13.2 Analysis of variance: testing for the equality of *k* population means

Analysis of variance can be used to test for the equality of *k* population means. The general form of the hypotheses tested is

$$H_0\colon \mu_1 = \mu_2 = \ldots = \mu_k$$
$$H_1\colon \text{Not all population means are equal}$$

where

$$\mu_j = \text{mean of the } j\text{th population}$$

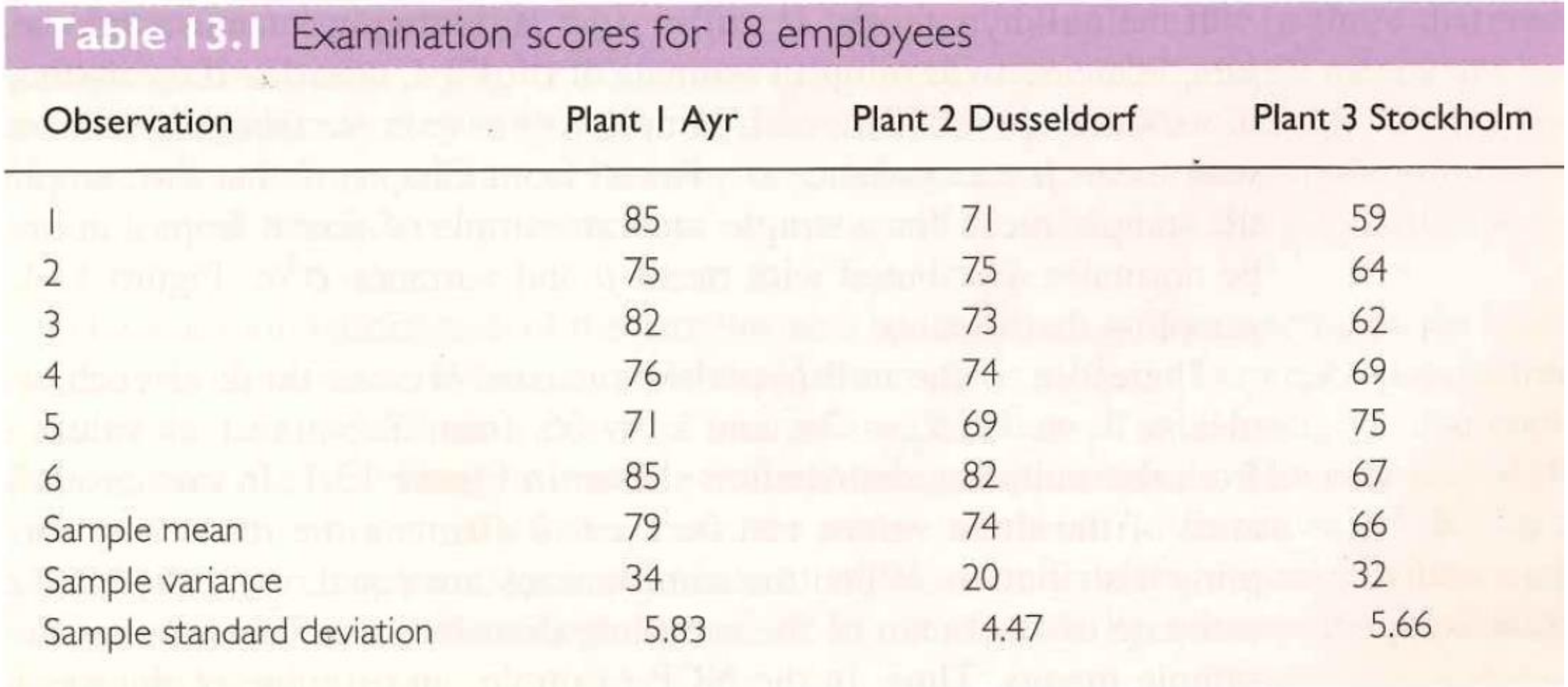
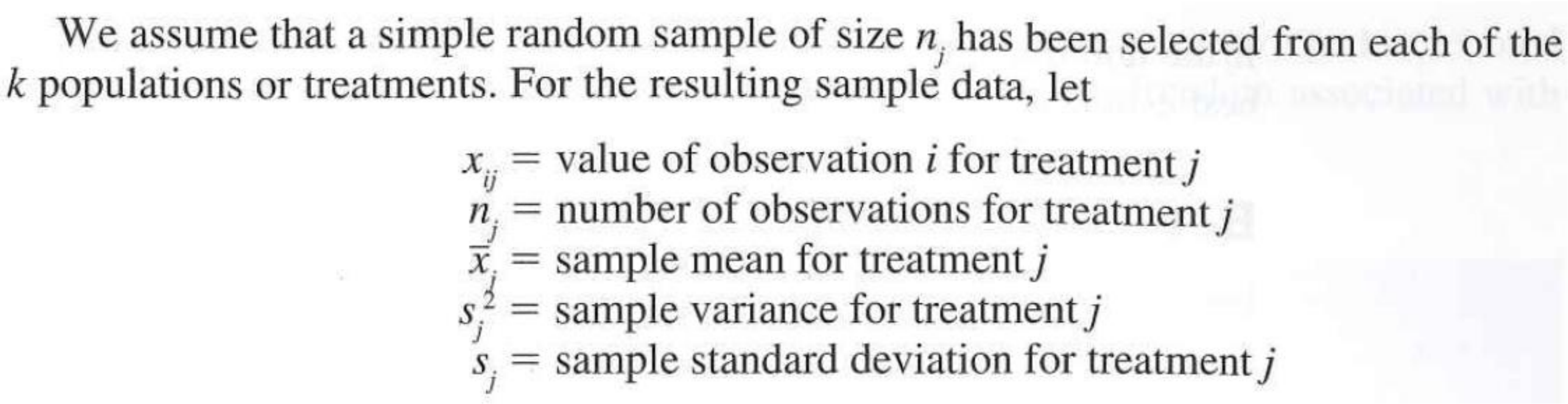


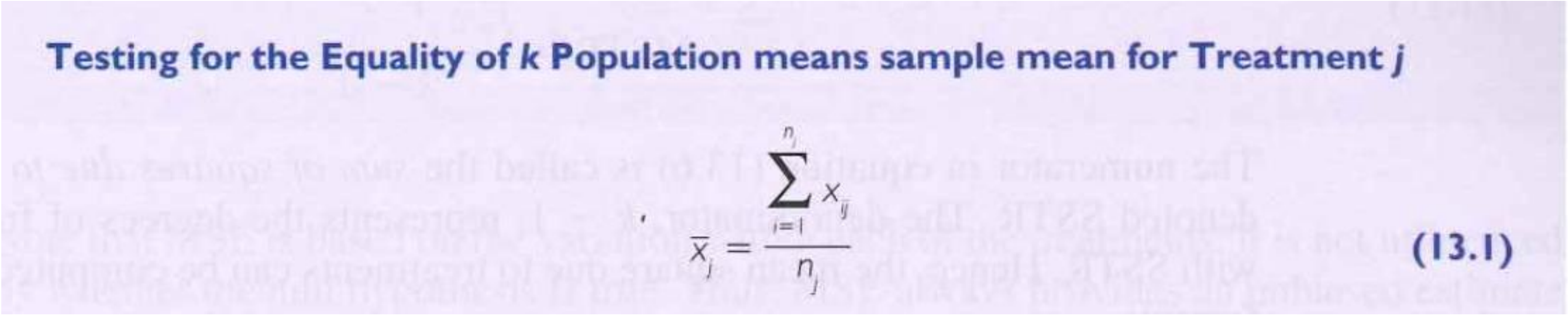
Table 13.1 Examination scores for 18 employees

Observation	Plant 1 Ayr	Plant 2 Dusseldorf	Plant 3 Stockholm
1	85	71	59
2	75	75	64
3	82	73	62
4	76	74	69
5	71	69	75
6	85	82	67
Sample mean	79	74	66
Sample variance	34	20	32
Sample standard deviation	5.83	4.47	5.66



