ANCOVA in R

The **Analysis of Covariance** (**ANCOVA**) is used to compare means of an outcome variable between two or more groups taking into account (or to correct for) variability of other variables, called covariates. In other words, ANCOVA allows to compare the adjusted means of two or more independent groups.

For example, you might want to compare "test score" by "level of education" taking into account the "number of hours spent studying". In this example: 1) `test score` is our outcome (dependent) variable; 2) `level of education` (high school, college degree or graduate degree) is our grouping variable; 3) `sudying time` is our covariate.

The **one-way ANCOVA** can be seen as an extension of the one-way ANOVA that incorporate a covariate variable. The **two-way ANCOVA** is used to evaluate simultaneously the effect of two independent grouping variables (A and B) on an outcome variable, after adjusting for one or more continuous variables, called covariates.

In this article, you will learn how to:

- **Compute and interpret** the one-way and the two-way ANCOVA in R
- **Check ANCOVA assumptions**
- **Perform post-hoc tests**, multiple pairwise comparisons between groups to identify which groups are different
- **Visualize the data** using box plots, add ANCOVA and pairwise comparisons p-values to the plot

Contents:

-

Practical Statistics in R II - Comparing Groups: Numerical Variables

Assumptions

ANCOVA makes several assumptions about the data, such as:

- **Linearity between the covariate and the outcome variable** at each level of the grouping variable. This can be checked by creating a grouped scatter plot of the covariate and the outcome variable.
- **Homogeneity of regression slopes**. The slopes of the regression lines, formed by the covariate and the outcome variable, should be the same for each group. This assumption evaluates that there is no interaction between the outcome and the covariate. The plotted regression lines by groups should be parallel.
- **The outcome variable should be approximately normally distributed**. This can be checked using the Shapiro-Wilk test of normality on the model residuals.
- **Homoscedasticity** or homogeneity of residuals variance for all groups. The residuals are assumed to have a constant variance (homoscedasticity)
- **No significant outliers** in the groups

Many of these assumptions and potential problems can be checked by analyzing the residual errors. In the situation, where the ANCOVA assumption is not met you can perform **robust ANCOVA** test using the WRS2 package.

Prerequisites

Make sure you have installed the following R packages:

- `tidyverse` for data manipulation and visualization
- `ggpubr` for creating easily publication ready plots
- `rstatix` for easy pipe-friendly statistical analyses
- `broom` for printing a nice summary of statistical tests as data frames
- `datarium` : contains required data sets for this chapter

Start by loading the following required packages:

```
1   library(tidyverse)
2   library(ggpubr)
3   library(rstatix)
4   library(broom)
```

One-way ANCOVA

Data preparation

We'll prepare our demo data from the `anxiety` dataset available in the datarium package.

Researchers investigated the effect of exercises in reducing the level of anxiety. Therefore, they conducted an experiment, where they measured the anxiety score of three groups of individuals practicing physical exercises at different levels (grp1: low, grp2: moderate and grp3: high).

The anxiety score was measured pre- and 6-months post-exercise training programs. It is expected that any reduction in the anxiety by the exercises programs would also depend on the participant's basal level of anxiety score.

In this analysis we use the pretest anxiety score as the covariate and are interested in possible differences between group with respect to the post-test anxiety scores.

```
1   # Load and prepare the data
2   data("anxiety", package = "datarium")
3   anxiety <- anxiety %>%
4     select(id, group, t1, t3) %>%
5     rename(pretest = t1, posttest = t3)
6   anxiety[14, "posttest"] <- 19
7   # Inspect the data by showing one random row by groups
8   set.seed(123)
9   anxiety %>% sample_n_by(group, size = 1)
10  ## # A tibble: 3 x 4
11  ##    id    group pretest posttest
12  ##    <fct> <fct>   <dbl>    <dbl>
13  ## 1 5      grp1     16.5    15.7
14  ## 2 27     grp2     17.8    16.9
15  ## 3 37     grp3     17.1    14.3
```

