

# Multivariate Power Analysis

*Aaron R. Caldwell*

*9/21/2019*

## **Multivariate ANOVA (MANOVA)**

A large proportion of research with within-subjects manipulations, or repeated measures, rely upon the “univariate” approach (Maxwell & Delaney). While this approach is valid, when corrections for sphericity are applied, it may not be the most powerful or informative analysis plan. Instead, researchers should consider a multivariate analysis (MANOVA). While the MANOVA is not “assumption free” it does not assume sphericity which makes it a very attractive analytical tool and the preferred method of analysis for some (Maxwell and Delaney).

## Sphericity?

In all the above examples, you will notice that the correlation is the same between repeated measures. For most experiments in real life, this will not be the case. The correlations between levels may vary. This is problematic because the univariate (ANOVA) approach assumes sphericity, which means, for all intents and purposes, that the correlations between factor-levels are equal and the standard deviations at each level are equal as well. This assumption is tenuous at best and is typically “adjusted” for by applying a sphericity correction (e.g, Greenhouse-Geisser). How bad can it get? Well, let’s simulate an example below.

```
design_result <- ANOVA_design("4w",
                             n = 29,
                             r = c(.05,.15,.25,.55, .65, .9
                                    ),
                             sd = 1,
                             mu= c(0,0,0,0))

#In order to simulate violations we MUST use ANOVA_power
power_result_s1 <- ANOVA_power(design_result, nsims = nsims, verbose = FALSE)
```

```
##           power effect_size
## anova_a   7.16  0.03445248
```

### INSERT ANOVA Result ABOUT HERE

As we can see, the actual type I error rate far exceeds the typical 5%!

Now, let’s pour gasoline on the fire and see what happens when we make the sphericity violation worse by varying the standard deviations.

```
design_result <- ANOVA_design("4w",
                             n = 29,
                             r = c(.05,.15,.25,.55, .65, .9
                                    ),
                             sd = c(1,3,5,7),
                             mu= c(0,0,0,0))

#In order to simulate violations we MUST use ANOVA_power
power_result_s2 <- ANOVA_power(design_result, nsims = nsims, verbose = FALSE)
```

```
##           power effect_size
## anova_a   8.37  0.03468564
```

### INSERT ANOVA Result ABOUT HERE

These inflated error rates are obviously a problem, which begs the question what do we do about them?

In our experience, most researchers default to using a Greenhouse-Geisser adjustment for sphericity, but this may not be the most statistical efficient way of dealing with violations of sphericity. Further, as we can see from the simulation (power\_result\_s2) the MANOVA maintains the type I error rate at 5%.

### Insert MANOVA Result About HERE

## MANOVA or Sphericity Adjustment?

In addition to adjusting for sphericity, one could also simply use the multivariate approach to repeated measures. While it is tempting to simply say one approach is superior to another, that is not the case when the sample size is small (Maxwell and Delaney, ppg 775). Some general guidelines were proposed by Algina and Keselman (1997):

MANOVA when  $\text{levels} \leq 4$ ,  $\text{epsilon} \leq .9$ ,  $n > \text{levels} + 15$  and  $5 \leq \text{levels} \leq 8$ ,  $\text{epsilon} \leq .85$ ,  $n > \text{levels} + 30$ .

However, `ANOVA_power` can make the solution easy. Simply simulate, like we have done above, for both the null situation (no differences) and with the hypothesized effect (to determine power), and see which approach best balances type I and II error rates.

As we saw above, the unadjusted repeated measures ANOVA has an elevated type I error rate, but the MANOVA analysis of the same data above approximately preserves the type I error rate. However, how does this perform relative to the sphericity corrections? Let's simulate again, and compare the results of the different corrections.

```
design_result <- ANOVA_design("4w",
                             n = 29,
                             r = c(.05,.15,.25,.55, .65, .9),
                             sd = c(1,3,5,7),
                             mu= c(0,0,0,0))

power_result_none <- ANOVA_power(design_result, nsims = nsims, verbose = FALSE)
```

```
##                power
## manova_(Intercept)  5.16
## manova_a           5.16
```

```
power_result_gg <- ANOVA_power(design_result, correction = "GG",
                               nsims = nsims, verbose = FALSE)
```

```
##          power effect_size
## anova_a  4.59  0.03390259
```

```
power_result_hf <- ANOVA_power(design_result, correction = "HF",
                               nsims = nsims, verbose = FALSE)
```

```
##          power effect_size
## anova_a  5.12  0.03469053
```

Both the sphericity corrections, Greenhouse-Geisser (GG) and Huynh-Feldt (HF), as well as the MANOVA were able to adequately control type I error rate. However, Greenhouse-Geisser seemed to be a tad conservative. We can now directly compare the power of the MANOVA or HF-adjusted approach. We can adjust the study design to the alternative, or hypothesized, model with the predicted means of 0, 0.75, 1.5, 3 instead of the null model.

```
design_result_power <- ANOVA_design("4w",
                                   n = 29,
                                   r = c(.05,.15,.25,.55, .65, .9)
```

```

),
sd = c(1,3,5,7),
mu= c(0,0.75,1.5,3))

power_result_hfeffect <- ANOVA_power(design_result_power, correction = "HF",
                                     nsims = nsims, verbose = FALSE)

##           power effect_size
## anova_a 60.48    0.1479382

##           power
## manova_(Intercept) 49.60
## manova_a           49.64

```

Well, it appears the MANOVA based approach has roughly ~10% lower power compared the HF-adjusted ANOVA. However, the study still appears to be underpowered so we can increase the sample size. We will continue to evaluate the MANOVA results because, as Algina and Keselman (1997) note, the power disadvantage of MANOVA is diminished with increased sample size.

```

design_result_power <- ANOVA_design("4w",
                                   n = 29,
                                   r = c(.05,.15,.25,.55, .65, .9),
                                   sd = c(1,3,5,7),
                                   mu= c(0,0.75,1.5,3))

power_result_hfeffect <- ANOVA_power(design_result_power, correction = "HF",
                                     nsims = nsims, verbose = FALSE)

##           power effect_size
## anova_a 84.69    0.1385088

##           power
## manova_(Intercept) 74.11
## manova_a           79.18

```

Again, the HF-adjusted analysis appears to be more powerful for *this very specific experimental design*. The difference in power between univariate and multivariate output is diminished when the sample size is increased.