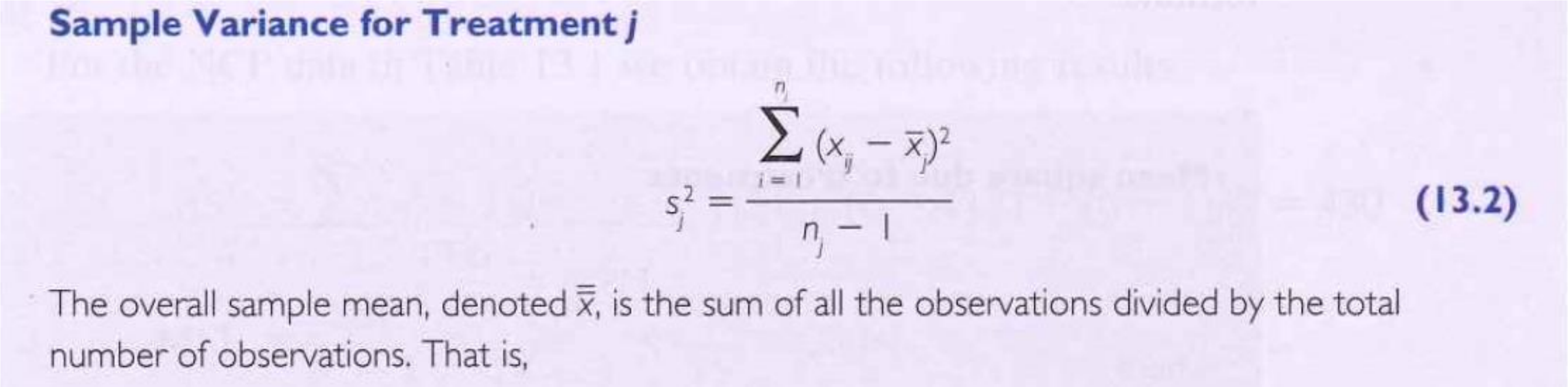
We assume that a simple random sample of size n_j has been selected from each of the k populations or treatments. For the resulting sample data, let

x_{ij} = value of observation i for treatment j
n_j = number of observations for treatment j
$\bar{x}_j$ = sample mean for treatment j
s_j^2 = sample variance for treatment j
s_j = sample standard deviation for treatment j



Testing for the Equality of *k* Population means sample mean for Treatment *j*

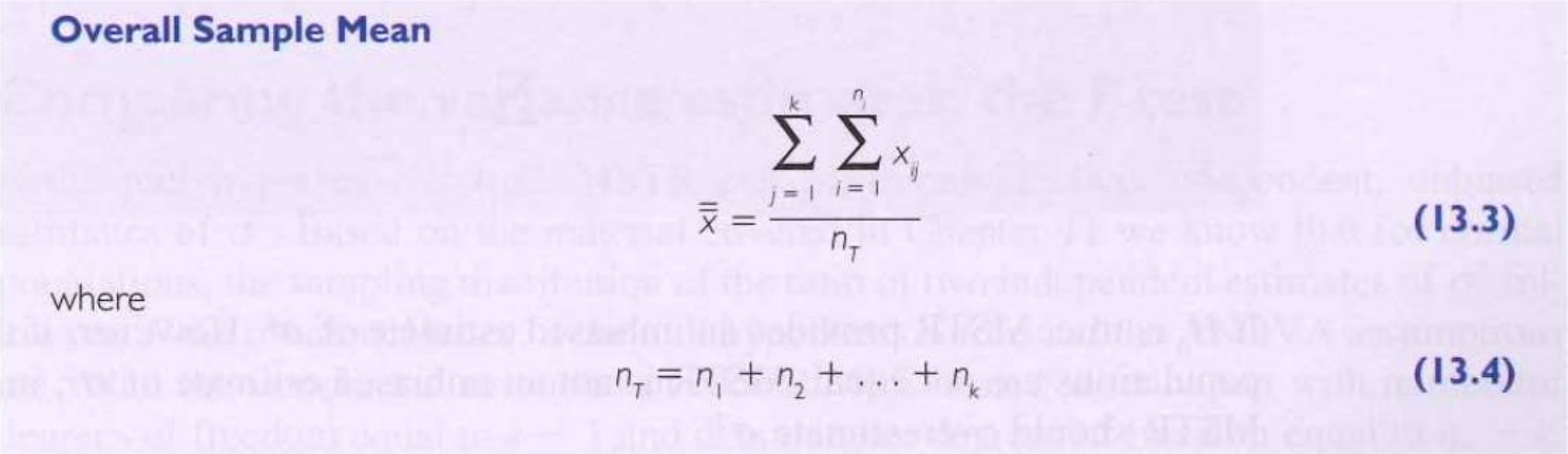
$$\bar{x}_j = \frac{\sum_{i=1}^{n_j} x_{ij}}{n_j} \qquad \textbf{(13.1)}$$



Sample Variance for Treatment *j*

$$s_j^2 = \frac{\sum_{i=1}^{n_j} (x_{ij} - \bar{x}_j)^2}{n_j - 1} \qquad \textbf{(13.2)}$$

The overall sample mean, denoted $\bar{\bar{x}}$, is the sum of all the observations divided by the total number of observations. That is,



Overall Sample Mean

$$\bar{\bar{x}} = \frac{\sum_{j=1}^{k} \sum_{i=1}^{n_j} x_{ij}}{n_T} \qquad \textbf{(13.3)}$$

where

$$n_T = n_1 + n_2 + \ldots + n_k \qquad \textbf{(13.4)}$$

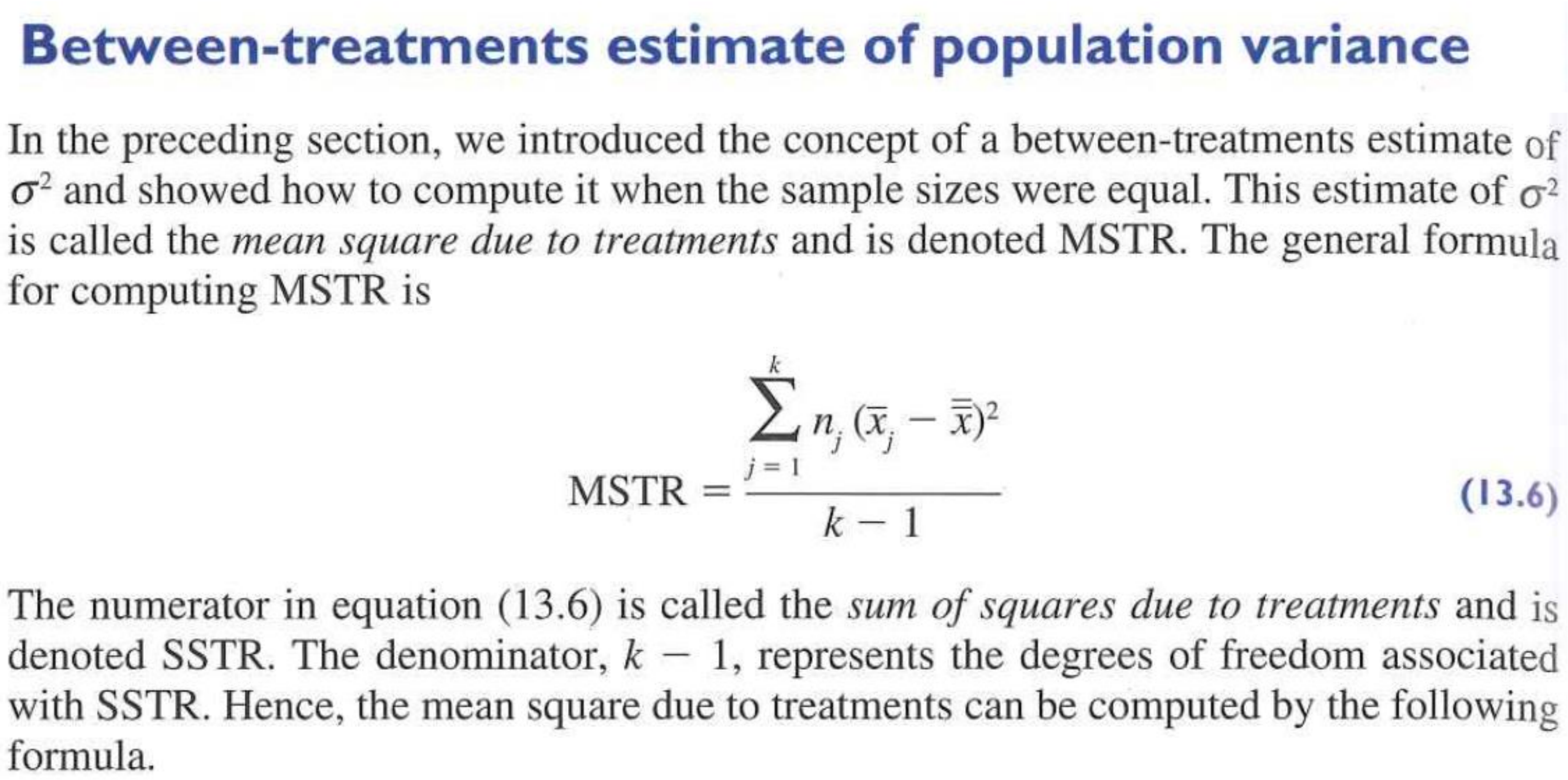
If the size of each sample is n, $n_T = kn$; in this case equation (13.3) reduces to