Check assumptions

Linearity assumption

- Create a scatter plot between the covariate (i.e., `pretest`) and the outcome variable (i.e., `posttest`)
- Add regression lines, show the corresponding equations and the R2 by groups

```
1   ggscatter(
2     anxiety, x = "pretest", y = "posttest",
3     color = "group", add = "reg.line"
4     )+
5     stat_regline_equation(
6       aes(label =  paste(..eq.label.., ..rr.label.., sep = "~~~~"), color =
    group)
7       )
```
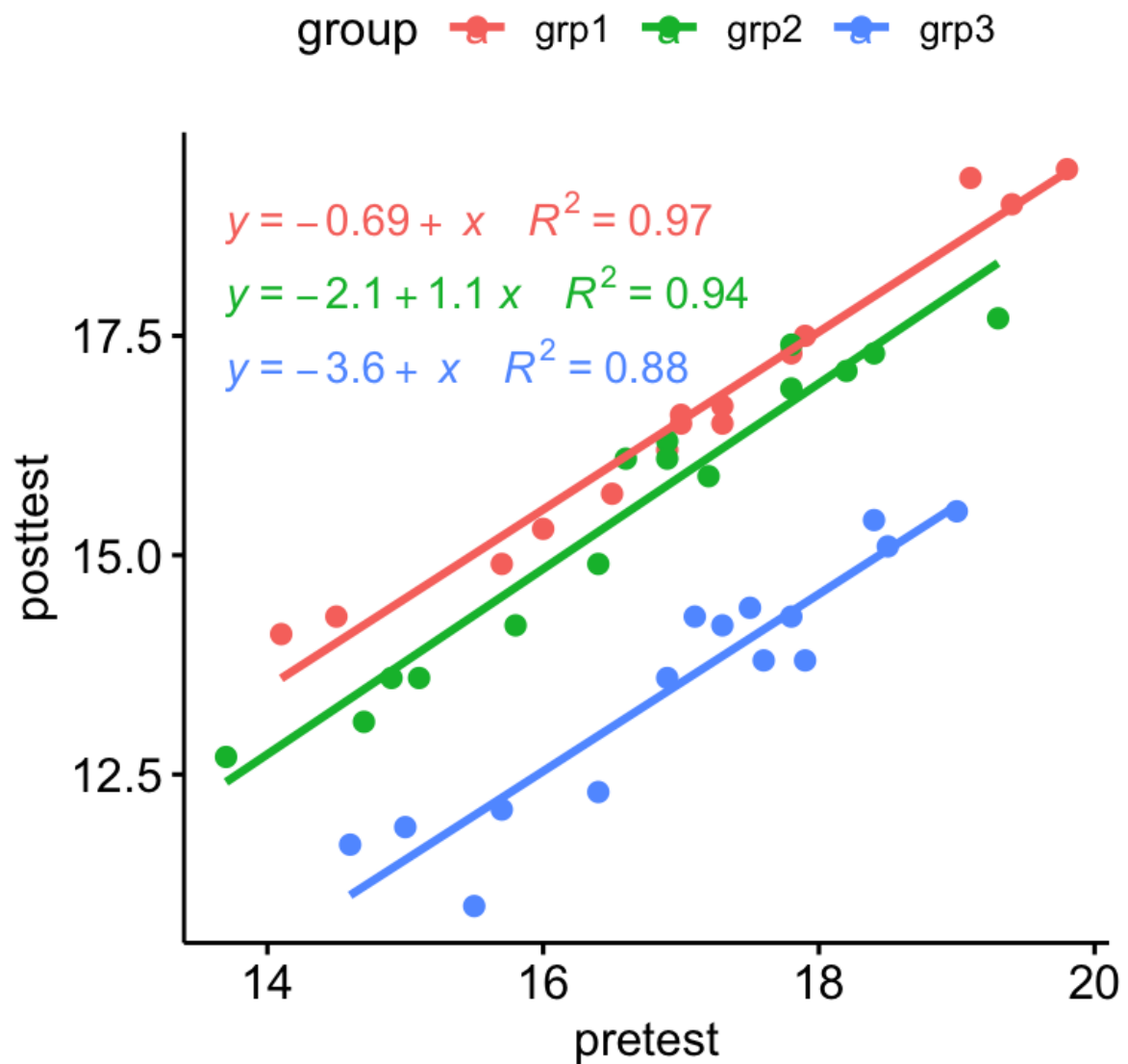
There was a linear relationship between pre-test and post-test anxiety score for each training group, as assessed by visual inspection of a scatter plot.

Homogeneity of regression slopes

This assumption checks that there is no significant interaction between the covariate and the grouping variable. This can be evaluated as follow:

```
anxiety %>% anova_test(posttest ~ group*pretest)
## ANOVA Table (type II tests)
##
##          Effect DFn DFd       F        p p<.05   ges
## 1         group   2  39 209.314 1.40e-21     * 0.915
## 2       pretest   1  39 572.828 6.36e-25     * 0.936
## 3 group:pretest   2  39   0.127 8.81e-01       0.006
```

There was homogeneity of regression slopes as the interaction term was not statistically significant, F(2, 39) = 0.13, p = 0.88.

Normality of residuals

You first need to compute the model using `lm()`. In R, you can easily augment your data to add fitted values and residuals by using the function `augment(model)` [broom package]. Let's call the output `model.metrics` because it contains several metrics useful for regression diagnostics.

```
1   # Fit the model, the covariate goes first
2   model <- lm(posttest ~ pretest + group, data = anxiety)
3   # Inspect the model diagnostic metrics
4   model.metrics <- augment(model) %>%
5     select(-.hat, -.sigma, -.fitted, -.se.fit) # Remove details
6   head(model.metrics, 3)
7   ## # A tibble: 3 x 6
8   ##   posttest pretest group .resid .cooksd .std.resid
9   ##      <dbl>   <dbl> <fct>  <dbl>   <dbl>      <dbl>
10  ## 1     14.1    14.1 grp1   0.550  0.101       1.46
11  ## 2     14.3    14.5 grp1   0.338  0.0310      0.885
12  ## 3     14.9    15.7 grp1  -0.295  0.0133     -0.750
13  # Assess normality of residuals using shapiro wilk test
14  shapiro_test(model.metrics$.resid)
15  ## # A tibble: 1 x 3
16  ##   variable            statistic p.value
17  ##   <chr>                   <dbl>   <dbl>
18  ## 1 model.metrics$.resid    0.975   0.444
```

The Shapiro Wilk test was not significant (p > 0.05), so we can assume normality of residuals

Homogeneity of variances

ANCOVA assumes that the variance of the residuals is equal for all groups. This can be checked using the Levene's test:

```
1   model.metrics %>% levene_test(.resid ~ group)
2   ## # A tibble: 1 x 4
3   ##     df1   df2 statistic     p
4   ##   <int> <int>     <dbl> <dbl>
5   ## 1     2    42      2.27 0.116
```

The Levene's test was not significant (p > 0.05), so we can assume homogeneity of the residual variances for all groups.

