ANOVA and Post-ANOVA Tests

Lab 10 R Notes: EXST 7014/15

Contents

0.1 Objectives

1. Use the **lm** function to construct ANOVA and post-ANOVA tests.

2. Use the **aov** function to construct ANOVA and post-ANOVA tests.

Analysis of variance (ANOVA) is the most commonly used technique for comparing the means of groups of measurement data. There are lots of different experimental designs that can be analyzed with different kinds of ANOVA. For this week's lab, the most basic type of ANOVA, one-way ANOVA will be introduced. In a one-way ANOVA, there is one continuous variable (Response Variable) and one categorical variable (Treatment). Multiple observations of the response variable are made for each level of the treatment.

One-way ANOVA corresponds to the completely randomized designed experiment (CRD) with one fixed treatment effect, with its linear model as the following:

$$Y_{ij} = \mu + \tau_i + \epsilon_{ij} = \mu_i + \epsilon_{ij} \quad (i = 1, 2, ..., t \;\; j = 1, 2, ..., n)$$

Where μ is the overall mean; $\tau_{\mathbf{i}}$ are the treatment level effects, and $\epsilon_{\mathbf{ij}}$ is the random error. $\mu_{\mathbf{i}}$ is the mean of the $\mathbf{i^{th}}$ level of treatment.

The null hypothesis test of one-way ANOVA is that the means of the response variable are the same for the different levels of treatment ($H_0 : \mu_1 = \mu_2 = \ldots = \mu_t$); the alternative hypothesis is that they are not all the same.

There are three assumptions need to be considered for ANOVA: the treatments are independently sampled; residuals or deviation of observations within groups should be normally distributed (evaluated by residual plot, normality test); and the variance from each level of treatment is the same (i.e. homogeneous variance, Bartlett test is preferred method to evaluate the homogeneity of variance).

ANOVA usually proceeds with an F-test of the MSTreatment ($\mathbf{d.f.} = \mathbf{t - 1}$) over MSError ($\mathbf{d.f.} = \mathbf{t(n - 1)}$). The MSError estimates a variance $\sigma^{\mathbf{2}}_{\epsilon}$, and MSTreatment estimates the same $\sigma^{\mathbf{2}}_{\epsilon}$ plus the difference between the levels of treatment ($\sigma^{\mathbf{2}}_{\epsilon} + \mathbf{n}\sigma^{\mathbf{2}}_{\tau}$). So the F-test can be written as the following:

$$F = MSTreatment/MSError = (\sigma_\epsilon^2 + n\sigma_\tau^2)/\sigma_\epsilon^2$$

It is One Tailed F-test since the variance of treatment is expected to be large if the null hypothesis is rejected.

Once the null hypothesis of ANOVA is rejected, i.e., the significant difference among the levels of treatment is concluded from F-Test of ANOVA, Post-ANOVA tests are needed to determine how the treatment levels are interrelated. Common Post hoc tests will be introduced in this lab.

0.2 Lab Setup

Run the following code to both install and load the required packages.

```
#' a function to only install needed but unavailable packages
#' and loads these packages after installation

ipak <- function(pkg){
    new.pkg <- pkg[!(pkg %in% installed.packages()[, "Package"])]
    if (length(new.pkg))
        install.packages(new.pkg, dependencies = TRUE)
    sapply(pkg, require, character.only = TRUE)
}

# use function to install and load packages
packages <- c('olsrr', 'car', 'lattice', 'agricolae', 'emmeans')
ipak(packages)


#' olsrr       # assesses model fit and variable diagnostics
#' car         # for Type II, III SS
#' lattice     # for boxplot with mean - bwplot function
#' agricolae   # for LSD Test
#' emmeans     # for post-hoc tests
```

0.3 The Data

The word lists are standard audiology tools for assessing hearing. They are calibrated to be equally difficult to perceive. However, the original calibration was performed with normal hearing subjects and no noise background. The experimenter wished to determine whether the lists were still equally difficult to understand in the presence of a noisy background. 24 subjects with normal hearing listened to standard audiology tapes of four standard 50-English-word lists at low volume with a noisy background. They repeated the words and were scored correct or incorrect in their perception of the words. The percent of words heard correctly was recorded and will be used as the response variable. Note that this is actually a randomized block design, which has not been covered. Therefore, we will pretend that the data was collected without control on the subjects and treat it as a CRD. Detailed information and references can be found at http://lib.stat.cmu.edu/DASL/Datafiles/Hearing.html. The variables in the dataset are:

- **SubjectID**: Code for each subject - 24 of them
- **List**: 4 standard lists (List1, List2, List3, List4).
- **score**: Score received on hearing test

```r
#' Download the data_lab10.txt file to your working directory
#' Create an object to host the data set
#'
#' @sep="" because the columns are seperated by 'space'
#' @colClasses= to specify the repective data classes of each column

hearing <- read.table('data_lab10.txt', header = TRUE,
                      colClasses = c("integer", "factor", "integer"))

str(hearing)  # get a structure (description) of your dataset

#View(hearing) # to view the hearing dataset in RStudio's GUI pane
```

It is imperative that you specify the categorical variable (that is, the column containing the various groups) with the factor function. There are so many ways to do this - one is shown above by indicating that the groups (i.e. second) column is of class, **factor**, in the **colClasses** argument in the **read.csv** function.