$$\bar{\bar{x}} = \frac{\sum_{j=1}^{k}\sum_{i=1}^{n_j} x_{ij}}{kn} = \frac{\sum_{j=1}^{k}\sum_{i=1}^{n_j} x_i / n}{k} = \frac{\sum_{j=1}^{k} \bar{x}_j}{k} \tag{13.5}$$

In other words, whenever the sample sizes are the same, the overall sample mean is just the average of the k sample means.

Because each sample in the NCP example consists of $n = 6$ observations, the overall sample mean can be computed by using equation (13.5). For the data in Table 13.1 we obtained the following result.

$$\bar{\bar{x}} = \frac{79 + 74 + 66}{3} = 73$$



Between-treatments estimate of population variance

In the preceding section, we introduced the concept of a between-treatments estimate of σ^2 and showed how to compute it when the sample sizes were equal. This estimate of σ^2 is called the *mean square due to treatments* and is denoted MSTR. The general formula for computing MSTR is

$$\text{MSTR} = \frac{\sum_{j=1}^{k} n_j (\bar{x}_j - \bar{\bar{x}})^2}{k - 1} \qquad \textbf{(13.6)}$$

The numerator in equation (13.6) is called the *sum of squares due to treatments* and is denoted SSTR. The denominator, $k - 1$, represents the degrees of freedom associated with SSTR. Hence, the mean square due to treatments can be computed by the following formula.

Mean square due to treatments

$$\text{MSTR} = \frac{\text{SSTR}}{k - 1} \qquad \textbf{(13.7)}$$

where

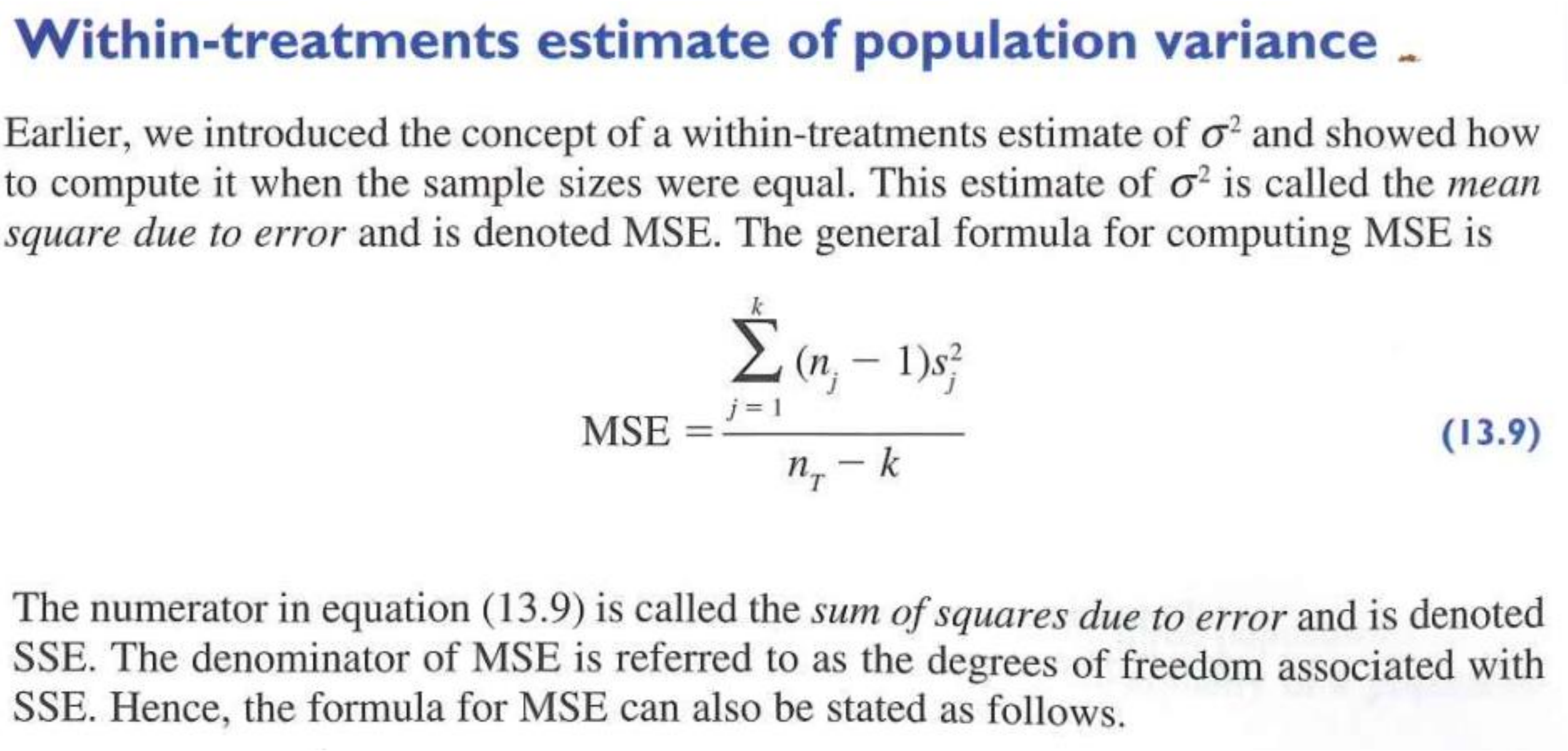
$$\text{SSTR} = \sum_{j=1}^{k} n_j (\bar{x}_j - \bar{\bar{x}})^2 \qquad \textbf{(13.8)}$$

If H_0 is true, MSTR provides an unbiased estimate of σ^2. However, if the means of the k populations are not equal, MSTR is not an unbiased estimate of σ^2; in fact, in that case, MSTR should overestimate σ^2.