Outliers

An outlier is a point that has an extreme outcome variable value. The presence of outliers may affect the interpretation of the model.

Outliers can be identified by examining the standardized residual (or studentized residual), which is the residual divided by its estimated standard error. Standardized residuals can be interpreted as the number of standard errors away from the regression line.

Observations whose standardized residuals are greater than 3 in absolute value are possible outliers.

```
1  model.metrics %>%
2    filter(abs(.std.resid) > 3) %>%
3    as.data.frame()
4  ## [1] posttest   pretest     group      .resid     .cooksd    .std.resid
5  ## <0 rows> (or 0-length row.names)
```

There were no outliers in the data, as assessed by no cases with standardized residuals greater than 3 in absolute value.

Computation

The orders of variables matters when computing ANCOVA. You want to remove the effect of the covariate first - that is, you want to control for it - prior to entering your main variable or interest.

The covariate goes first (and there is no interaction)! If you do not do this in order, you will get different results.

```
1  res.aov <- anxiety %>% anova_test(posttest ~ pretest + group)
2  get_anova_table(res.aov)
3  ## ANOVA Table (type II tests)
4  ##
5  ##     Effect DFn DFd   F        p p<.05   ges
6  ## 1 pretest   1  41 598 4.48e-26     * 0.936
7  ## 2   group   2  41 219 1.35e-22     * 0.914
```

After adjustment for pre-test anxiety score, there was a statistically significant difference in post-test anxiety score between the groups, $F_{(2, 41)} = 218.63$, $p < 0.0001$.

Post-hoc test

Pairwise comparisons can be performed to identify which groups are different. The Bonferroni multiple testing correction is applied. This can be easily done using the function `emmeans_test()` [rstatix package], a wrapper around the `emmeans` package, which needs to be installed. Emmeans stands for **estimated marginal means** (aka least square means or adjusted means).

```
1   # Pairwise comparisons
2   library(emmeans)
3   pwc <- anxiety %>%
4     emmeans_test(
5       posttest ~ group, covariate = pretest,
6       p.adjust.method = "bonferroni"
7       )
8   pwc
9   ## # A tibble: 3 x 8
10  ##   .y.      group1 group2    df statistic        p    p.adj p.adj.signif
11  ## * <chr>    <chr>  <chr>  <dbl>     <dbl>    <dbl>    <dbl> <chr>
12  ## 1 posttest grp1   grp2      41      4.24 1.26e- 4 3.77e- 4 ***
13  ## 2 posttest grp1   grp3      41     19.9  1.19e-22 3.58e-22 ****
14  ## 3 posttest grp2   grp3      41     15.5  9.21e-19 2.76e-18 ****
15  # Display the adjusted means of each group
16  # Also called as the estimated marginal means (emmeans)
17  get_emmeans(pwc)
18  ## # A tibble: 3 x 8
```

```
19  ##    pretest group emmean     se     df conf.low conf.high method
20  ##      <dbl> <fct>  <dbl> <dbl> <dbl>    <dbl>     <dbl> <chr>
21  ## 1    16.9 grp1    16.4 0.106    41     16.2      16.7 Emmeans test
22  ## 2    16.9 grp2    15.8 0.107    41     15.6      16.0 Emmeans test
23  ## 3    16.9 grp3    13.5 0.106    41     13.2      13.7 Emmeans test
```

Data are adjusted mean +/- standard error. The mean anxiety score was statistically significantly greater in grp1 (16.4 +/- 0.15) compared to the grp2 (15.8 +/- 0.12) and grp3 (13.5 +/_ 0.11), p < 0.001.

Report

An ANCOVA was run to determine the effect of exercises on the anxiety score after controlling for basal anxiety score of participants.

After adjustment for pre-test anxiety score, there was a statistically significant difference in post-test anxiety score between the groups, $F(2, 41) = 218.63$, $p < 0.0001$.

Post hoc analysis was performed with a Bonferroni adjustment. The mean anxiety score was statistically significantly greater in grp1 (16.4 +/- 0.15) compared to the grp2 (15.8 +/- 0.12) and grp3 (13.5 +/_ 0.11), p < 0.001.

```
1  # Visualization: line plots with p-values
2  pwc <- pwc %>% add_xy_position(x = "group", fun = "mean_se")
3  ggline(get_emmeans(pwc), x = "group", y = "emmean") +
4    geom_errorbar(aes(ymin = conf.low, ymax = conf.high), width = 0.2) +
5    stat_pvalue_manual(pwc, hide.ns = TRUE, tip.length = FALSE) +
6    labs(
7      subtitle = get_test_label(res.aov, detailed = TRUE),
8      caption = get_pwc_label(pwc)
9    )
```

