Chapter 4

4.1

The ANCOVA is a form of statistical control, and was specifically developed in order to reduce unexplained error variation. To do this a covariate is included. This is a source of variation that is believed to influence the dependent variable, but has not been included in the regular design. It is sometimes also referred to as the concomitant variable. The covariate is used to (1) reduce error variation, (2) take into account any pre-existing level mean difference of the covariate, (3) take into account the relationship between the covariate and dependent variable, and (4) create a more precise and less biased estimate of level/group effects. An ANCOVA model is more powerful than an ANOVA model as long as it succeeds in reducing the error variation. (See chapter 6 for an extensive comparison of ANOVA and ANCOVA). The covariate has its uses, but for every covariate one degree of freedom is lost from the error term. This leads to a larger critical value for the F-test, and raises the difficulty of finding a statistically significant F-test statistic.

In an ANCOVA model there is a single independent variable with two or more levels/categories (meaning the independent variable is either nominal or ordinal). These levels are fixed and decided by the researcher beforehand. Subjects are then assigned to one of this levels (unless you do ANCOVA with repeated measures).

A situation is referred to as a true experimental design if the researcher can indeed randomly assign subjects to these levels. When this is not the case, and the researcher thus has no control over to which level the subject belongs, it is a quasi-experimental design. There are two possible reasons why there may be such a lack of control. First, there may have already been groups before the research is started. Such groups are called intact groups. Second, it may be the case that subjects cannot be assigned to certain levels because of distinctions already made (if levels are divided because of income level, or education level for example).

Also important to consider in ANCOVA is the measurement scales of the variables. Namely, it is assumed in ANCOVA that the dependent variable is measured at the interval level at least. This same assumption applies to the covariate. And, as covered above, the independent variable must be nominal or ordinal.

The dependent variable is statistically adjusted in ANCOVA in order to remove the uncontrolled effects of the covariate. This is done by adjusting the level means of the dependent variable so they now represent levels with the same means in the covariate.

4.2

Each observation of the dependent variable is designated as Y_{ij}. In the subscript the "j" stands for the level that the observation belongs to, and the "i" for the observation/identification number within that level. The number of levels of the independent variable are demarcated by "J", and the number of observations in group "j" are indicated as "n_j". X_{ij} indicates each observation on the covariate.

4.3

ANCOVA model is a form of the general linear model (GLM), and can be written as follows in terms of population parameters:

$Y_{ij} = \mu_Y + \alpha_j + \beta_w (X_{ij} - \mu_X) + \varepsilon_{ij}$

Here Y_{ij} is the observed score on the dependent variable for individual i in group j. μ_Y is the grand population mean for the dependent variable Y. α_j is the group effects for group j. β_w is the within-groups regression slope from the regression of Y on X. μ_X is the grand population mean for the covariate X. And ε_{ij} is the random residual error for individual i in group j.

This residual error can be caused by individual differences, measurement error and/or other (unknown) factors.

The sum of the group effects is 0 here, just like in ANOVA.

The H0 here states that all of the adjusted means are equal.

4.4

The summary table for ANCOVA is as follows:

Source	SS	df	MS	F
Between adjusted	SSbetw(adj)	J – 1	MSbetw(adj)	MSbetw(adj)/MSwith(adj)
Within adjusted	SSwith(adj)	N – J -1	MSwith(adj)	
Covariate	SScov	1	MScov	MScov/MSwith(adj)
Total	SStotal	N -1		

Here the "between" source represents the systematically studying of the independent variable, and the "within" represents the error/residual.

For the test of difference between the adjusted means the critical value is $\alpha F_{J-1,N-J-1}$. For the test of the covariate the critical value is $\alpha F_{1,N-J-1}$. In each case the H0 is rejected if the F-test statistic is greater than the F critical value. In order for the ANCOVA to be rightfully used it is important that the F-test statistic of the test of the covariate is greater than the F critical value (because this shows that the covariate is significantly related to the dependent variable).

If the F-test statistic for the test of difference between the adjusted means is greater than the F critical value, however, this leads to uncertainty of how the means are different, though only if there are more than two levels. In this case an MCP can be used to find out which means are different.