For the NCP data in Table 13.1, we obtain the following results.

$$\text{SSTR} = \sum_{j=1}^{k} n_j (\bar{x}_j - \bar{\bar{x}})^2 = 6(79 - 73)^2 + 6(74 - 73)^2 + 6(66 - 73)^2 = 516$$

$$\text{MSTR} = \frac{\text{SSTR}}{k - 1} = \frac{516}{2} = 258$$

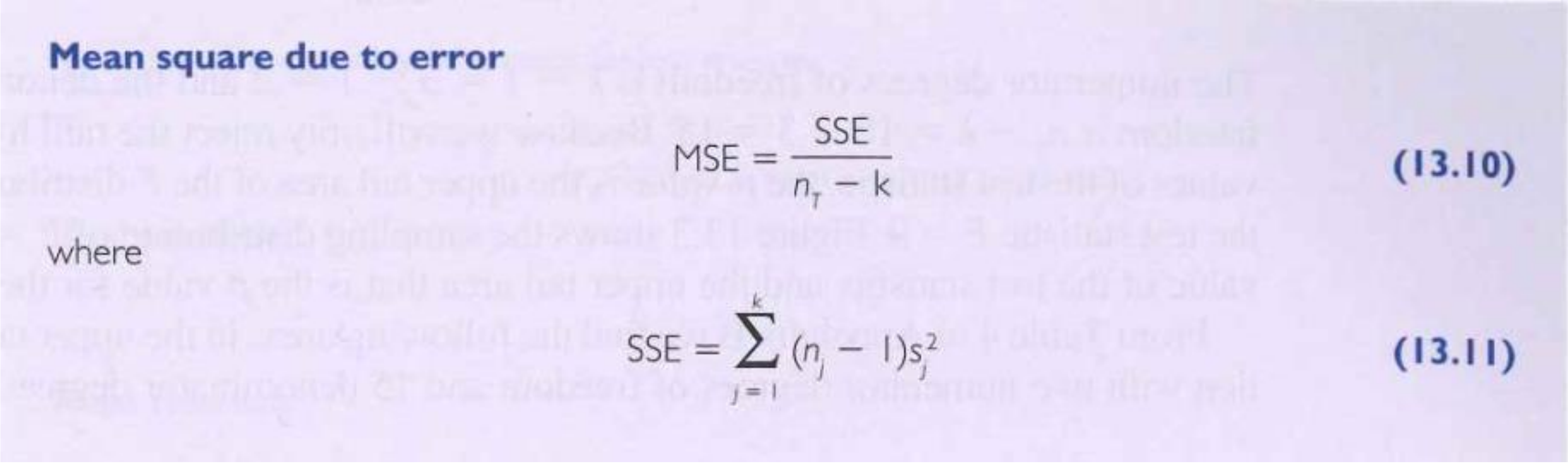


Within-treatments estimate of population variance

Earlier, we introduced the concept of a within-treatments estimate of σ^2 and showed how to compute it when the sample sizes were equal. This estimate of σ^2 is called the *mean square due to error* and is denoted MSE. The general formula for computing MSE is

$$\text{MSE} = \frac{\sum_{j=1}^{k} (n_j - 1)s_j^2}{n_T - k} \qquad \textbf{(13.9)}$$

The numerator in equation (13.9) is called the *sum of squares due to error* and is denoted SSE. The denominator of MSE is referred to as the degrees of freedom associated with SSE. Hence, the formula for MSE can also be stated as follows.

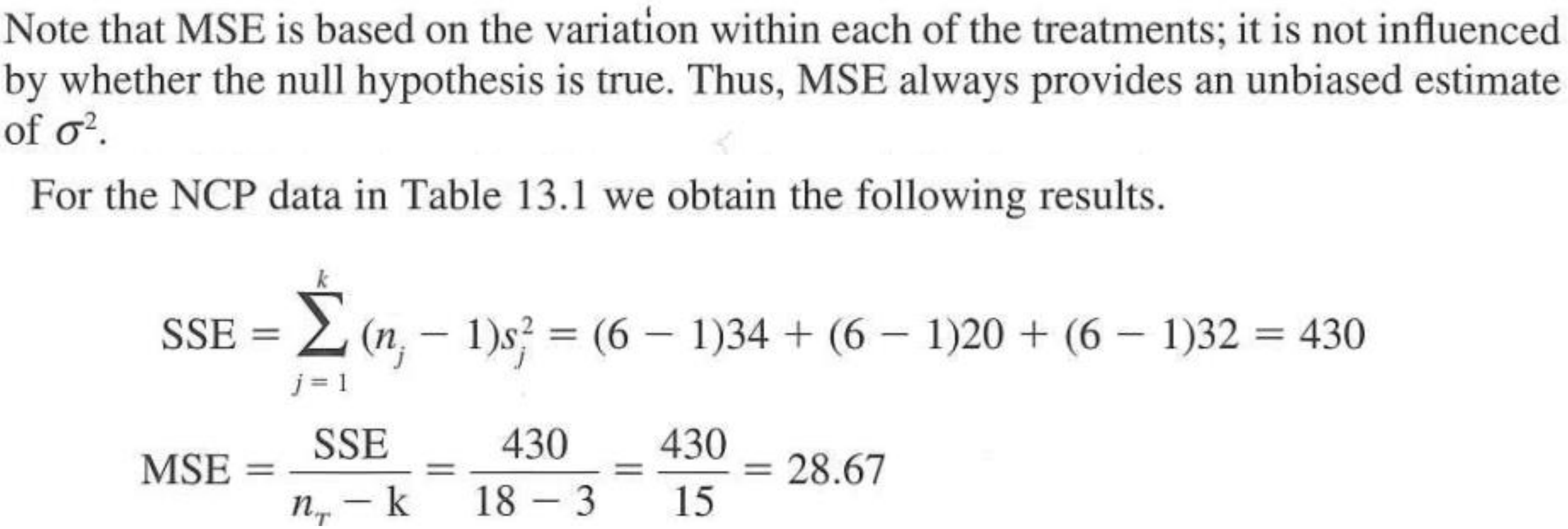


Mean square due to error

$$\text{MSE} = \frac{\text{SSE}}{n_T - k} \qquad \textbf{(13.10)}$$

where

$$\text{SSE} = \sum_{j=1}^{k} (n_j - 1)s_j^2 \qquad \textbf{(13.11)}$$



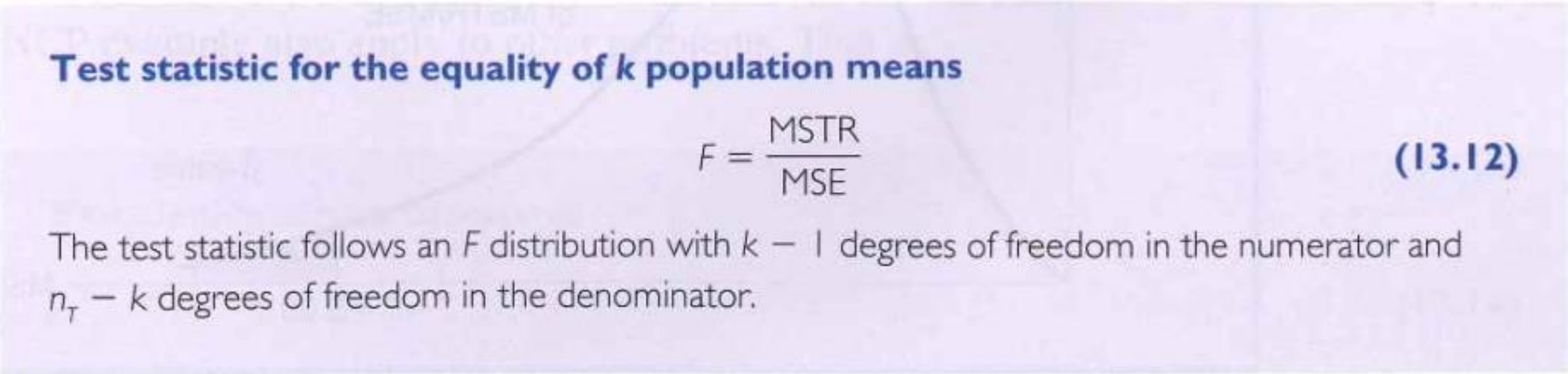
Note that MSE is based on the variation within each of the treatments; it is not influenced by whether the null hypothesis is true. Thus, MSE always provides an unbiased estimate of σ^2.

