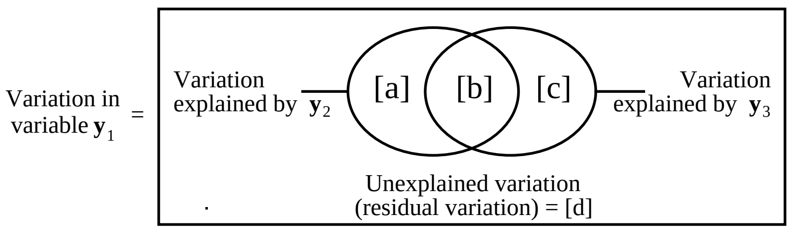
Variation partitioning **Box 4.1**

Variation partitioning, which is presented in detail in Subsection 10.3.5, provides a general framework to illustrate the similarities and difference between the coefficient of multiple determination, the partial and semipartial correlation coefficients, and the corresponding F-statistics.

Three variables only, $\mathbf{y}_1$, $\mathbf{y}_2$, and $\mathbf{y}_3$, are considered in this example. In the following Venn diagram, the rectangle represents the total sum of squares of variable $\mathbf{y}_1$:



In the multiple regression of $\mathbf{y}_1$ on $\mathbf{y}_2$ and $\mathbf{y}_3$, $\hat{y}_1 = b_0 + b_2 y_2 + b_3 y_3$ (this is an application of eq. 10.15), the coefficient of multiple determination, which is the square of the coefficient of multiple correlation, is:

$$R^2_{1.23} = \frac{[\mathrm{a+b+c}]}{[\mathrm{a+b+c+d}]} \qquad \text{with} \quad F = \frac{[\mathrm{a+b+c}]/2}{[\mathrm{d}]/(n-3)}$$

The partial correlation of $\mathbf{y}_1$ with $\mathbf{y}_2$ while controlling for the effect of $\mathbf{y}_3$ is:

$$r_{12.3} = \sqrt{\frac{[\mathrm{a}]}{[\mathrm{a+d}]}} \qquad \text{with} \quad F = \frac{[\mathrm{a}]/1}{[\mathrm{d}]/(n-3)}$$

The semipartial correlation of $\mathbf{y}_1$ with $\mathbf{y}_2$ in the presence of $\mathbf{y}_3$ is:

$$r_{1(2.3)} = \sqrt{\frac{[\mathrm{a}]}{[\mathrm{a+b+c+d}]}} \qquad \text{with} \quad F = \frac{[\mathrm{a}]/1}{[\mathrm{d}]/(n-3)}$$

The coefficients of partial and semipartial correlation receive the same sign as the corresponding coefficient of partial regression.

The test of a partial regression coefficient, b_2 or b_3, is the same (i.e., it has the same F-statistic) as the test of the corresponding partial correlation coefficient, $r_{12.3}$ or $r_{13.2}$. The F-statistic is always the ratio of two *independent* portions of the variation of $\mathbf{y}_1$, each one divided by its degrees of freedom; see eqs. 4.39 and 4.40.