Chap 6: Analysis of Variance

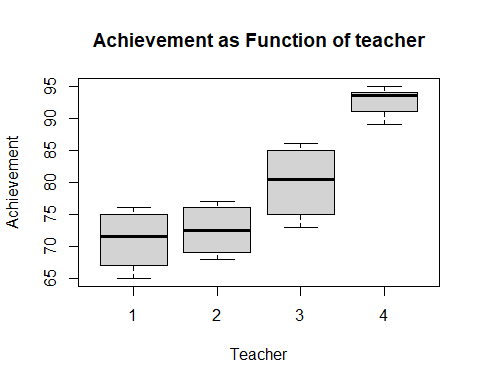
## Learning Objectives

* ANOVA Analysis: Fixed Effects, Random Effects, Mixed Models and Repeated Measures
* Logic behind ANOVA
* Verification of Assumptions in ANOVA model with Inferential Tests and Graphs
* Distinguishing between fixed effects, random effects and mixed models in ANOVA
* Evaluating pairwise comparisons with Tukeys test (Studentized Range Distribution), as post-hoc method, after the F-Test for significance
* Using aov() and lm() to run ANOVA in R
* Understand and read interaction measures in tables and graphs
* Perform Simple Effects analyses in R , as a follow up to interaction
* Use and interpret statistics like AIC and BIC in ANOVA model outputs
* Run random effects and mixed effects model in R
* Distinguish between repeated-measures models and between-subjects model
* Run Repeated-Measures models in R

library(readxl)  
achiev <- read\_excel('D:/june2023\_excel\_workbooks/achiev.xlsx')  
achiev

## # A tibble: 24 × 3  
## ac teach text  
## <dbl> <dbl> <dbl>  
## 1 70 1 1  
## 2 67 1 1  
## 3 65 1 1  
## 4 75 1 2  
## 5 76 1 2  
## 6 73 1 2  
## 7 69 2 1  
## 8 68 2 1  
## 9 70 2 1  
## 10 76 2 2  
## # ℹ 14 more rows

attach(achiev)  
boxplot(ac~teach,data = achiev, main ='Achievement as Function of teacher',  
 xlab = 'Teacher',ylab = 'Achievement')

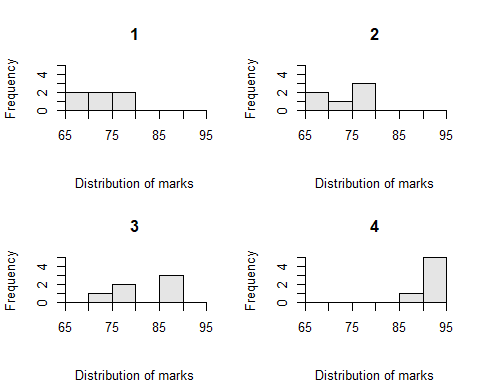


## Evaluating Assumptions

library(FSA)

## ## FSA v0.9.5. See citation('FSA') if used in publication.  
## ## Run fishR() for related website and fishR('IFAR') for related book.

f.teach <- factor(teach)  
hist(ac~f.teach, data = achiev,sub='Distribution of marks', xlab = '')



* ANOVA assumes all data is more or less normal, since the model assumes errors are for all data points aross all levels.
* However, as can be seen from the histograms across the 4 levels of teacher, the assumption of normality seems like a stretch
* But another issue here is that the data (in each level/group) is only 6 points, hence the density is actually too low to infer whether or not a particular distribution is suitable for every level.
* As a result, non-graphical- and hence quantitative tests- may be used to infer of the assumption of normality holds.

