SAS Tutorial for STAT 350 Lab 8

Author: Leonore Findsen, Cheng Li

Example: (Data Set: eduproduct.txt)

Evaluation of a New Educational Product. A company designed two versions of a new educational product to improve children's reading comprehension. They would like to claim that children will acquire better reading comprehension skills utilizing either of the two versions than with the traditional approach. You are helping the company to prepare the marketing material, which will include the results of a study conducted to compare the two versions of the new product with the traditional method. The standard method is called Basal (B), and the two variations of the new method are called DRTA (D) and Strat (S).

Education researchers randomly divided 66 children into three groups of 22. Each group was taught by one of the three methods. The response variable is a measure of reading comprehension called Comp that was obtained by a test taken after the instruction was completed.

a) Make side-by-side boxplots and an effects plot of the data. Also, make a table containing the sample size, mean, and standard deviation for each treatment group. From this information, do you think that all of the means are the same? Be sure to comment on each of the plots.
b) Examine the assumptions necessary for ANOVA. Is it appropriate to continue the analysis? Be sure to state each of the assumptions including those ones that you are assuming to be true, and comment on each of them using the appropriate plots/data. Remember, you need to generate the normal probability plots and histograms for each population.
c) Report the results of the ANOVA significance test (four steps) using a significance level of 0.05. Are your results in this part consistent with part a)?
d) Use an appropriate multiple-comparison method to determine which of the educational method(s) helps reading comprehension the best. To determine which method is best, we need to compare all pairs of treatment methods. Explain why you chose this multiple-comparison method. Present a graphical representation of the results if appropriate for your method. Write a short statement for your conclusion in complete English sentences.

Write a short report explaining the treatment effects of this three educational methods. Be sure to answer the question posed in this question and how far the findings of this study can be generalized. This paragraph should be written in complete English sentences and should be understandable to someone who has not taken a course in Statistics.

Solution:

```
* Read the data;
data ed;
  infile 'W:\STAT350\eduproduct.txt' firstobs = 2 delimiter = '09'x;
  input Subject GroupName $ Comp;
run;

* Some of the following procedures need the data sorted;
proc sort data = ed;
  by GroupName;
run;
```

SAS Tutorial for STAT 350 Lab 8

Author: Leonore Findsen, Cheng Li

```
* a) Effects plot and table of descriptive statistics;
* side-by-side boxplots are included automatically in proc glm;
* proc means gives the descriptive statistics (group size, mean, SD);
* We also use it to save the group averages that will be used to
*   create the effects plot;
proc means data = ed;
  var Comp; * response variable;
  /* Using the class statement, instead of the by statement, */
  /* would put the descriptive statistics in one table.*/
  class GroupName;  * categorical variable;
  /* create a data set named 'means,' in which we save the   */
  /* group means to the variable 'average'.                  */
  output out = means mean = average;
run;

* Effects plot;
symbol1 v = dot i = join; * to make the effects plot 'pretty';
proc gplot data = means;
  plot average * GroupName; * y = numeric (average);
                            * x = categorical (GroupName);
run;

* b) This time, we need to create the histograms and QQ plots manually;
* Histogram;
proc sgplot data = ed;
  by GroupName; * to create the histogram for each of the groups;
  histogram Comp;
  density Comp; * to add the normal density curve;
  density Comp / type = kernel; * to add the kernel density curve;
run;

* QQ plot;
proc univariate data = ed noprint;
  /* noprint: only thing printed are the requested graphs. */
  /* We are generating the descriptive statistics in proc means. */
  by GroupName; * to create the QQ plot for each of the groups;
  qqplot Comp / normal(mu = est sigma = est);
run;

* c), d) ANOVA and multiple comparisons;
proc glm data = ed alpha = 0.05; * glm stands for general linear model;
  class GroupName; * categorical variable;
  model Comp = GroupName; * response variable = categorical variable;
  means GroupName / Tukey cldiff; * multiple comparisons;
run;
```

a) Make side-by-side boxplots and an effects plot of the data. Also, make a table containing the sample size, mean, and standard deviation for each treatment group. From this information, do you think that all of the means are the same? Be sure to comment on each of the plots.