Anova, $F(2,41) = 218.63$, $p = <0.0001$, $\eta_g^2 = 0.91$

pwc: **Emmeans test**; p.adjust: **Bonferroni**
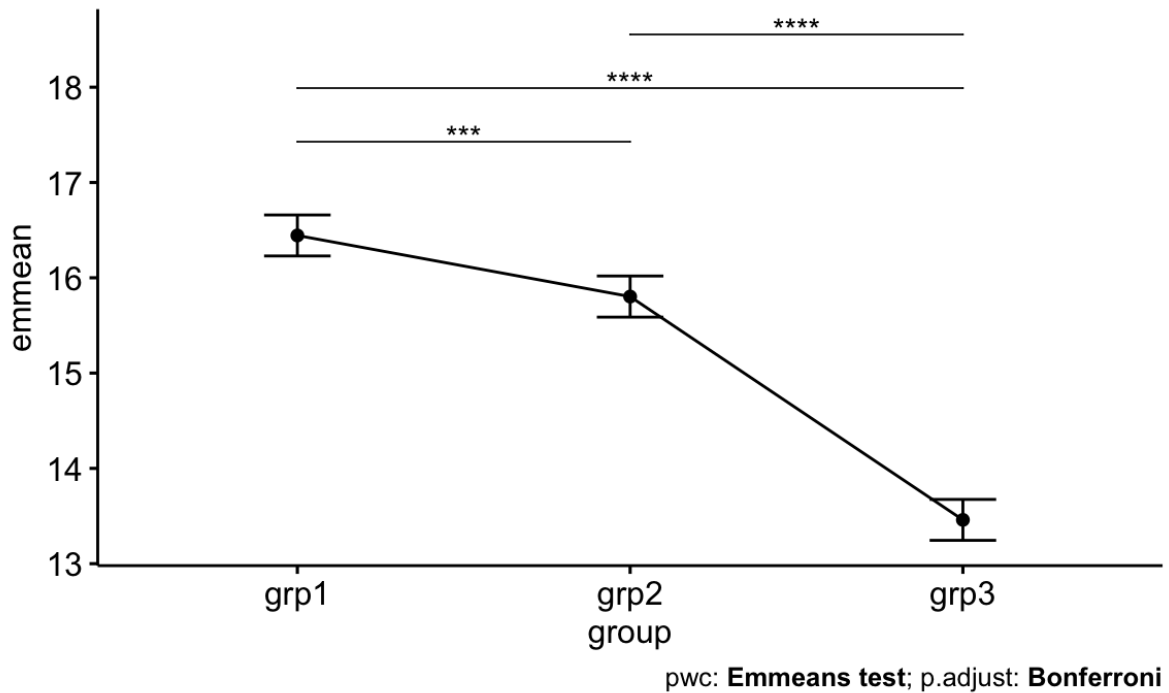
Two-way ANCOVA

Data preparation

We'll use the `stress` dataset available in the datarium package. In this study, a researcher wants to evaluate the effect of `treatment` and `exercise` on stress reduction `score` after adjusting for `age`.

In this example: 1) `stress score` is our outcome (dependent) variable; 2) `treatment` (levels: no and yes) and `exercise` (levels: low, moderate and high intensity training) are our grouping variable; 3) `age` is our covariate.

Load the data and show some random rows by groups:

```
 1  data("stress", package = "datarium")
 2  stress %>% sample_n_by(treatment, exercise)
 3  ## # A tibble: 6 x 5
 4  ##       id score treatment exercise    age
 5  ##    <int> <dbl> <fct>     <fct>      <dbl>
 6  ## 1     8  83.8 yes       low           61
 7  ## 2    15  86.9 yes       moderate      55
 8  ## 3    29  71.5 yes       high          55
 9  ## 4    40  92.4 no        low           67
10  ## 5    41 100   no        moderate      75
11  ## 6    56  82.4 no        high          53
```