4.5

Partitioning the sums of squares is essential in all GLM's. The firsts step in doing this is partitioning the total variation into its relevant parts/sources of variation, which is done as follows:

SStotal = SSbew(adj) + SSwith(adj) + SScov

After this statistical software is used to carry out the remaining computations.

4.6

The adjusted mean is denoted as Ȳj', and is estimated as follows:

$$\bar{Y}'_{.j} = \bar{Y}_{.j} - b_w \left(-X_{.j} - -X_{..} \right)$$

Here bw is the within-levels regression slope, with the part between brackets indicating the difference between the level mean and the overall mean of the covariate (this difference is thus the level/group effect). If bw is zero, or the means are the same, no adjustment will be made

If you want to use an MCP in an ANCOVA situation, the procedures first need to be adjusted for use with a covariate. These adapted procedures then involve a different form of the standard error of a contrast (as the contrasts are formed here on the base of adjusted means).

The procedures of power, confidence intervals (CI's), and effect size measures work the same in ANCOVA as they do in ANOVA, with the exception that they are based on adjusted means when used in ANCOVA.

4.7

Introduction of a covariate means the introduction of more assumptions than included in a regular ANOVA (for these ANOVA assumptions, see chapters 1 and 3). These new assumptions are:

- Linearity: This assumption is that the regression of Y on X is linear. Use of the usual ANCOVA procedure is not appropriate if this regression is not linear (as ANCOVA also fits a straight line to the data). Non-linear data will thus lead to the group effects being biased, and adjustments made in SSwith and SSbetw will be smaller. Generally violations of this assumption can be detected by examining a scatterplot of Y versus X. For full accuracy this also needs to be done for each level/category in the independent variable. The two alternatives when non-linearity is found are transformations and non-linear ANCOVA.

- Independence of the covariate and the independent variable: This is a condition of the ANCOVA model, which requires that the covariate and the independent variable be independent. Use of a covariate may be problematic if this covariate is affected by the treatment itself. This may result in (1) deletion of part of the treatment effect, (2) production of an inflated treatment effect, or (3) alteration of the covariate scores (if the treatment is administered before the covariate data is obtained). For this last reason it is important to keep an eye out for possible covariate variables before starting the study. Whether this condition is violated can be assessed through examination of the mean differences in the covariate across the levels of the independent variable. The condition is likely met when the levels are not statistically significant in the covariate.

- The covariate being measured without error: This assumption is especially important when it comes to research in education and the behavioural sciences, where there is often a considerable measurement error when variables are measured. It is also of interest in randomized experiments, where the reduction in unexplained variation is small, and the F-test not very powerful, due to an underestimated bw. It is possible to detect obvious violations of this assumption by computing the reliability of the covariate before conducting the study (or computing it from prior research). The validity of the covariate can also be considered here.

- Homogeneity of the regression slopes: This assumption states that the slope of the regression line between the dependent variable and covariate is the same for each category of the independent variable. This assumption is important because it allows the use of bw as the within-levels regression slope. It also allows the testing for group intercept differences (which is the central aim of ANCOVA, as it's the same as testing for differences between adjusted means). If this assumption is violated this represents a type of interaction, and interpretation of any kind of result is not really possible (the size of this effect differs per model though). To test for violation of this assumption a formal statistical procedure is often used. It is possible, however, to first simply look whether the slopes look similar by examining a scatterplot (also known as the eyeball method). There are five alternatives if this assumption is violated: (1) Use the covariate as a blocking variable. (2) Analyse each group/level separately (subsets of the level do need to have equal slopes). This is a more undesirable method. (3) Utilize the interaction terms between the covariate and the independent variable, and use this to conduct a regression analysis. (4) Use the Johnson & Neyman (1936) technique, determining the values of X that are related to significant level differences on Y. (5) Use more-modern robust methods.