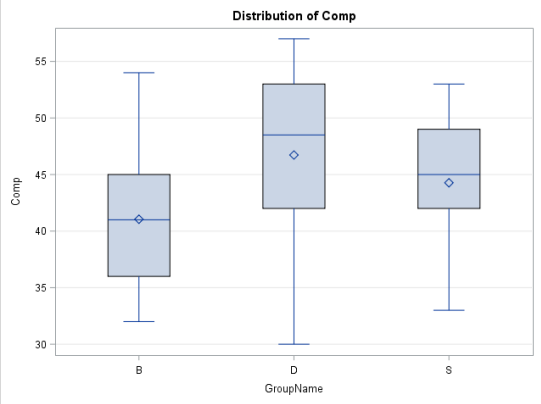
SAS Tutorial for STAT 350 Lab 8

Author: Leonore Findsen, Cheng Li

Solution:

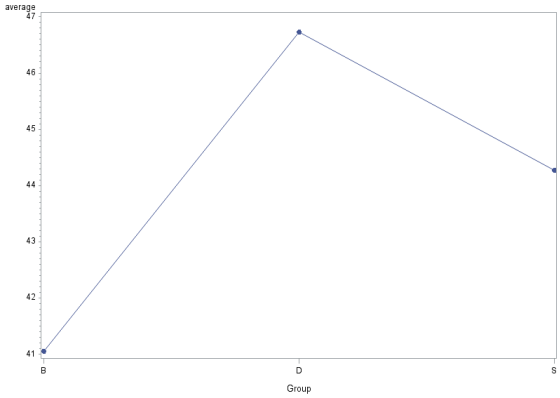
Note: The side-by-side boxplots are generated automatically from proc glm (the procedure used for running ANOVA). The rest of the graphs needs to be generated separately using other procedures.

side-by-side boxplots:



From this plot, I would state that all of the means could be the same, since the boxes overlap a fair amount.

effects plot:



From the effects plot, we can get a sense of the ranking of the means, if they turn out to be statistically different.

Calculating the sample size, mean and standard deviation per category:

The SAS System

The MEANS Procedure

Analysis Variable : Comp						
Group	N Obs	N	Mean	Std Dev	Minimum	Maximum
B	22	22	41.0454545	5.6355781	32.0000000	54.0000000
D	22	22	46.7272727	7.3884196	30.0000000	57.0000000
S	22	22	44.2727273	5.7667505	33.0000000	53.0000000

When using the commands here, there is no need to create another table. Note that SAS automatically alphabetizes the group names.

It appears group B may have a smaller mean than the other two, but inference needs to be performed (both c) and d)) to determine if B is lower than the other treatment groups or they are all the same.

SAS Tutorial for STAT 350 Lab 8

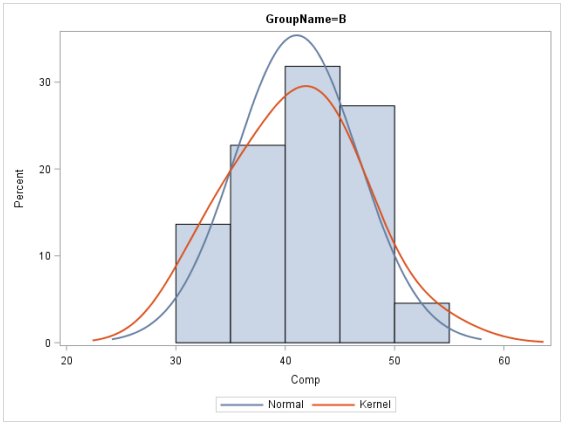
Author: Leonore Findsen, Cheng Li

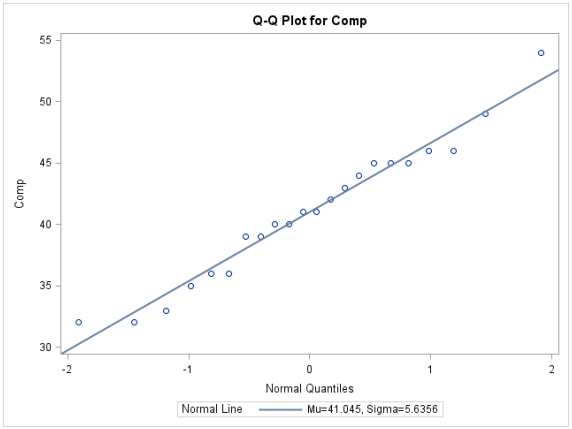
b) Examine the assumptions necessary for ANOVA. Is it appropriate to continue the analysis? Be sure to state each of the assumptions including those ones that you are assuming to be true, and comment on each of them using the appropriate plots/data. Remember, you need to generate the normal probability plots and histograms for each population.