Check assumptions

Linearity assumption

- Create a scatter plot between the covariate (i.e., `age`) and the outcome variable (i.e., `score`) for each combination of the groups of the two grouping variables:
- Add smoothed loess lines, which helps to decide if the relationship is linear or not

```
1  ggscatter(
2    stress, x = "age", y = "score",
3    facet.by  = c("exercise", "treatment"),
4    short.panel.labs = FALSE
5    )+
6    stat_smooth(method = "loess", span = 0.9)
```
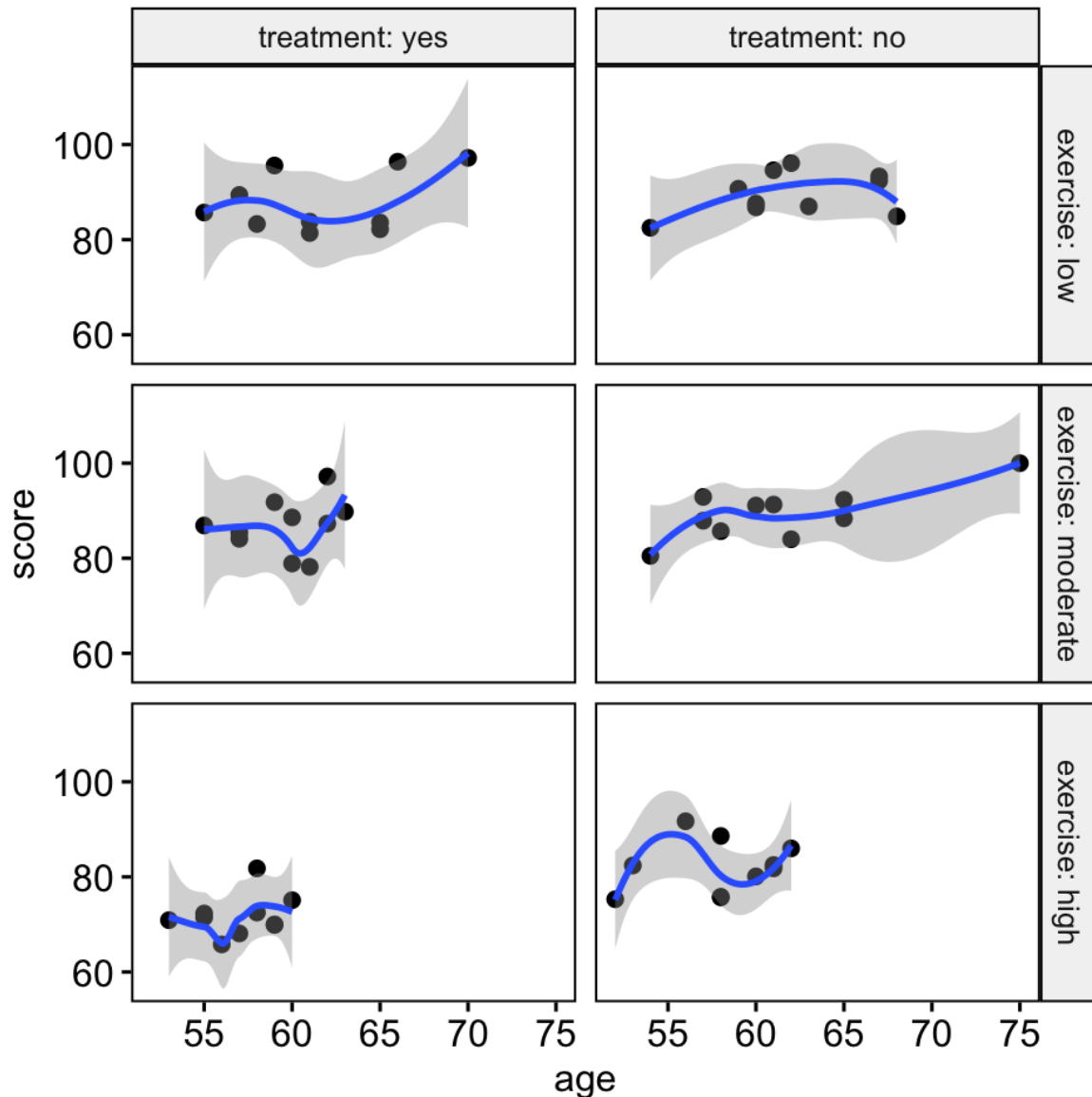

There was a linear relationship between the covariate (age variable) and the outcome variable (score) for each group, as assessed by visual inspection of a scatter plot.

Homogeneity of regression slopes

This assumption checks that there is no significant interaction between the covariate and the grouping variables. This can be evaluated as follow:

```
1  stress %>%
```

```
 2    anova_test(
 3      score ~ age + treatment + exercise +
 4        treatment*exercise + age*treatment +
 5        age*exercise + age*exercise*treatment
 6    )
 7  ## ANOVA Table (type II tests)
 8  ##
 9  ##                       Effect DFn DFd      F        p p<.05      ges
10  ## 1                        age   1  48  8.359 6.00e-03     * 0.148000
11  ## 2                  treatment   1  48  9.907 3.00e-03     * 0.171000
12  ## 3                   exercise   2  48 18.197 1.31e-06     * 0.431000
13  ## 4         treatment:exercise   2  48  3.303 4.50e-02     * 0.121000
14  ## 5              age:treatment   1  48  0.009 9.25e-01       0.000189
15  ## 6               age:exercise   2  48  0.235 7.91e-01       0.010000
16  ## 7 age:treatment:exercise      2  48  0.073 9.30e-01       0.003000
```

Another simple alternative is to create a new grouping variable, say `group`, based on the combinations of the existing variables, and then compute ANOVA model:

```
1  stress %>%
2    unite(col = "group", treatment, exercise) %>%
3    anova_test(score ~ group*age)
4  ## ANOVA Table (type II tests)
5  ##
6  ##       Effect DFn DFd      F        p p<.05   ges
7  ## 1      group   5  48 10.912 4.76e-07     * 0.532
8  ## 2        age   1  48  8.359 6.00e-03     * 0.148
9  ## 3 group:age    5  48  0.126 9.86e-01       0.013
```

There was homogeneity of regression slopes as the interaction terms, between the covariate (`age`) and grouping variables (`treatment` and `exercise`), was not statistically significant, p > 0.05.

