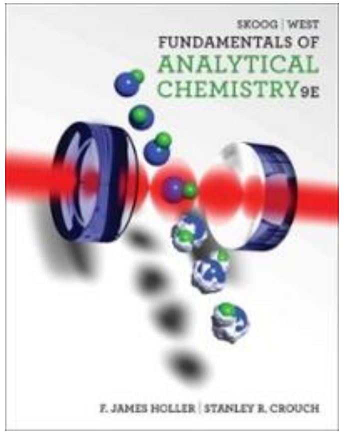
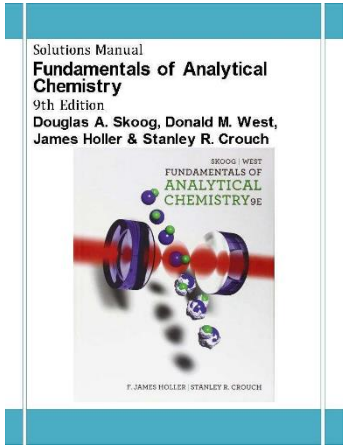
STATISTICS IN CHEMISTRY

These materials have been prepared by H. Elif Kormalı Ertürün for educational purposes only (as lecture notes) using the following resources. Responsibility for reproducing any part of these materials in any form or by any means or stored in a retrieval system for different purposes, rests with the third person performing the action.

1.



2.



1. Skoog DA, West DM, Holler FJ, Crouch SR. Fundamentals of Analytical Chemistry. Nelson Education; 2013.
2. Skoog DA, West DM, Holler FJ, Crouch SR. Solutions Manual of Fundamentals of Analytical Chemistry. Nelson Education; 2013.

Comparison of Precision, Analysis of Variance (ANOVA)

Comparison of Precisions, F test:

It is very important to know and compare the precisions of the results obtained from the applied analysis method. The comparison of the two populations is performed by a statistical method called the F test. The variances of each population are calculated. The F value is calculated by proportioning the large variance to the numerator and the small variance to the denominator.

$$\sigma_A^2 > \sigma_B^2 \quad \text{so} \quad F = \frac{\sigma_A^2}{\sigma_B^2}$$

The obtained value is compared to the critical value (the value found in the F test table, based on the degrees of freedom of the numerator and denominator for a given confidence level). If the calculated value is less than the theoretical value, it is interpreted that there is no significant difference between the precision of that confidence level.

Variance Analysis

The comparison of means from the binary population was performed by applying t-test as previously described. The t-test is no longer performed when the number of populations increases, ie when there are 3 or more. Instead, multiple variance analysis, known as ANOVA, is used for analysis of variance.

Typical examples of ANOVA testing are as follows:

Whether there is a difference between the means of the analysis found by 5 different analysts from the same sample,

Whether there is a difference between the analysis results obtained at four different temperatures,

Whether there is a difference between the results obtained by 3 different analysis methods of the same sample,

Whether there is a difference between the results obtained at different pH 4.

The differences in the elements compared are called factors or applications. In the first example, the analyzer is the temperature in the second, the method in the third, and the pH in the fourth. The averages obtained separately from the factor of interest indicate the level. For example, in the first example, the factor is the analyzer. Since there are 5 analysts, it is called 5 levels. In the ANOVA test, factor levels are often referred to as groups.

In this lesson, we will focus on single factor ANOVA.

Single Factor ANOVA

The analysis results of the *I* number of samples, the averages, $\bar{x}_1$, $\bar{x}_2$, $\bar{x}_3$.......$\bar{x}_I$ and the variances S_1^2, S_2^2, S_3^2....S_I^2.... In addition, the general average $\bar{\bar{x}}$ den obtained from all data is calculated from the equation given below.

$$\bar{\bar{x}} = \frac{\bar{x}_1 + \bar{x}_2 + \bar{x}_3 + \cdots . + \bar{x}_I}{I}$$

The quantities related to the sum of squares are calculated for use in the F test. One of these is the sum of squares related to the factor (SSF), which we can perceive as interlaboratory. The other is the sum of squares related to errors (SSE), which we can perceive as in-laboratory.

Sum of squares related to factor (SSF) It is calculated using the equation given below:

$$SSF= N_1(\bar{x}_1 - \bar{\bar{x}})^2 + N_2(\bar{x}_2 - \bar{\bar{x}})^2 + N_3\,(\bar{x}_3 - \bar{\bar{x}})^2 + \ldots\ldots\ldots + N_I\,(\bar{x}_I - \bar{\bar{x}})^2$$

Sum of squares related to errors (SSE) It is calculated using the equation given below:

$$SSE= (N_1 - 1)\,s_1^2 + (N_2 - 1)\,s_2^2 + (N_3 - 1)\,s_3^2 + \ldots\ldots\ldots + (N_I - 1)\,s_I^2$$

Total Squares Total (SST)

$$SST = SSF + SSE$$

If we show the total number of analyzes in all groups with *N* and the number of groups with I, the degrees of freedom of each sum of squares are calculated as given below.

degree of freedom for SST : (N-1)
degree of freedom for SSF : (N-I)
degree of freedom for SSE : (I - 1)

From the sum of squares, the average sum of squares (MS) is calculated as follows.

Factor-related mean squares, $MSF = \frac{\mathrm{SSF}}{\mathrm{I}-1}$

Mean squares related to error, $MSE = \frac{\mathrm{SSE}}{\mathrm{N}-\mathrm{I}}$

F test application

$$F = \frac{\mathrm{MSF}}{\mathrm{MSE}}$$

The calculated value is compared with the critical value. If $F_{calculated} > F_{critical}$, the null hypothesis is rejected.

If the null hypothesis is rejected, the smallest significant difference must be calculated to find out which results are different from the others.

If the null hypothesis is rejected, the smallest significant difference must be calculated to find out which results are different from the others.

$$LSD = \sqrt{\frac{2 \ x \ MSE}{N_g}}$$

Here;

t: t value of MSE in the degree of freedom and the calculated confidence level,

N_g: Number of analyzes in a group.