For the NCP data in Table 13.1 we obtain the following results.

$$\text{SSE} = \sum_{j=1}^{k} (n_j - 1)s_j^2 = (6-1)34 + (6-1)20 + (6-1)32 = 430$$

$$\text{MSE} = \frac{\text{SSE}}{n_T - \text{k}} = \frac{430}{18-3} = \frac{430}{15} = 28.67$$

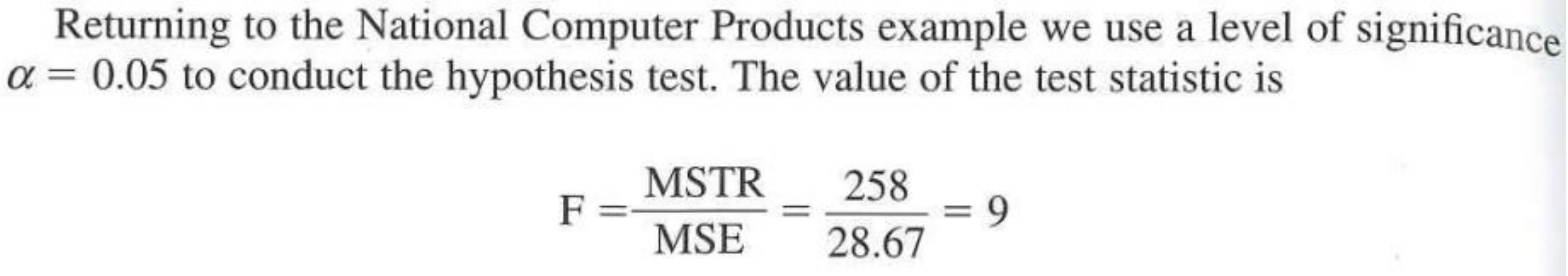
Comparing the variance estimates: the *F* test



Test statistic for the equality of *k* population means

$$F = \frac{\text{MSTR}}{\text{MSE}} \qquad \textbf{(13.12)}$$

The test statistic follows an *F* distribution with $k - 1$ degrees of freedom in the numerator and $n_T - k$ degrees of freedom in the denominator.

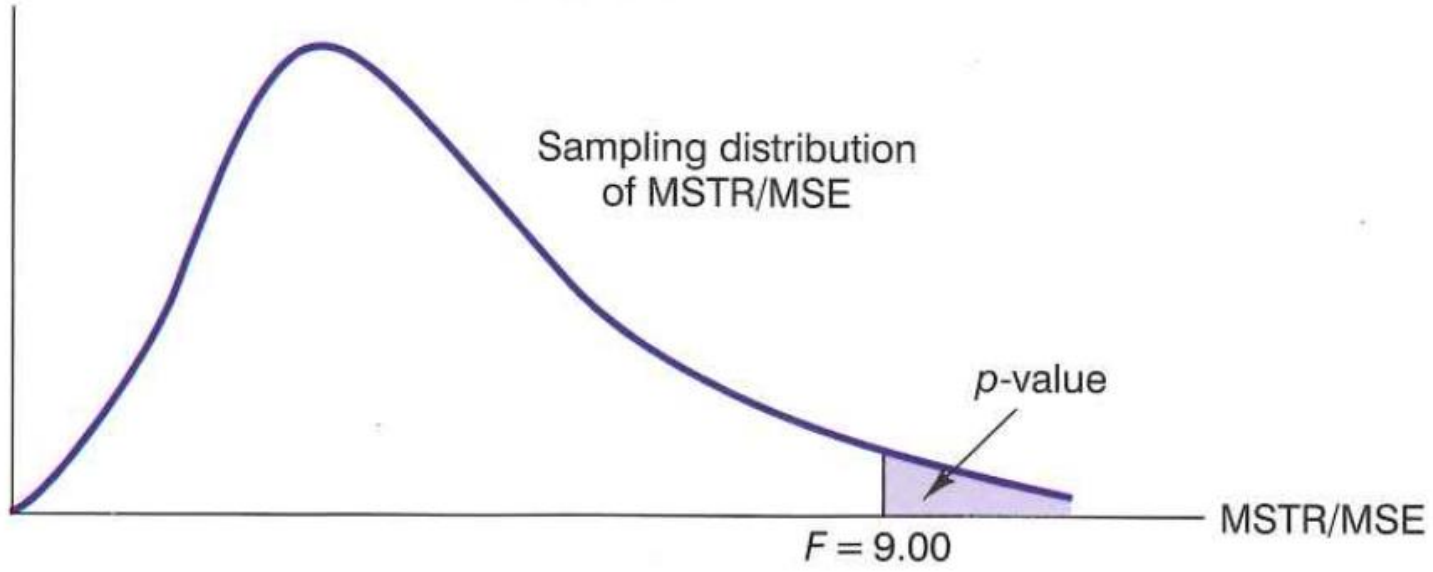


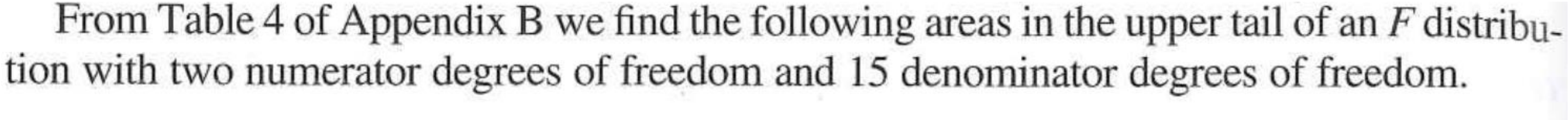
Returning to the National Computer Products example we use a level of significance $\alpha = 0.05$ to conduct the hypothesis test. The value of the test statistic is

$$\text{F} = \frac{\text{MSTR}}{\text{MSE}} = \frac{258}{28.67} = 9$$

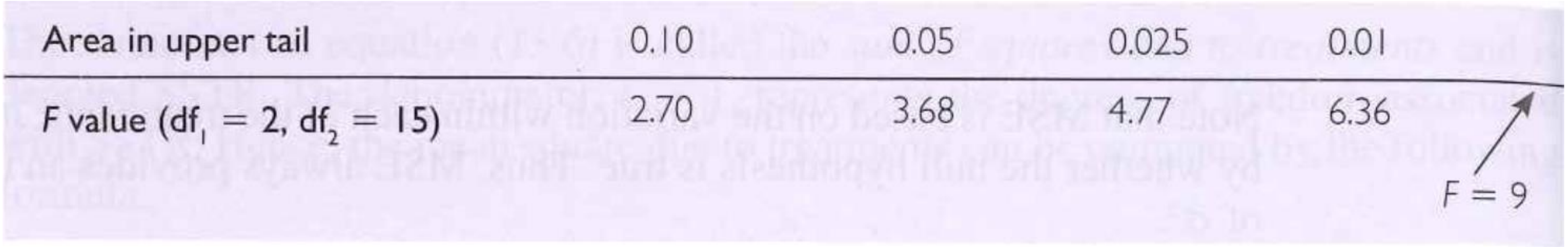
The numerator degrees of freedom is $k - 1 = 3 - 1 = 2$ and the denominator degrees of freedom is $n_T - k = 18 - 3 = 15$. Because we will only reject the null hypothesis for large values of the test statistic, the p-value is the upper tail area of the F distribution to the right of the test statistic $F = 9$. Figure 13.3 shows the sampling distribution of F = MSTR/MSE, the value of the test statistic, and the upper tail area that is the p-value for the hypothesis test.

Figure 13.3 Computation of p-value using the sampling distribution of MSTR/MSE



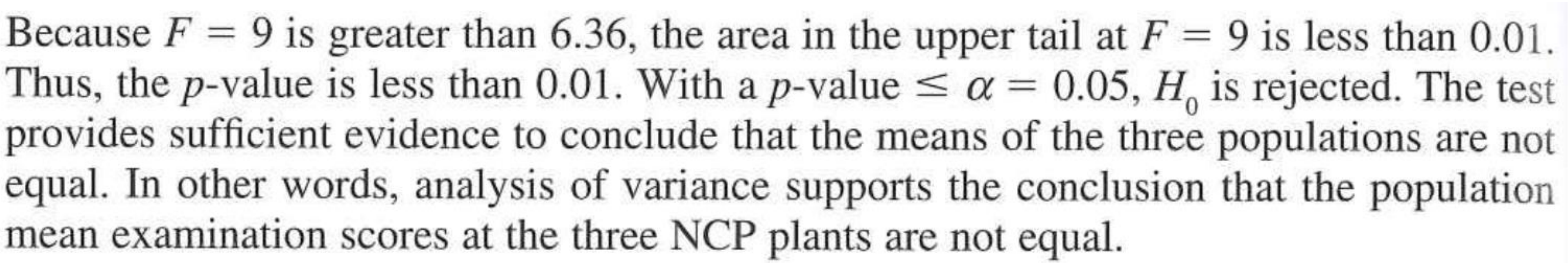


From Table 4 of Appendix B we find the following areas in the upper tail of an F distribution with two numerator degrees of freedom and 15 denominator degrees of freedom.