Normality of residuals

```
 1  # Fit the model, the covariate goes first
 2  model <- lm(score ~ age + treatment*exercise, data = stress)
 3  # Inspect the model diagnostic metrics
 4  model.metrics <- augment(model) %>%
 5    select(-.hat, -.sigma, -.fitted, -.se.fit) # Remove details
 6  head(model.metrics, 3)
 7  ## # A tibble: 3 x 7
 8  ##   score   age treatment exercise .resid .cooksd .std.resid
 9  ##   <dbl> <dbl> <fct>     <fct>     <dbl>   <dbl>      <dbl>
10  ## 1  95.6    59 yes       low        9.10  0.0647       1.93
11  ## 2  82.2    65 yes       low       -7.32  0.0439      -1.56
12  ## 3  97.2    70 yes       low        5.16  0.0401       1.14
13  # Assess normality of residuals using shapiro wilk test
14  shapiro_test(model.metrics$.resid)
15  ## # A tibble: 1 x 3
16  ##   variable            statistic p.value
17  ##   <chr>                   <dbl>   <dbl>
18  ## 1 model.metrics$.resid    0.982   0.531
```

The Shapiro Wilk test was not significant (p > 0.05), so we can assume normality of residuals

Homogeneity of variances

ANCOVA assumes that the variance of the residuals is equal for all groups. This can be checked using the Levene's test:

```
1  levene_test(.resid ~ treatment*exercise, data = model.metrics)
```

The Levene's test was not significant (p > 0.05), so we can assume homogeneity of the residual variances for all groups.

Outliers

Observations whose standardized residuals are greater than 3 in absolute value are possible outliers.

```
1  model.metrics %>%
2    filter(abs(.std.resid) > 3) %>%
3    as.data.frame()
4  ## [1] score      age        treatment  exercise   .resid     .cooksd
   .std.resid
5  ## <0 rows> (or 0-length row.names)
```

There were no outliers in the data, as assessed by no cases with standardized residuals greater than 3 in absolute value.

Computation

```
1   res.aov <- stress %>%
2     anova_test(score ~ age + treatment*exercise)
3   get_anova_table(res.aov)
4   ## ANOVA Table (type II tests)
5   ##
6   ##              Effect DFn DFd     F       p p<.05   ges
7   ## 1              age   1  53  9.11 4.00e-03     * 0.147
8   ## 2        treatment   1  53 11.10 2.00e-03     * 0.173
9   ## 3         exercise   2  53 20.82 2.13e-07     * 0.440
10  ## 4 treatment:exercise 2  53  4.45 1.60e-02     * 0.144
```

After adjustment for age, there was a statistically significant interaction between treatment and exercise on the stress score, $F_{(2, 53)} = 4.45$, $p = 0.016$. This indicates that the effect of exercise on score depends on the level of exercise, and vice-versa.

Post-hoc test

A statistically significant two-way interactions can be followed up by **simple main effect analyses**, that is evaluating the effect of one variable at each level of the second variable, and vice-versa.

In the situation, where the interaction is not significant, you can report the main effect of each grouping variable.

A **significant two-way interaction** indicates that the impact that one factor has on the outcome variable depends on the level of the other factor (and vice versa). So, you can decompose a significant two-way interaction into:

- **Simple main effect**: run one-way model of the first variable (factor A) at each level of the second variable (factor B),
- **Simple pairwise comparisons**: if the simple main effect is significant, run multiple pairwise comparisons to determine which groups are different.

For a **non-significant two-way interaction**, you need to determine whether you have any statistically significant **main effects** from the ANCOVA output.

In this section we'll describe the procedure for a significant three-way interaction

Simple main effect analyses for treatment

Analyze the simple main effect of `treatment` at each level of `exercise`. Group the data by `exercise` and perform one-way ANCOVA for `treatment` controlling for `age`:

```
 1  # Effect of treatment at each level of exercise
 2  stress %>%
 3    group_by(exercise) %>%
 4    anova_test(score ~ age + treatment)
 5  ## # A tibble: 6 x 8
 6  ##   exercise Effect      DFn   DFd      F         p `p<.05`   ges
 7  ##   <fct>    <chr>     <dbl> <dbl>  <dbl>     <dbl> <chr>    <dbl>
 8  ## 1 low      age           1    17   2.25  0.152    ""       0.117
 9  ## 2 low      treatment     1    17   0.437 0.517    ""       0.025
10  ## 3 moderate age           1    17   6.65  0.02     *        0.281
11  ## 4 moderate treatment     1    17   0.419 0.526    ""       0.024
12  ## 5 high     age           1    17   0.794 0.385    ""       0.045
13  ## 6 high     treatment     1    17  18.7   0.000455 *        0.524
```

Note that, we need to apply Bonferroni adjustment for multiple testing corrections. One common approach is lowering the level at which you declare significance by dividing the alpha value (0.05) by the number of tests performed. In our example, that is 0.05/3 = 0.016667.

Statistical significance was accepted at the Bonferroni-adjusted alpha level of 0.01667, that is 0.05/3. The effect of treatment was statistically significant in the high-intensity exercise group (p = 0.00045), but not in the low-intensity exercise group (p = 0.517) and in the moderate-intensity exercise group (p = 0.526).

Compute pairwise comparisons between treatment groups at each level of exercise. The Bonferroni multiple testing correction is applied.

```
1   # Pairwise comparisons
2   pwc <- stress %>%
3     group_by(exercise) %>%
4     emmeans_test(
5       score ~ treatment, covariate = age,
6       p.adjust.method = "bonferroni"
7       )
8   pwc %>% filter(exercise == "high")
9   ## # A tibble: 1 x 9
10  ##   exercise .y.   group1 group2    df statistic         p     p.adj
    p.adj.signif
11  ##   <fct>    <chr> <chr>  <chr>  <dbl>     <dbl>     <dbl>     <dbl> <chr>

12  ## 1 high     score yes    no        53     -4.36 0.0000597 0.0000597 ****
```

In the pairwise comparison table, you will only need the result for "exercises:high" group, as this was the only condition where the simple main effect of treatment was statistically significant.