### Inferential Tests for Normality

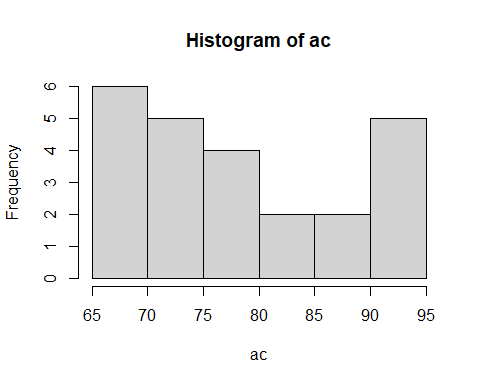
### Shapiro-wilk Test for normality

shapiro.test(ac)

##   
## Shapiro-Wilk normality test  
##   
## data: ac  
## W = 0.90565, p-value = 0.02842

The p-value of this test is <0.05 , hence is significant. Thus we can reject the Null Hypothesis that the data is normal–> Normality assumption is violated.

hist(ac)



* The above histogram of the variable **ac** now gives a visual confirmation that the data is likely not mormal

### Homogeneity of the Variances

We wish to test the following assumption in ANOVA,

The 3 common tests that are used are the *Fligner–Killeen* test, *Bartlett* test, and *Levene’s* test.

* Fligner-Killeen test is non-parametiric and hence robust to violations of normality and other distribution based assumptions.

fligner.test(ac~f.teach, data = achiev)

##   
## Fligner-Killeen test of homogeneity of variances  
##   
## data: ac by f.teach  
## Fligner-Killeen:med chi-squared = 10.813, df = 3, p-value = 0.01278

Here again, p-value is 0.012 << 0.05, so we again reject the null hypothesis that the varianecs are equal.

* Caution: We have rejected the assumption pf homogeneous variance at a level of 5%(p-value), but a more stringent level of say 1% would fail to reject .

Now let’s look at the actual values of sample variances for each level of the data

aggregate(ac~f.teach, FUN = var)

## f.teach ac  
## 1 1 19.600000  
## 2 2 15.500000  
## 3 3 35.200000  
## 4 4 5.066667

**ISSUE** : The largest variance is around 7 times that of the smallest variance . Hence, we are inclined to reject the assumption of equal variances, prompting a more robust test ~ such as the *Welch test* , in place of the usual standard F-Test for ANOVA

* We will subsequently compare unadjusted (for variance homogeneity) ANOVA test vs Welch and see the impact of violation of the assumption on our test.

Let’s also look at the mean responses by each level of teacher in the data

aggregate(ac~f.teach, FUN = mean)

## f.teach ac  
## 1 1 71.00000  
## 2 2 72.50000  
## 3 3 80.00000  
## 4 4 92.66667

Initial Observations: \* Means for 1 & 2 are very close \* teacher 3 and 4 are far apart with 4 having the highest mean at 92%

## Performing ANOVA using aov()

As our teacher and achievement data is 1-way with *teach* as the factor (with 4 levels), the anova will require treating teach as a factor variable.

f.teach<- factor(teach)  
f.teach

## [1] 1 1 1 1 1 1 2 2 2 2 2 2 3 3 3 3 3 3 4 4 4 4 4 4  
## Levels: 1 2 3 4

Next we again compute the means of each level of teach, this time by using **tapply()** function

tapply(ac, f.teach, mean)

## 1 2 3 4   
## 71.00000 72.50000 80.00000 92.66667

we can also compute associated medians

tapply(ac, f.teach, median)

## 1 2 3 4   
## 71.5 72.5 80.5 93.5

as well as variances

tapply(ac, f.teach, var)

## 1 2 3 4   
## 19.600000 15.500000 35.200000 5.066667

WE WILL ALSO REQUIRE THE **GRAND MEAN** OVER ALL THE DATA: Simply given by,

mean(ac)

## [1] 79.04167

* Balanced data: Same number of samples (data points) for each cell(2-way) or each level(1-way) and Unbalanced means different sample sizes across levels or cells.
* In the balanced data case, the overall mean (grand mean ) is the same as the average of all the level-wise means of the data. However, if the data was **unbalanced** the grand mean would just be the average of all data points, this would be different from the average of the 4 group means

### ANOVA and its summary table

anova.fit<- aov(ac~f.teach, data = achiev)

summary(anova.fit)

## Df Sum Sq Mean Sq F value Pr(>F)   
## f.teach 3 1764.1 588.0 31.21 9.68e-08 \*\*\*  
## Residuals 20 376.8 18.8   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Conclusion: F value is wayyy above the expected value of 1, with a vary significant p-value as indicated by the signif codes.