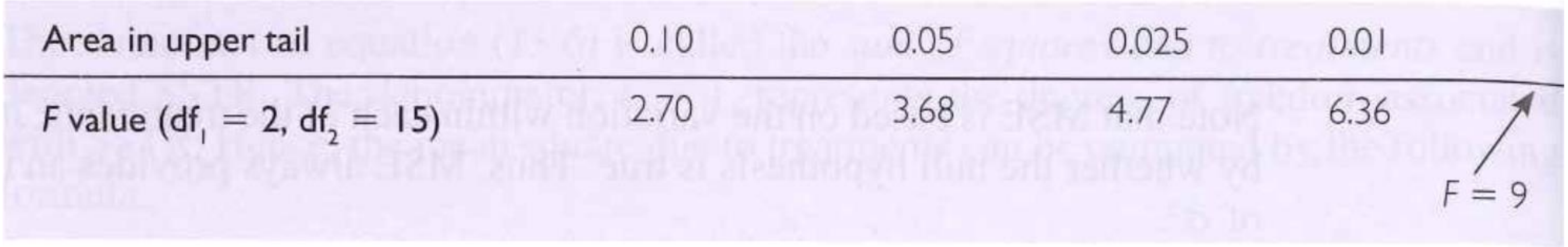


Area in upper tail	0.10	0.05	0.025	0.01
F value ($df_1 = 2$, $df_2 = 15$)	2.70	3.68	4.77	6.36

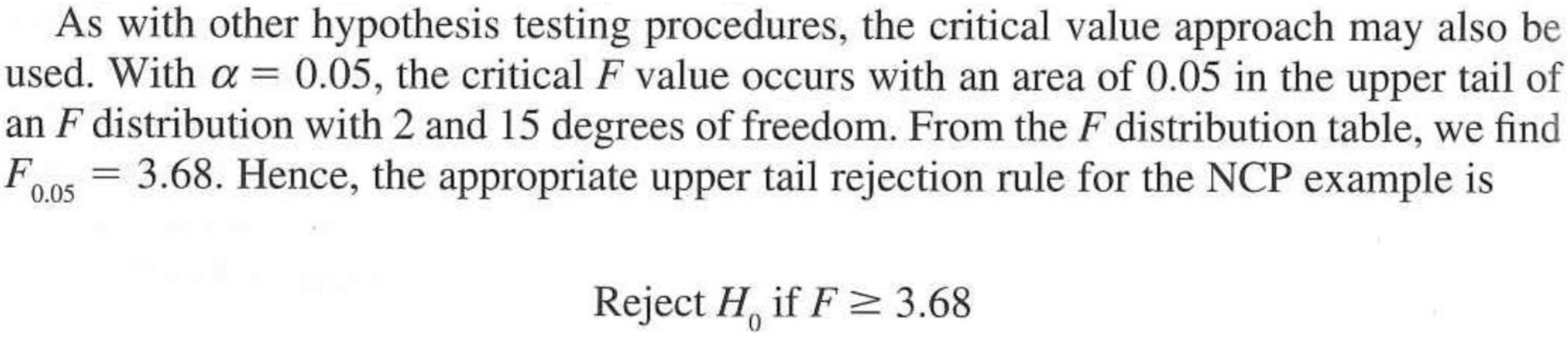
$F = 9$



Because $F = 9$ is greater than 6.36, the area in the upper tail at $F = 9$ is less than 0.01. Thus, the p-value is less than 0.01. With a p-value $\leq \alpha = 0.05$, H_0 is rejected. The test provides sufficient evidence to conclude that the means of the three populations are not equal. In other words, analysis of variance supports the conclusion that the population mean examination scores at the three NCP plants are not equal.

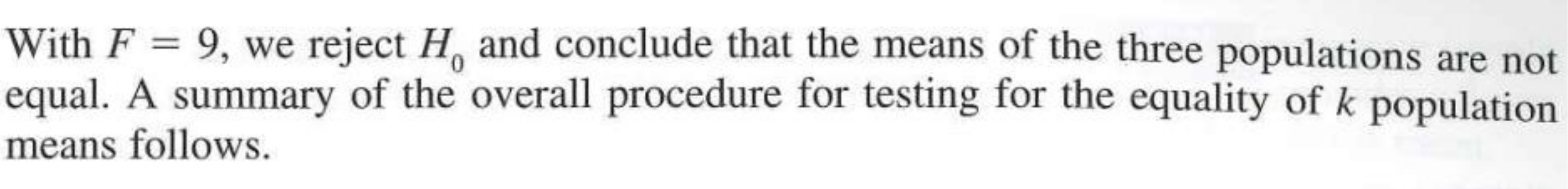


Area in upper tail	0.10	0.05	0.025	0.01
F value ($df_1 = 2$, $df_2 = 15$)	2.70	3.68	4.77	6.36

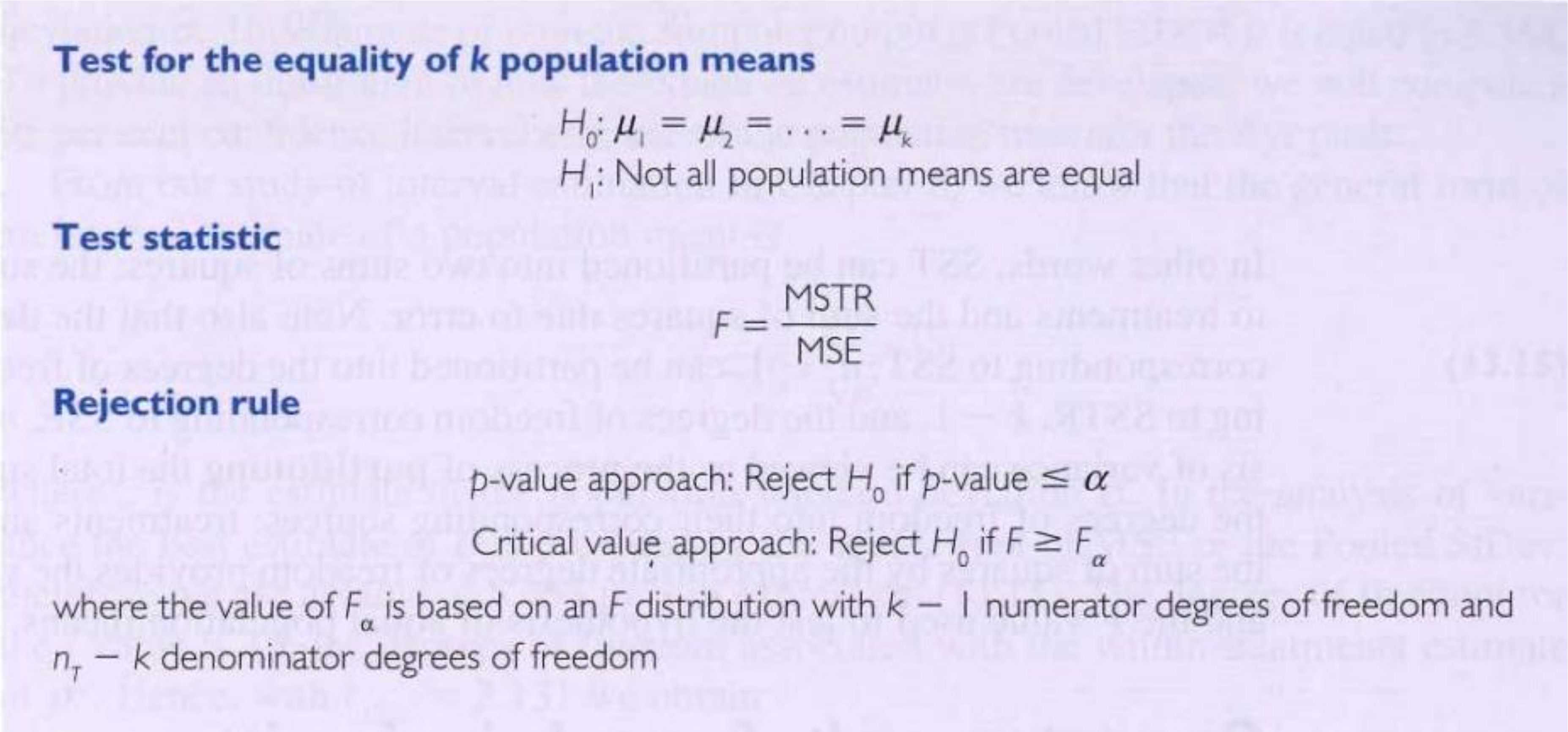


As with other hypothesis testing procedures, the critical value approach may also be used. With $\alpha = 0.05$, the critical F value occurs with an area of 0.05 in the upper tail of an F distribution with 2 and 15 degrees of freedom. From the F distribution table, we find $F_{0.05} = 3.68$. Hence, the appropriate upper tail rejection rule for the NCP example is

$$\text{Reject } H_0 \text{ if } F \geq 3.68$$



With $F = 9$, we reject H_0 and conclude that the means of the three populations are not equal. A summary of the overall procedure for testing for the equality of k population means follows.



