7	60	308	7,128/5	41,184/7	20,800
λ	2	1	2/3	12/7	10/7	1	3	5/6	7/12	3/20

	k = 10					k = 11				
Coeff.	b_1	b_2	b_3	b_4	b_5	b_1	b_2	b_3	b_4	b_5
W_1	−9	6	−42	18	−6	−5	15	−30	6	−3
W_2	−7	2	14	−22	14	−4	6	6	−6	6
W_3	−5	−1	35	−17	−1	−3	−1	22	−6	1
W_4	−3	−3	31	3	−11	−2	−6	23	−1	−4
W_5	−1	−4	12	18	−6	−1	−9	14	4	−4
W_6	1	−4	−12	18	6	0	−10	0	6	0
W_7	3	−3	−31	3	11	1	−9	−14	4	4
W_8	5	−1	−35	−17	1	2	−6	−23	−1	4
W_9	7	2	−14	−22	−14	3	−1	−22	−6	−1
W_{10}	9	6	42	18	6	4	6	−6	−6	−6

Table 31.8
Orthogonal polynomials with equal intervals (*Continued*)

	$k = 10$					$k = 11$				
Coeff.	b_1	b_2	b_3	b_4	b_5	b_1	b_2	b_3	b_4	b_5
W_{11}						5	15	30	6	3
$\lambda' S$	330	132	8,580	2,860	780	110	858	4,290	286	1
										5
										6
λS	165	264	5,148	6,864	7,800	110	858	5,148	3,432	6
										2
										4
										0
S	165/2	528	15,444/5	82,368/5	78,000	110	858	30,886/5	41,184	2
										4
										9
										6
										0
										0
λ	2	1/2	5/3	5/12	1/10	1	1	5/6	1/12	1
										4
										0

	$k = 12$					$k = 13$				
Coeff.	b_1	b_2	b_3	b_4	b_5	b_1	b_2	b_3	b_4	b_5
W_1	−11	55	−33	33	−33	−6	22	−11	99	−22
W_2	−9	25	3	−27	57	−5	11	0	−66	33
W_3	−7	1	21	−33	21	−4	2	6	−96	18
W_4	−5	−17	25	−13	−29	−3	−5	8	−54	−11
W_5	−3	−29	19	12	−44	−2	−10	7	11	−26
W_6	−1	−35	7	28	−20	−1	−13	4	64	−20
W_7	1	−35	−7	28	20	0	−14	0	84	0
W_8	3	−29	−19	12	44	1	−13	−4	64	20
W_9	5	−17	−25	−13	29	2	−10	−7	11	26
W_{10}	7	1	−21	−33	−21	3	−5	−8	−54	11
W_{11}	9	25	−3	−27	57	4	2	−6	−96	−18
W_{12}	11	55	33	33	33	5	11	0	−66	−33
W_{13}						6	22	11	99	22
$\lambda' S$	572	12,012	5,148	8,008	15,912	182	2,002	572	68,068	6,188
λS	286	4,004	7,722	27,456	106,080	182	2,002	3,432	116,688	106,080
S	143	4,004/3	11,583	658,944/7	707,200	182	2,002	20,592	1,400,256/7	12,729,600/7
λ	2	3	2/3	7/24	3/20	1	1	1/6	7/12	7/120

$$S_{A_l} = \frac{(W_1A_1 + W_2A_2 + W_2A_2 + W_3A_3 + W_4A_4)^2}{r(\lambda^2 S)}$$

$$= \frac{[-3(223) - 209 + 190 + 3(172)]^2}{(5)(20)}$$

$$= \frac{(-172)^2}{100} = 296 \qquad (31.109)$$

$$S_{A_q} = \frac{(223 - 209 - 190 + 172)^2}{(5)(4)} = \frac{(-4)^2}{20} = 1 \qquad (31.110)$$

$$S_{A_c} = \frac{[-223 + 3(209) - 3(190) + 172]^2}{(5)(20)} = \frac{6^2}{100} = 0 \qquad (31.111)$$

$$S_T = 43^2 + 47^2 + \cdots + 35^2 - \text{CF} = 348$$

$$S_e = S_T - S_{A_l} - S_{A_q} - S_{Ac}$$

$$= 51 \qquad (31.112)$$

It is known from Table 31.9 that only the linear term of A is significant; hence, the relationship between temperature, A, and tensile strength, y, can be deemed to be linear within the range of our experimental temperature changes. The estimate of the linear coefficient of temperature, b, is

$$b_1' = \frac{-3(223) - 209 + 190 + 3(172)}{r(\lambda S)h}$$

$$= \frac{-172}{(5)(10)(15)} = -0.23 \qquad (31.113)$$

Table 31.9
ANOVA table

Source	f	S	V	ρ (%)
A				
Linear	1	296	296	84.2
Quadratic	1	1	1	
Cubic	1	0	0	
e	16	51	3.2	
Total	19	348		100.0
(e)	(18)	(52)	(2.9)	(15.8)

The relationship between temperature and tensile strength is then

$$y = b_0' + b_1' (A - \overline{A})$$

$$= \frac{794}{20} - 0.23(A - 27.5)$$

$$= 39.70 - 0.23(A - 27.5) \qquad (31.114)$$

b_1: 1 *linear*

b_2: 2 *quadratic*

$$\hat{b} = \frac{W_1A_1 + \cdots + W_aA_a}{r(\lambda S)h^i} \qquad (31.115)$$

$$\mathrm{Var}(\hat{b}_i) = \frac{\sigma^2}{rSh^{2i}}$$

$$S_{\hat{b}_i} = \frac{(W_1A_1 + \cdots + W_aA_a)^2}{r(\lambda^2 S)}$$