Thus it would be correct to reject , and conclude that there is an actual differential effect across the levels.

### Differential treatment effect

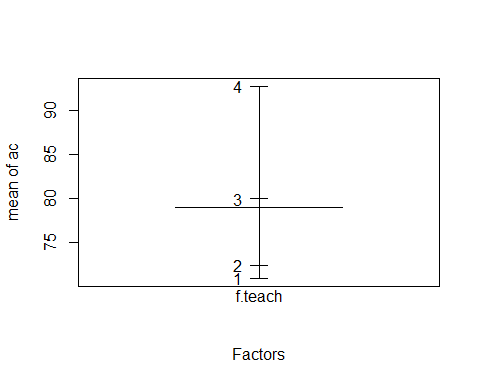
model.tables(anova.fit)

## Tables of effects  
##   
## f.teach   
## f.teach  
## 1 2 3 4   
## -8.042 -6.542 0.958 13.625

### Plotting the results of the 1-way ANOVA

A graph indicating the level of treatment effects, relative to the grand mean

plot.design(ac~f.teach)



### Post-hoc tests and further analysis

To know which pair of levels in f.teach are significantly different, requires a post-hoc analysis after the anova F-Test. This is usually the Tukey test based on the studentized range distribution.

Tu\_HSD<-TukeyHSD(anova.fit)  
Tu\_HSD

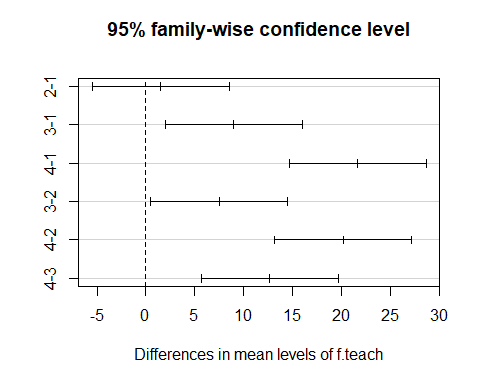
## Tukey multiple comparisons of means  
## 95% family-wise confidence level  
##   
## Fit: aov(formula = ac ~ f.teach, data = achiev)  
##   
## $f.teach  
## diff lwr upr p adj  
## 2-1 1.50000 -5.5144241 8.514424 0.9313130  
## 3-1 9.00000 1.9855759 16.014424 0.0090868  
## 4-1 21.66667 14.6522425 28.681091 0.0000002  
## 3-2 7.50000 0.4855759 14.514424 0.0334428  
## 4-2 20.16667 13.1522425 27.181091 0.0000006  
## 4-3 12.66667 5.6522425 19.681091 0.0003278

Which returns 6 confidence intervals corresponding to 6 unique mean comparisons.

* 1st row[pairs 1&2] - diff of 1.5, and wp 95% pf being betweem -5.5 and 8.5 (contains 0) . P-value of 0.93 (very high). Conclusion: fail to reject , the difference in means between level 1 and 2 is not significant.
* 2nd row [pairs 3&1 ] - diff of 9 , CI (1.98, 16.01) with 95% prob. The p-value is 0.009 (<0.05), hence the difference is statistically significant.

### Visualizing the Tukey HSD test of multiple comparisons

plot(Tu\_HSD)



## Doing ANOVA with lm()

ANOVA one of many *general linear models* and is infact a speacial case of the *regression model*. As such, ANOVA can be performed by functions such as **lm()** which are used to evaluate linear models, besides dedicated functions such *aov()*.