Test for the equality of k population means

$$H_0: \mu_1 = \mu_2 = \ldots = \mu_k$$
$$H_1: \text{Not all population means are equal}$$

Test statistic

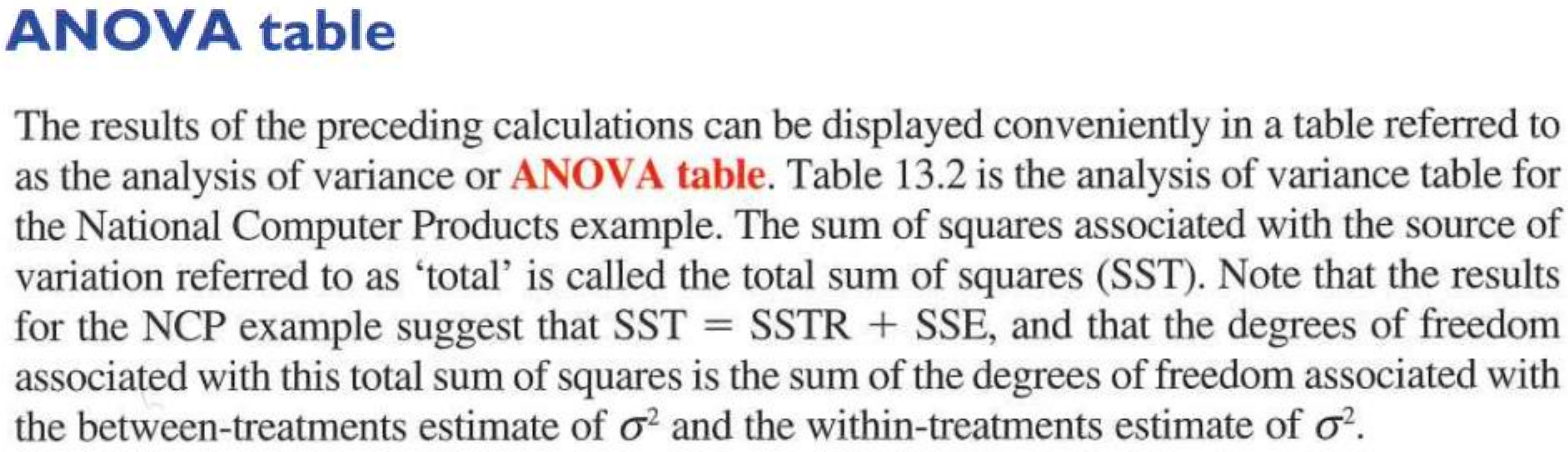
$$F = \frac{\text{MSTR}}{\text{MSE}}$$

Rejection rule

p-value approach: Reject H_0 if p-value $\leq \alpha$

Critical value approach: Reject H_0 if $F \geq F_\alpha$

where the value of F_α is based on an F distribution with $k - 1$ numerator degrees of freedom and $n_T - k$ denominator degrees of freedom



ANOVA table

The results of the preceding calculations can be displayed conveniently in a table referred to as the analysis of variance or **ANOVA table**. Table 13.2 is the analysis of variance table for the National Computer Products example. The sum of squares associated with the source of variation referred to as 'total' is called the total sum of squares (SST). Note that the results for the NCP example suggest that SST = SSTR + SSE, and that the degrees of freedom associated with this total sum of squares is the sum of the degrees of freedom associated with the between-treatments estimate of σ^2 and the within-treatments estimate of σ^2.

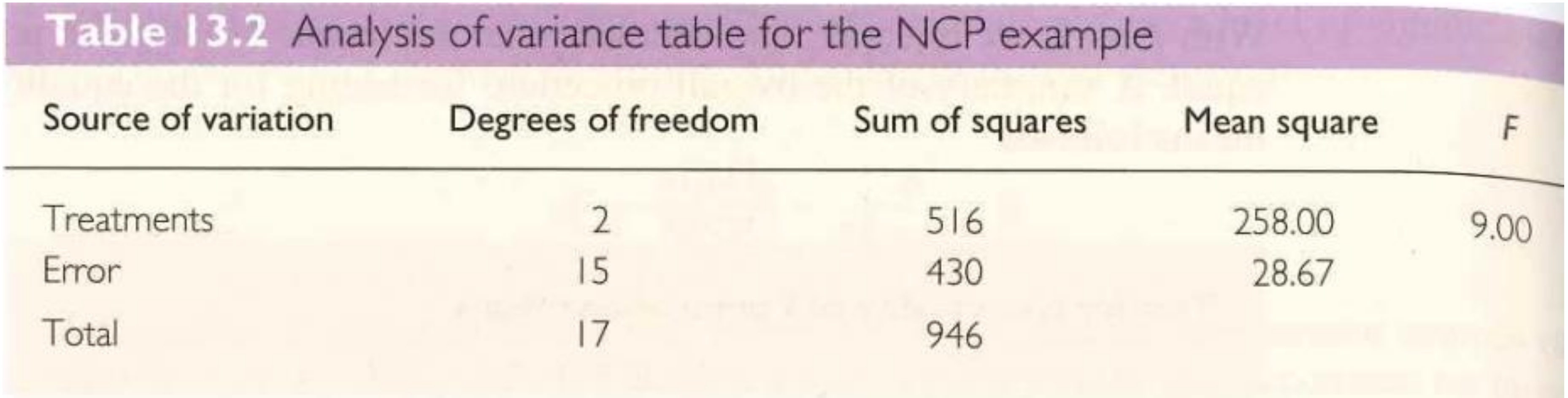
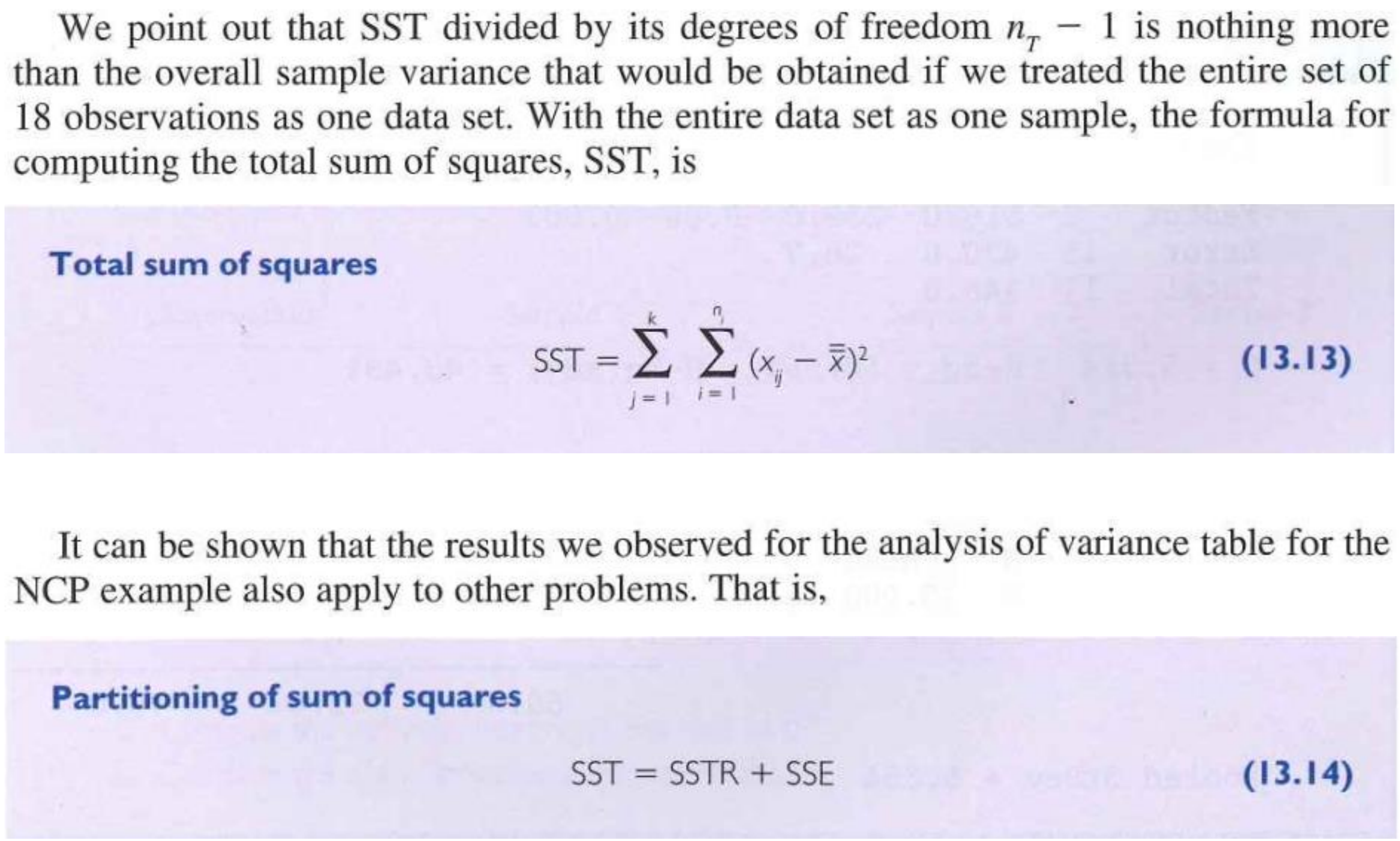