The pairwise comparisons between treatment:no and treatment:yes group was statistically significant in participant undertaking high-intensity exercise (p < 0.0001).

Simple main effect for exercise

You can do the same post-hoc analyses for the `exercise` variable at each level of `treatment` variable.

```
1   # Effect of exercise at each level of treatment
2   stress %>%
3     group_by(treatment) %>%
4     anova_test(score ~ age + exercise)
5   ## # A tibble: 4 x 8
6   ##   treatment Effect     DFn   DFd     F        p `p<.05`   ges
7   ##   <fct>     <chr>    <dbl> <dbl> <dbl>    <dbl> <chr>   <dbl>
8   ## 1 yes       age          1    26  2.37 0.136    ""      0.083
9   ## 2 yes       exercise     2    26 17.3  0.0000164 *      0.572
10  ## 3 no        age          1    26  7.26 0.012    *       0.218
11  ## 4 no        exercise     2    26  3.99 0.031    *       0.235
```

Statistical significance was accepted at the Bonferroni-adjusted alpha level of 0.025, that is 0.05/2 (the number of tests). The effect of exercise was statistically significant in the treatment=yes group (p < 0.0001), but not in the treatment=no group (p = 0.031).

Perform multiple pairwise comparisons between `exercise` groups at each level of `treatment`. You don't need to interpret the results for the "no treatment" group, because the effect of `exercise` was not significant for this group.

```
1   pwc2 <- stress %>%
2     group_by(treatment) %>%
3     emmeans_test(
4       score ~ exercise, covariate = age,
5       p.adjust.method = "bonferroni"
6       ) %>%
7     select(-df, -statistic, -p) # Remove details
8   pwc2 %>% filter(treatment == "yes")
```

```
 9  ## # A tibble: 3 x 6
10  ##    treatment .y.   group1   group2       p.adj p.adj.signif
11  ##    <fct>     <chr> <chr>    <chr>        <dbl> <chr>
12  ## 1 yes        score low      moderate 1          ns
13  ## 2 yes        score low      high     0.00000113   ****
14  ## 3 yes        score moderate high     0.000000466  ****
```

There was a statistically significant difference between the adjusted mean of low and high exercise group (p < 0.0001) and, between moderate and high group (p < 0.0001). The difference between the adjusted means of low and moderate was not significant.

Report

A two-way ANCOVA was performed to examine the effects of treatment and exercise on stress reduction, after controlling for age.

There was a statistically significant two-way interaction between treatment and exercise on score concentration, whilst controlling for age, $F(2, 53) = 4.45$, $p = 0.016$.

Therefore, an analysis of simple main effects for exercise and treatment was performed with statistical significance receiving a Bonferroni adjustment and being accepted at the $p < 0.025$ level for exercise and $p < 0.0167$ for treatment.

The simple main effect of treatment was statistically significant in the high-intensity exercise group ($p = 0.00046$), but not in the low-intensity exercise group ($p = 0.52$) and the moderate-intensity exercise group ($p = 0.53$).

The effect of exercise was statistically significant in the `treatment=yes` group ($p < 0.0001$), but not in the `treatment=no` group ($p = 0.031$).

All pairwise comparisons were computed for statistically significant simple main effects with reported p-values Bonferroni adjusted. For the `treatment=yes` group, there was a statistically significant difference between the adjusted mean of low and high exercise group ($p < 0.0001$) and, between moderate and high group ($p < 0.0001$). The difference between the adjusted means of low and moderate exercise groups was not significant.

- Create a line plot:

```
1  # Line plot
2  lp <- ggline(
3    get_emmeans(pwc), x = "exercise", y = "emmean",
4    color = "treatment", palette = "jco"
5    ) +
6    geom_errorbar(
7      aes(ymin = conf.low, ymax = conf.high, color = treatment),
8      width = 0.1
9      )
```

- Add p-values

```
1   # Comparisons between treatment group at each exercise level
2   pwc <- pwc %>% add_xy_position(x = "exercise", fun = "mean_se",
    step.increase = 0.2)
3   pwc.filtered <- pwc %>% filter(exercise == "high")
4   lp +
5   stat_pvalue_manual(
6     pwc.filtered, hide.ns = TRUE, tip.length = 0,
7     bracket.size = 0
8     ) +
9   labs(
10    subtitle = get_test_label(res.aov,  detailed = TRUE),
11    caption = get_pwc_label(pwc)
12  )
```