* Basically, regression on the mean responses of each group can be done via lm(), against the teach, by considering it as a categorical variable ( like in the case of multinomial or logistic regression). The key idea is to perform regression with the levels of factor variable as independent predictors.

anova.lm <- lm(ac~f.teach)  
summary(anova.lm)

##   
## Call:  
## lm(formula = ac ~ f.teach)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -7.0000 -3.7500 0.8333 3.6250 6.0000   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 71.000 1.772 40.066 < 2e-16 \*\*\*  
## f.teach2 1.500 2.506 0.599 0.55620   
## f.teach3 9.000 2.506 3.591 0.00183 \*\*   
## f.teach4 21.667 2.506 8.646 3.44e-08 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 4.341 on 20 degrees of freedom  
## Multiple R-squared: 0.824, Adjusted R-squared: 0.7976   
## F-statistic: 31.21 on 3 and 20 DF, p-value: 9.677e-08

CAUTION:In the summary output, intercept is actually the value for f.teach1, since lm is not particularly designed for ANOVA, there will be some differences in the nomenclature.

## Factorial ANOVA

WHEN THERE IS MORE THAN 1 INDEPENDENT VARIABLE AS LEVELS IN ANY ANOVA DESIGN, WE REFER TO IT AS FACTORIAL ANOVA.

### A Key advantage of the Factorial ANOVA model

The question: Why use a different model, why not just use multiple 1 way models ? Reason: the factorial model can actually account for **interaction** terms between the 2(or more) factor variables that we consider. This is very important as we need not make a model that has factors that necessarily independent of each other and do not react.

NOTE: An interaction between two factors is indicated by nonparallel lines. In other words, mean differences on one factor are not consistent across levels of a second factor.

## Example of factorial ANOVA

head(achiev)

## # A tibble: 6 × 3  
## ac teach text  
## <dbl> <dbl> <dbl>  
## 1 70 1 1  
## 2 67 1 1  
## 3 65 1 1  
## 4 75 1 2  
## 5 76 1 2  
## 6 73 1 2

* First factor is teach, second is text

Again, need to turn these columns into actual factors

f.teach<-factor(teach)  
f.text<-factor(text)  
f.teach

## [1] 1 1 1 1 1 1 2 2 2 2 2 2 3 3 3 3 3 3 4 4 4 4 4 4  
## Levels: 1 2 3 4

f.text

## [1] 1 1 1 2 2 2 1 1 1 2 2 2 1 1 1 2 2 2 1 1 1 2 2 2  
## Levels: 1 2

Now for the 2-way factorial model

fit.factorial <- aov(ac~f.teach + f.text + f.teach:f.text, data = achiev)  
summary(fit.factorial)

## Df Sum Sq Mean Sq F value Pr(>F)   
## f.teach 3 1764.1 588.0 180.936 1.49e-12 \*\*\*  
## f.text 1 5.0 5.0 1.551 0.231   
## f.teach:f.text 3 319.8 106.6 32.799 4.57e-07 \*\*\*  
## Residuals 16 52.0 3.3   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

library(phia)

## Loading required package: car

## Loading required package: carData

## Registered S3 methods overwritten by 'car':  
## method from  
## hist.boot FSA   
## confint.boot FSA

##   
## Attaching package: 'car'

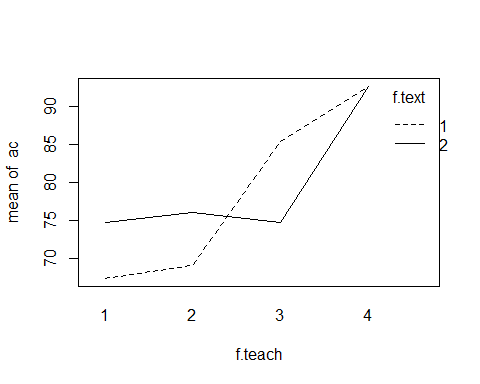
## The following object is masked from 'package:FSA':  
##   
## bootCase

fit.means <- interactionMeans(fit.factorial)  
fit.means

## f.teach f.text adjusted mean std. error  
## 1 1 1 67.33333 1.040833  
## 2 2 1 69.00000 1.040833  
## 3 3 1 85.33333 1.040833  
## 4 4 1 92.66667 1.040833  
## 5 1 2 74.66667 1.040833  
## 6 2 2 76.00000 1.040833  
## 7 3 2 74.66667 1.040833  
## 8 4 2 92.66667 1.040833

Now we can plot this fitted values to closely look at the interactions

interaction.plot(f.teach, f.text, ac)



plot(fit.means)