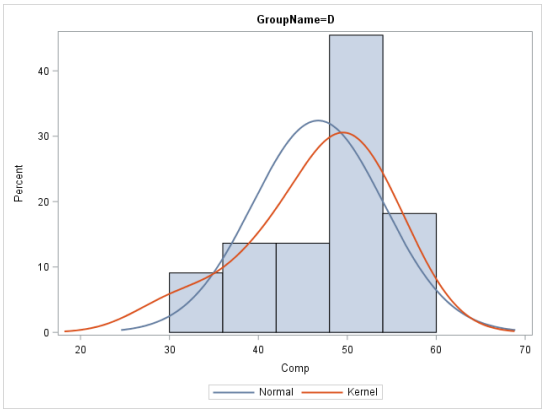
Solution:

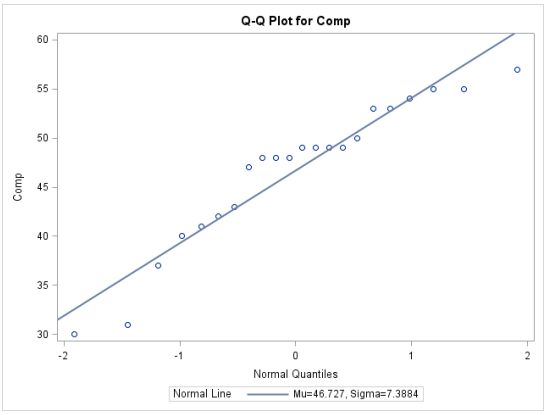
SRS / independent errors:
Assumed to be true.

Normality:

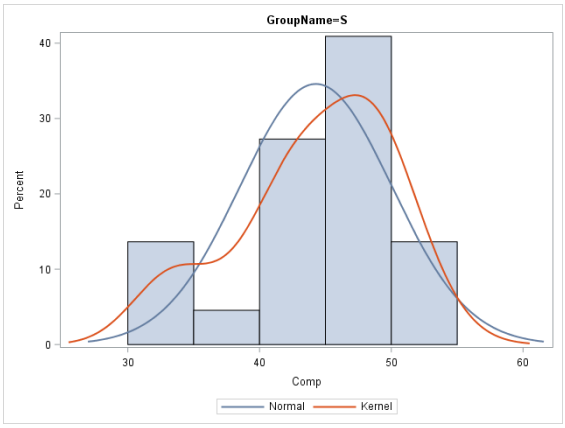


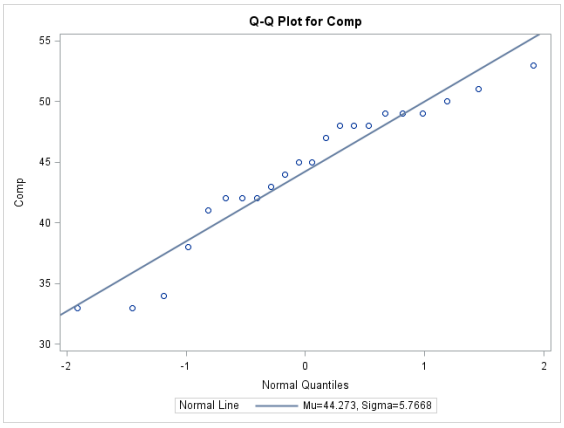






Author: Leonore Findsen, Cheng Li





The group name is not on the QQ plots, but in the output; therefore, be sure that you keep them with their respective histograms or label them in your report so that you know which plot is for which group.

At the given sample size, these distributions are close enough to being normal.

Constant standard deviation

Analysis Variable : Comp						
Group	N Obs	N	Mean	Std Dev	Minimum	Maximum
B	22	22	41.0454545	5.6355781	32.0000000	54.0000000
D	22	22	46.7272727	7.3884196	30.0000000	57.0000000
S	22	22	44.2727273	5.7667505	33.0000000	53.0000000

$$\frac{s_{max}}{s_{min}} = \frac{7.3884196}{5.6355781} = 1.31 < 2$$

Therefore, the constant standard deviation assumption is valid.