Some of the assumptions that are the same as in ANOVA need to be treated differently, however. This are:

- Independence: This assumption can be met by: (1) making sure that the assignment of individuals to levels remains separate throughout the design of the experiment, and (2) making sure that individuals are kept separate from each other with the use of experimental control, in order to make sure that the scores of the dependent variable Y are independent across subjects. Random assignment helps achieve independence. As in ANOVA, use of independent random samples is very important here. Violation of this assumption increases both Type 1 and Type 2 errors, due to the sensitivity of the F-ratio to this assumption. A violation may also affect the standard errors of the sample adjusted means, thus influencing any inferences that were made about those means. Independence can best be assessed through examination of residual plots by groups. The residuals will be placed randomly if the assumption is satisfied, and will form a cyclical pattern if it is violated. This assumption is generally violated in scenarios where there is time series data, observation within blocks, or replication. Sadly there is not an easy way to "fix" the violation of this assumption.

- Homogeneity of variance: This assumption refers to the situation where the variances of each population are the same. Violation of this assumption will likely lead to bias in the SSwith-term, an increased possibility of Type-1 errors, and possibly even an increase in the likelihood of Type-2 errors. With equal, or nearly equal n's across the levels, however, the violation is negligible. But if the larger n's are associated with the smaller or larger variances (when the observed α is larger or smaller than the stated α) it is problematic. Violation of this assumption can be detected by examining a plot of Y versus the covariate X. Another option is to use formal statistics (see chapter 1 for more details).

- Normality: This assumption is that all of the populations follow the normal distribution. As the F-test is luckily quite robust here, though, there is only cause for worry if there is serious non-normality. Frequency distributions or normal probability plots can be used to detect violation of this assumption. When this assumption is violated, the data can be normalized with transformations (see chapter 1 for more details).

- Fixed independent variable: This assumption states that researcher sets the levels of the independent variable. If this assumption is maintained it results in a fixed-effects model, if it is violated it results in a random-effects model.

4.8

4.9

Randomization is the procedure in an experiment where individuals are randomly assigned to groups (Note: this is not the same as random selection, as that is about sampling). Designs where randomization is used are referred to as true experiments. True experiments in ANCOVA are always more powerful than experiments that do not meet the same requirements. In a true experiment the probability that the levels differ on the covariate is equal to α. This means that is very unlikely that group/level means will be any different in the covariate, meaning that any adjustment in the level means will be small. This is important as it means that the error term is likely to be greatly reduced.

As discussed earlier, randomization is not always possible, and the designs where this is the case are referred to as quasi-experimental. In such experiments it is more likely that there are statistically significant differences between the level means in the covariate. Adjustment in the level means can thus be quite large in such experiments. This can have several effects:

- The groups/levels are likely to be different when it comes to other important characteristics, which may not have been controlled for statistically or experimentally.

- It is less likely that the homogeneity of regression slopes assumption is going to be met.

- Part of the treatment effect may be removed when adjusting for the covariate.

- Equating levels within the covariate could be an extrapolation that is beyond the range of possible values that occur in one of the particular levels.

- When extrapolating beyond the range of scores, the slopes may not be equal, even if they were for the range of X's obtained.

- The standard errors of the adjusted means could increase, meaning that the tests of the adjusted means are not significant anymore.

- A differential growth in the levels may occur, confounding the results.

ANCOVA can still be used in quasi-experiments though, any researcher should simply exercise more caution when interpreting results in this situation.

4.10

Just like ANOVA models can be extended into more complex models, complex ANCOVA models can also be created. Any of the following designs can be part of an ANCOVA: factorial designs; fixed-, random-, and mixed-effects designs; repeated measures and split-plot (mixed) designs; hierarchical designs; and randomized block designs.

Such designs work the same in ANCOVA as they do in ANOVA.

4.11

Nonparametric ANCOVA models are an alternative that can be considered when assumptions of normality, homogeneity of variance, and/or linearity have been violated in a serious manner. It is also to be used if the dependent variable is measured at the ordinal level.

4.12

See p. 155-179 for an extensive overview of computing ANCOVA and using "G*Power" in SPSS.

4.13

See p. 179-181 for a guide to writing about ANCOVA