Anova, $F(2,53) = 4.45$, $p = 0.016$, $\eta_g^2 = 0.14$

pwc: **Emmeans test**; p.adjust: **Bonferroni**
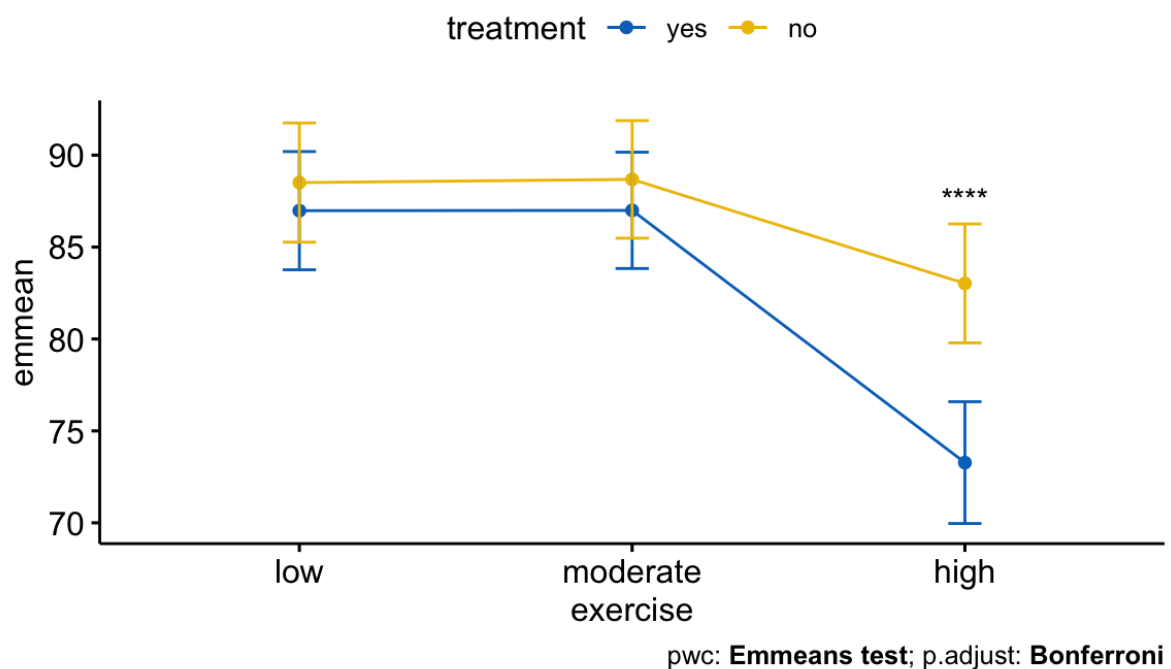
```
1   # Comparisons between exercises group at each treatment level
2   pwc2 <- pwc2 %>% add_xy_position(x = "exercise", fun = "mean_se")
3   pwc2.filtered <- pwc2 %>% filter(treatment == "yes")
4   lp +
5   stat_pvalue_manual(
6     pwc2.filtered, hide.ns = TRUE, tip.length = 0,
7     step.group.by = "treatment", color = "treatment"
8     ) +
9   labs(
10    subtitle = get_test_label(res.aov,  detailed = TRUE),
11    caption = get_pwc_label(pwc2)
12  )
```

Anova, $F(2,53) = 4.45$, $p = 0.016$, $\eta^2_g = 0.14$

pwc: **Emmeans test**; p.adjust: **Bonferroni**
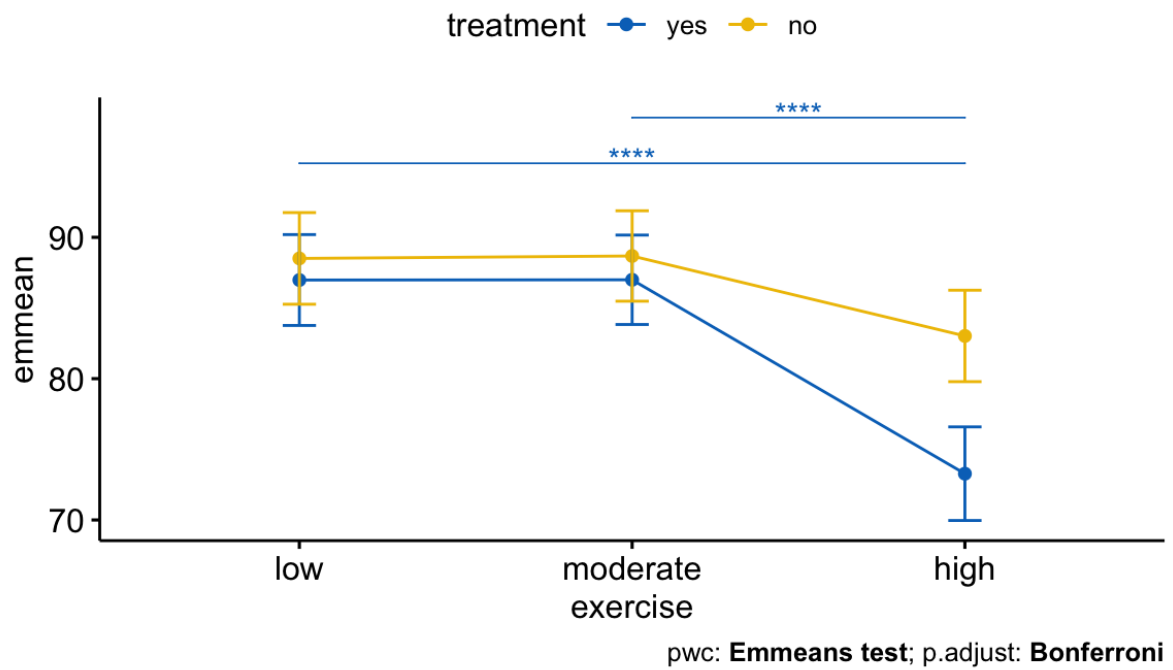
Summary

This article describes how to compute and interpret one-way and two-way ANCOVA in R. We also explain the assumptions made by ANCOVA tests and provide practical examples of R codes to check whether the test assumptions are met or not.