COMPUTATIONAL FORMULAS FOR ANOVA

The deviation formulas for the sums of squares—Equations (14.9), (14.11), and (14.12)—are perfectly adequate methods for computing the analysis of variance, whether or not the \( n_j \) are equal. However, the ANOVA deviation formulas can be somewhat cumbersome to compute when data sets are large. We present here computational formulas that are somewhat easier because they do not require the calculation of deviations for each point. The alert reader will recognize that the present situation is entirely analogous to that encountered in Chapter 5, when for the same reason we presented mean square deviation and computational formulas for the standard deviation and variance.

These computational formulas for the sums of square are applicable whether or not the \( n_j \) are equal. The total sum of squares is

\[
SS_T = \sum \frac{X^2}{n_G} - \left( \frac{\sum X}{n_G} \right)^2 \tag{14.9a}
\]

where \( n_G \) is the total number of observations: \( n_G = n_1 + n_2 + \ldots + n_k \). The sum of squares between groups is

\[
SS_B = \left[ \left( \frac{\sum X^2_1}{n_1} \right)^2 + \left( \frac{\sum X^2_2}{n_2} \right)^2 + \cdots + \left( \frac{\sum X^2_k}{n_k} \right)^2 \right] - \left( \frac{\sum X}{n_G} \right)^2 \tag{14.12a}
\]

Note that the last terms in Equations (14.9a) and (14.12a) are identical and so need to be computed only once.

The sum of squares within groups is

\[
SS_W = \sum \frac{X^2}{n} - \left[ \frac{\left( \sum X^2_1 \right)}{n_1} + \frac{\left( \sum X^2_2 \right)}{n_2} + \cdots + \frac{\left( \sum X^2_k \right)}{n_k} \right] \tag{14.11a}
\]

Note that the terms in brackets in Equations (14.12a) and (14.11a) are identical, so they need to be computed only once. Similarly, the first terms in Equations (14.9a) and (14.11a) are identical.

We use Equation (14.10b) to verify our computations:

\[
SS_T = SS_W + SS_B
\]
Table 1 summarizes the computation using the given formulas for the antidepressant drug data from Table 14.1. Note that we get identical results to those we obtained in the text.

**Table 1** Using the computational formulas for ANOVA on the antidepressant drug data

<table>
<thead>
<tr>
<th>$X_{i1}$</th>
<th>$X_{i1}^2$</th>
<th>$X_{i2}$</th>
<th>$X_{i2}^2$</th>
<th>$X_{i3}$</th>
<th>$X_{i3}^2$</th>
<th>$X_{i4}$</th>
<th>$X_{i4}^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>29</td>
<td>841</td>
<td>35</td>
<td>1225</td>
<td>23</td>
<td>529</td>
<td>33</td>
<td>1089</td>
</tr>
<tr>
<td>32</td>
<td>1024</td>
<td>34</td>
<td>1156</td>
<td>25</td>
<td>625</td>
<td>36</td>
<td>1296</td>
</tr>
<tr>
<td>26</td>
<td>676</td>
<td>28</td>
<td>784</td>
<td>26</td>
<td>676</td>
<td>30</td>
<td>900</td>
</tr>
<tr>
<td>28</td>
<td>784</td>
<td>32</td>
<td>1024</td>
<td>26</td>
<td>676</td>
<td>32</td>
<td>1024</td>
</tr>
<tr>
<td>25</td>
<td>625</td>
<td>37</td>
<td>1369</td>
<td>25</td>
<td>625</td>
<td>32</td>
<td>1024</td>
</tr>
<tr>
<td>28</td>
<td>784</td>
<td>35</td>
<td>1225</td>
<td>19</td>
<td>361</td>
<td>32</td>
<td>1024</td>
</tr>
<tr>
<td>168</td>
<td>4734</td>
<td>201</td>
<td>6783</td>
<td>144</td>
<td>3492</td>
<td>195</td>
<td>6357</td>
</tr>
</tbody>
</table>

$\sum X_{ij} = 168 + 201 + 144 + 195 = 708$

$\sum X_{ij}^2 = 4734 + 6783 + 3492 + 6357 = 21,366$

$$SS_T = \sum X_{ij}^2 - \frac{(\sum X_{ij})^2}{n_G} = 21,366 - \frac{(708)^2}{24} = 21,366 - 20,886 = 480.00$$

$$SS_B = \left[ \frac{(\sum X_{i1})^2}{n_1} + \frac{(\sum X_{i2})^2}{n_2} + \cdots + \frac{(\sum X_{iG})^2}{n_G} \right] - \frac{(\sum X_{ij})^2}{n_G}$$

$$= \left[ \frac{(168)^2}{6} + \frac{(201)^2}{6} + \frac{(144)^2}{6} + \frac{(195)^2}{6} \right] - \frac{(708)^2}{24}$$

$$= (4704.0 + 6733.5 + 3456.0 + 6337.5) - 20,886.0$$

$$= 21,231.0 - 20,886.0 = 345.00$$

$$SS_W = \sum X_{ij}^2 - \left[ \frac{(\sum X_{i1})^2}{n_1} + \frac{(\sum X_{i2})^2}{n_2} + \cdots + \frac{(\sum X_{iG})^2}{n_G} \right]$$

$$= 21,366 - 21,231.0 = 135.00$$

Check: $SS_T = 480.00 = SS_W + SS_B = 135.00 + 345.00$

**EXERCISE**

1. Use the computational formulas to compute the ANOVA for the chimpanzee grooming data of the textbook’s Table 14.6. Show that the computational formulas give the same results shown in Table 14.7.