Because all of the assumptions are valid, it is appropriate to continue the analysis.

c) Report the results of the ANOVA significance test (four steps) using a significance level of 0.05. Are your results in this part consistent with part a)?

SAS Tutorial for STAT 350 Lab 8

Author: Leonore Findsen, Cheng Li

Solution:

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	357.303030	178.651515	4.48	0.0152
Error	63	2511.681818	39.867965		
Corrected Total	65	2868.984848			

Step 1: Define parameters
μ_B is the population mean Comp score for the Basal method (control).
μ_D is the population mean Comp score for the DRTA method.
μ_S is the population mean Comp score for the Strat method.

Step 2: State the hypotheses
H_0: $\mu_B = \mu_D = \mu_S$
H_a: at least two μ_i's are different.

Step 3: Find the *test statistic, p-value, report DF*

$F_{ts} = 4.48$
DF1 = 2, DF2 = 63
p-value = 0.0152

Step 4: Conclusion:
$\alpha = 0.05$
Since 0.0152 < 0.05, we should reject H_0

The data provides evidence (p-value = 0.0152) to the claim that the population mean values of at least one of the education methods is different from the rest.

In part a), I stated that the methods might have been different, but I wasn't sure because the three groups looked potentially the same from the boxplot. Here, the test based on ANOVA indicated a difference. Therefore, the conclusions are partially consistent. Keep in mind that the results of the test are more objective than the subjective analysis of the plots.

d) Use an appropriate multiple-comparison method to determine which of the educational method(s) helps reading comprehension the best. To determine which method is best, we need to compare all pairs of treatment methods. Explain why you chose this multiple-comparison method. Present a graphical representation of the results if appropriate for your method. Write a short statement for your conclusion in complete English sentences.

Solution:

The Tukey method was chosen because we want to compare all of the means in a pairwise fashion.

Author: Leonore Findsen, Cheng Li

Tukey's Studentized Range (HSD) Test for Comp

Note: This test controls the Type I experimentwise error rate.

Alpha	0.05
Error Degrees of Freedom	63
Error Mean Square	39.86797
Critical Value of Studentized Range	3.39448
Minimum Significant Difference	4.5696

Comparisons significant at the 0.05 level are indicated by ***.

Group Comparison	Difference Between Means	Simultaneous 95% Confidence Limits		
D - S	2.455	-2.115	7.024	
D - B	5.682	1.112	10.251	***
S - D	-2.455	-7.024	2.115	
S - B	3.227	-1.342	7.797	
B - D	-5.682	-10.251	-1.112	***
B - S	-3.227	-7.797	1.342	

What is included in the **red box** is the only output that needs to be presented.

Note that although the comparison of each pair is given twice, the order of the two groups does not influence the conclusions so either of the two may be used to create the visual comparison.

In this case, we have statistical evidence that D and B are different but we do not have evidence that B is different from S nor that D is different from S.

The following is the graphical representation (to determine the values of the means, please look at the table that you created in part a). You must construct this by hand.

B	S	D
41.0454545	44.2727273	46.7272727

Note that the above is a table was created in Word where I used the Border Painter in Table Tools → Design to only include the appropriate Borders.

This says that that, in terms of statistical significance, methods D and S are the same and methods S and B are the same (because they are joined by an underline). Since we are interested in the method that leads to the largest value, we should eliminate method B because it is significantly different from method D: The second underline only joins S and D. Meanwhile, note that method S is not significantly different from the worst method, B. Therefore, I would choose educational method D out of S and D.

e) Write a short report explaining the treatment effects of this three educational methods. Be sure to answer the question posed in this question and how far the findings of this study can be generalized. This paragraph should be written in complete English sentences and should be understandable to someone who has not taken a course in Statistics.

Solution:

The original question asked us to determine the best educational method and whether the new methods D and S are better than the traditional method, B. We determined that the ANOVA assumptions were valid, so we performed an ANOVA analysis to infer about the population reading comprehension scores, Comp, associated with these methods. These results show that method S is not statistically different from method B (traditional method). However, we have evidence that method D is better than the original method. The study can be generalized to other children who are at similar levels of reading comprehension as the children in the study.

Remember when discussing the conclusion in other situations, first determine whether you are interested in the lowest number(s) or the highest number(s).