Unbalanced two-factor ANOVA

The term "unbalanced" means that the sample sizes n_{ij} are not all equal. A balanced design is one in which all $n_{ij} = n$.

In the unbalanced case, there are two ways to define sums of squares for factors A and B.

The method SAS calls Type III sums of squares goes by the name "partial SS" or "adjusted SS." It's the default in many programs, like Stata. For this method, here are the full and reduced models being tested for Factor A, Factor B, and the AB interaction:

(I've left the subscripts off in all of the following model statements. They should be obvious.)

Factor A:
Full model is $Y = \mu + \alpha + \beta + \alpha\beta + \varepsilon$
Reduced model is $Y = \mu + \beta + \alpha\beta + \varepsilon$

Factor B:
Full model is $Y = \mu + \alpha + \beta + \alpha\beta + \varepsilon$
Reduced model is $Y = \mu + \alpha + \alpha\beta + \varepsilon$

AB interaction:
Full model is $Y = \mu + \alpha + \beta + \alpha\beta + \varepsilon$
Reduced model is $Y = \mu + \alpha + \beta + \varepsilon$

The method SAS calls Type I sums of squares is usually called sequential sums of squares. It is the default in R. To get it in programs like Stata and Minitab, you need to ask for the "sequential" sums of squares.

Factor A:
Full model is $Y = \mu + \alpha + \varepsilon$
Reduced model is $Y = \mu + \varepsilon$

Factor B:
Full model is $Y = \mu + \alpha + \beta + \varepsilon$
Reduced model is $Y = \mu + \alpha + \varepsilon$

AB interaction: Same as above.

For Type I SS, notice that it matters what order you use to name the factors, whereas it doesn't matter for Type III SS.