0.4 Fitting the ANOVA model

one-way ANOVA, in essence, is a linear model. This means that ANOVA models can be fit in R using linear modelling functions in R:

- lm
- aov
- lmer (for Linear Mixed and multilevel modelling - from the lme4 package)
- lme (for Linear and Nonlinear Mixed Effects Models- from the nlme package)

For this dataset, because there are no random treatment effects, and we are running a one-way ANOVA, we are going to use the **lm** and **aov** functions in R.

The codes below produce a visualization(boxplot) of the hearing score for the various the levels (or groups).

```r
#' Open a blank png file to later populate with the plot
png(filename = "plot_01.png")

#' Create the plot
boxplot(score ~ list, data = hearing,
        boxwex=0.5, col="lightsteelblue",
        main="Distribution of Hearing Score by List",
        ylab="score")
means <- tapply(hearing$score, hearing$list, mean)
points(means,col="red", pch=18, cex=1.2)

#' close the png function portal
dev.off()

#' Alternative way to chart boxplot with mean
library(lattice)
bwplot(score ~ list, data = hearing)
```

Because we are fitting the ANOVA model with the **lm** function all attributes of the lm function can me called on/applied to the model object (in this case, **lmMod** in the rchunk below)

```
#' ANOVA as a linear model
lmMod <- lm(score ~ list, data=hearing)
summary(lmMod)

#' ANOVA model
aovLM <- aov(lmMod)
summary(aovLM)

#' Assess Model goodness-of-fit
car::Anova(lmMod, type=3, test.statistic="F")

#' Note that you can equally read the ANOVA information for
#'    model goodness-of-fit test from the summary(lmMod)
```

0.5 Checking Assumptions

0.5.1 Test for Homogeneity of Variance

```
# Test for Homogeneity of Variance
bartlett.test(score ~ list, data=hearing)
```

0.5.2 Testing for Normality

```
#' (Residuals vs. Fitted Values) and QQ Plots
plot(lmMod, col=hearing$list, which=c(1,2))
```

```
#' Test for Normality of Residuals
ols_test_normality(lmMod)
```

0.6 Post-Hoc Tests

If we reject the null for ANOVA, then the next step is to run pairwise tests to determine which groups reported significantly differently means.

0.6.1 Least Significant Difference(LSD) Test

```
N <- nrow(na.omit(hearing)) # total sample size
k <- length(unique(hearing$list)) # number of groups
```

```r
# sample size per group in a balanced design (since sizes are equal)
n <- length(hearing$list) / k

mse <- function(lmModel){sum((lmModel$residuals^2)/(N-k))}

(with(hearing, LSD.test(y=score, trt=list, DFerror=(N - k), MSerror=mse(lmMod))))
```

For the **$groups** output above, groups with that share the same letter do not differ in population means significantly.

0.6.2 Tukey's Honest Significant Difference(HSD) Test

```r
#' Pairwise Tukey Test
TukeyHSD(aovLM, conf.level = 0.95)

#' plot(tukeyTest, comparisons = TRUE)
#' Set plot/chart area margins
#'    par(oma=c(b,l, t,r))
#'    b=bottom, l=left, t=top and r=right margins
par(oma=c(0,5,0,0)  # this was done to allow room for
plot(TukeyHSD(aovLM, conf.level = 0.95),las=1, col = "red")
```

0.6.3 Other R functions for Post-ANOVA Tests

The following functions offer some flexibility in the choice of p-value adjustment or post-hoc test, and choice of degree of freedom approximation method specification (e.g. Satterthwaite, Kenward Roger, etc.)

```r
#' bonferroni-adjusted pairwise.t.test function from base R
with(hearing, pairwise.t.test(score, list, p.adjust.method = "bonferroni"))

#' library(emmeans)
emmeans(lmMod, pairwise ~ list,
        adjust = "tukey",
        lmer.df = "satterthwaite")
```

Notice that for the above functions, you can change the p-value adjustment method to your preference, say, **scheffe, dunnet's**, etc.

```r
#' Check adjustment options
?pairwise.t.test

#' Get exhaustive notes with examples on group comparisons using the emmeans package
#'    Search Ouptput would be in the Help Subpane in RStudio
vignette(topic="comparisons", package = "emmeans")
```

0.7 Lab Assignment

Your assignment is to perform necessary analysis using either SAS or R to answer the following questions Only print the graphs and tables that you think are relevant to your answers.

0.7.1 Question 1

Write the linear model to test the hypothesis that there is no treatment effect. Clearly describe each term in the model, and the range of the subscripts. Write the null hypothesis that you are testing.

0.7.2 Question 2

Fit a model to test the hypothesis you stated in Question 1. Report your results, including your F-value, p-value, and conclusions.

0.7.3 Question 3

List the assumptions necessary for your analysis and determine whether they have been violated.
Include any relevant SAS or R output in you report.

0.7.4 Question 4

If there is a significant treatment effect, describe which pairs of means are different. Explain which adjustment method you chose.