Anova versus anova in R

Summary: For anything beyond one-way ANOVA, **we are generally interested in the Anova(, type = 3) results because the tests are reasonable even with unbalanced data and/or significant interactions.** Type 3 tests are the default for most statistical software except R. When calculating Type3 tests for a multi-way ANOVA model (multiple factors plus interaction(s)), it is necessary to set the contrasts option in R before running the Type 3 tests.

The anova() function is part of the stats package (base R). When applied to an individual lm object, the anova() function will produce a sequential (Type 1) ANOVA table. The resulting table will show tests produced by fitting a sequence of models to the data. The results depend on the order the variables are listed.

The Anova() function is part of the car package. The Anova() function will produce unique/marginal (Type 2 or 3) ANOVA table. The resulting table will show tests for adding one of the predictors to a model that includes all the others. The results do NOT depend on the order the variables are listed.

When there is only a single predictor, Type 1, 2 and 3 tests are the same. Fine to use anova(). This is what we did in STAT511 for one-way ANOVA.

When there are multiple predictors but the data is balanced (equal number of observations per treatment combination), Type 1, 2 and 3 tests are the same.

When there are multiple predictors but no interactions, the Type2 and 3 tests are the same.

Type2 and 3 tests differ in the way they handle interactions. Most software programs (SAS, JMP, SPSS, etc) return Type 3 tests by default. However, the default for Anova() is Type 2 tests. We are generally interested in the Type 3 results because the tests are reasonable even with missing data or significant interactions.

When calculating Type3 tests for a multi-way ANOVA model (multiple factors plus interactions), it is necessary to set the contrasts option in R. Because the multi-way ANOVA model is over-parameterized, it is necessary to choose a contrasts setting that sums to zero, otherwise the ANOVA analysis will give incorrect results with respect to the expected hypothesis. The default contrasts type does not satisfy this requirement. Instead use: options(contrasts = c("contr.sum","contr.poly"))

When fitting mixed models (with fixed and random effects), we will use the lmerTest package for creating an ANOVA table. After that package has been loaded, anova(, ddf="Kenward-Roger") will calculate Type 3 tests. This is not a typo, after loading the lmerTest package, the anova() function will return Type 3 tests for mixed models.

Here are some references discussing the differences between Type 2 and 3 tests:

An R Companion to Applied Regression by Fox and Weisberg, pages 192-197 (2nd Edition).

Discovering Statistics Using R by Field, Miles and Field, pages 476-477 (1st Edition).

https://mcfromnz.wordpress.com/2011/03/02/anova-type-iiiiii-ss-explained/