Correcting for Violations of Sphericity

Box Correction

The effect of violation of assumption was clearly shown by Box (1954). For any arbitrary covariance matrix which deviates from the required form, the mean square ratio is approximately distributed as F with reduced degrees of freedom. The amount of the reduction is dictated by a multiplicative factor, ε, which depends on the population covariance matrix. When the matrix fulfills the condition noted above, ε = 1.0; otherwise, ε < 1.0, with a minimum of $\frac{1}{a-1}$. Box (1954) also showed that when sphericity fails, the number of degrees of freedom of the F_A ratio depends directly upon the degree of sphericity (*i.e., ε*) and are equal to $v_1 = \varepsilon(A - 1)$ and $v_2 = \varepsilon(A - 1)(S - 1)$.

In SPSS, three corrections are generated: the Greenhouse-Geisser (1959), the Huynh-Feldt (1976), and the lower-bound. Each of these corrections have been developed to alter the degrees of freedom and produce an F-ratio where the Type I error rate is reduced. The actual F-ratio does not change as a result of applying the corrections; only the degrees of freedom.

Lower bound correction.

This is a correction for a violation of the sphericity assumption. The correction works by using a higher F-critical value to reduce the likelihood of a Type I error. In the non-SPSS world, the lower bound correction is really referred to as the "Geisser-Greenhouse" correction in which $df_A = 1$ and $df_{AxS} = s - 1$ are used instead of the usual $df_A = a - 1$ and $df_{AxS} = (a - 1)(s - 1)$. Under this correction to the dfs, the F-critical values will be larger and it will be harder to detect significance, thus reducing Type I error. Unfortunately, this correction approach tends to overcorrect, so there are too many Type II errors with this procedure (i.e., low power). This correction assumes the maximum degree of heterogeneity among the cells.

Geisser-Greenhouse correction (1959).

The Geisser-Greenhouse correction referred to in SPSS is another variant on the procedure described under Huynh & Feldt. A slightly different correction factor (epsilon) is computed, which corrects the degrees of freedom slightly more than the Huynh & Feldt correction. So, significance tests with this correction will be a little more conservative (higher p-value) than those using the Huynh-Feldt correction.

Of the three corrections, Huynh-Feldt is considered the least conservative, while Greenhouse-Geisser is considered more conservative and the lower-bound correction is the most conservative. Girden recommended that when epsilon is > .75, the Huynh-Feldt correction should be applied and when epsilon is < .75 or nothing is known about sphericity, the Greenhouse-Geisser correction should be applied.

Huynh & Feldt Correction (1976).

This correction is based on a similar correction by Box (1954). In both cases, an adjustment factor based on the amount of variance heterogeneity (i.e., how much the variances are unequal) is computed (the adjustment factor is called epsilon). Then, both df_{AxS} and df_A are adjusted by this factor, so that the F-critical will be somewhat larger. The correction is not as severe as the "lower bound" correction.

Recommendations

- It is more often the rule than the exception that sphericity is violated in repeated measures designs. For this reason, all repeated measures designs should be exposed to tests of violations of sphericity. If sphericity is violated then the researcher must decide whether a multivariate or univariate analysis is preferred.
- If univariate methods are chosen then the omnibus ANOVA must be corrected appropriately depending on the level of departure from sphericity.
- If pairwise comparisons are required the Bonferroni method should probably be used to control the Type 1 error rate when group sizes are equal.
- Do not apply the sphericity and compound symmetry assumptions when there are only two levels (or cells) of the within-subjects factor (e.g., pre vs. post test only).
- If F is non significant, do not worry about the corrections, because the corrections will only increase the p-values).
- If the F is significant using all three approaches, do not worry about the corrections. That is, if the most conservative approach is still significant, there is no increased risk of Type I error.
- If the various correction tests lead to different conclusions, there is no one perfect solution. Prefer Huynh and Feldt correction, because (a) it addresses the Type I error problem and is more powerful than the "lower bound" correction, (b) it does not perform perfectly but it might be safest to report the results from all correction approaches when they do not show the same result.
- With large sample sizes, a small departure from sphericity might be significant, but the correction in these situations should be relatively minor and there is not likely to be difference in the conclusions drawn from the significance tests using the various corrections.

- Transform the dependent variable scores using a square root transformation or raise the score to a fractional power and use a multivariate analysis of variance (MANOVA) test. For the transformation approach, you may have to try different transformations until the sphericity problem is resolved. The transformation approach may be quite helpful in resolving the problem, but the researcher will have more difficulty in interpreting the results.

- Under some conditions, the MANOVA approach can be more powerful than the within-subjects ANOVA, and the MANOVA test does not assume sphericity. MANOVA assumes that the data have a "multivariate" normal distribution—that the analysis variable is jointly normally distributed when all levels are considered together. Algina and Kesselman (1997) suggest that for few levels and large sample sizes, the MANOVA approach may be more powerful. Their guidelines are to use MANOVA if The number of levels is:
    - less than or equal to 4 ($a < 4$) and $n > a + 15$, or
    - between 5 and 8 ($5 < a < 8$) and $n > a + 30$.