Table 13.2 Analysis of variance table for the NCP example

Source of variation	Degrees of freedom	Sum of squares	Mean square	F
Treatments	2	516	258.00	9.00
Error	15	430	28.67	
Total	17	946		



We point out that SST divided by its degrees of freedom $n_T - 1$ is nothing more than the overall sample variance that would be obtained if we treated the entire set of 18 observations as one data set. With the entire data set as one sample, the formula for computing the total sum of squares, SST, is

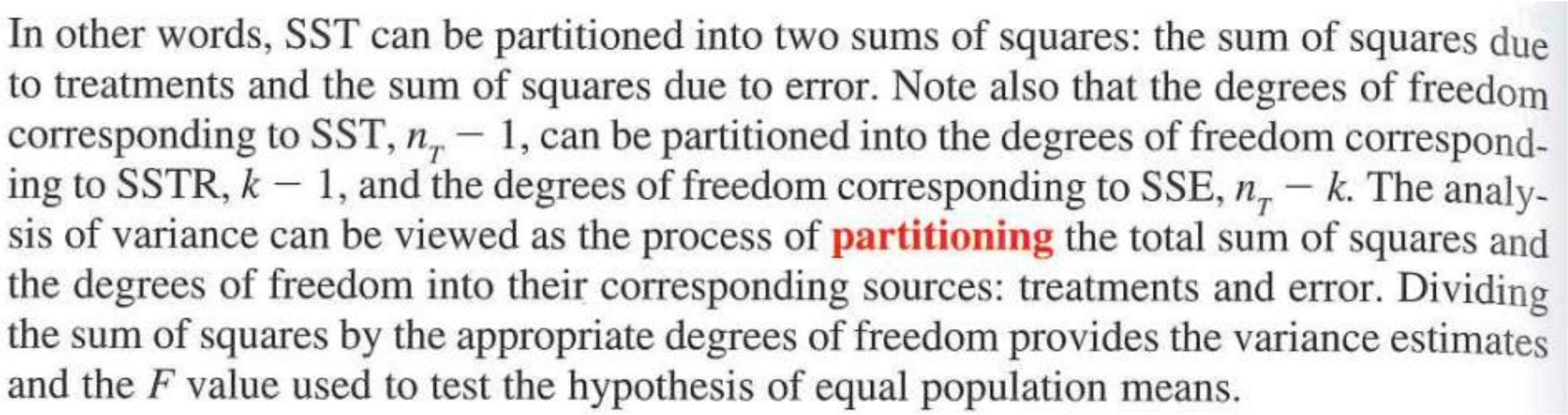
Total sum of squares

$$SST = \sum_{j=1}^{k} \sum_{i=1}^{n_j} (x_{ij} - \bar{\bar{x}})^2 \tag{13.13}$$

It can be shown that the results we observed for the analysis of variance table for the NCP example also apply to other problems. That is,

Partitioning of sum of squares

$$SST = SSTR + SSE \tag{13.14}$$



In other words, SST can be partitioned into two sums of squares: the sum of squares due to treatments and the sum of squares due to error. Note also that the degrees of freedom corresponding to SST, $n_T - 1$, can be partitioned into the degrees of freedom corresponding to SSTR, $k - 1$, and the degrees of freedom corresponding to SSE, $n_T - k$. The analysis of variance can be viewed as the process of **partitioning** the total sum of squares and the degrees of freedom into their corresponding sources: treatments and error. Dividing the sum of squares by the appropriate degrees of freedom provides the variance estimates and the F value used to test the hypothesis of equal population means.

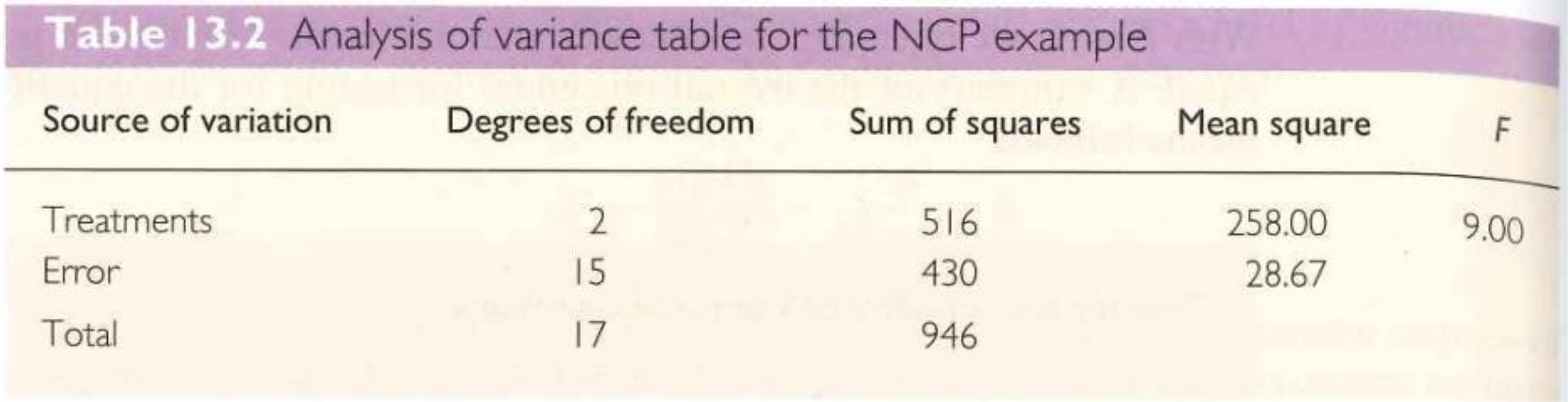


Table 13.2 Analysis of variance table for the NCP example

Source of variation	Degrees of freedom	Sum of squares	Mean square	F
Treatments	2	516	258.00	9.00
Error	15	430	28.67	
Total	17	946		