Introductory Computing for Statistics

Lecture 4:t-test & ANOVA

Xiao Li

October 28, 2017

Highlights from Lecture 3

- LABEL and TITLE statement
- PROC UNIVARIATE, PROC MEANS, PROC FREQ
- PROC CHART, PROC PLOT

Today's topic

1. t-test: PROC MEANS, PROC TTEST
    - One-sample t-test
    - Paired t-test
    - Two-sample t-test
2. ANOVA: PROC GLM, PROC ANOVA
    - One-way ANOVA
    - Two-way ANOVA

One-sample t-test

One-sample t-test is usually used to test whether the **mean** of a normally distributed population has a value specified in a null hypothesis. The null and alternative hypotheses are

$$H_0 : \mu = 0 \text{ vs. } H_1 : \mu \neq 0$$

General form: One-sample t-test

```
PROC MEANS Data=dataset N MEAN STD T PRT;
VAR variables;
RUN;
```

- The option T requests t-statistic.
- The option PRT requests the p-value for the two-sided test.

One-sample t-test

One-sample t-test is usually used to test whether the **mean** of a normally distributed population has a value specified in a null hypothesis. The null and alternative hypotheses are

$$H_0 : \mu = 0 \text{ vs. } H_1 : \mu \neq 0$$

General form: One-sample t-test

```
PROC MEANS Data=dataset N MEAN STD T PRT;
VAR variables;
RUN;
```

- The option T requests t-statistic.
- The option PRT requests the p-value for the two-sided test.

One-sample t-test-Variations

1. If the true mean is **not zero**. i.e

$$H_0 : \mu = \mu_0 \text{ vs. } H_1 : \mu \neq \mu_0$$

   where $\mu_0 \neq 0$, then just subtract μ_0 from the tested variable in the DATA step.

2. One-sided test. Report the modified p-value from two-sided test as following:
   - $H_0 : \mu = \mu_0$ vs. $H_1 : \mu > \mu_0$, then one-sided probability is $p/2$ if $t > 0$, and is $1 - p/2$ if $t < 0$.
   - $H_0 : \mu = \mu_0$ vs. $H_1 : \mu < \mu_0$, then one-sided probability is $1 - p/2$ if $t > 0$, and is $p/2$ if $t < 0$.

One-sample t-test-Variations

1. If the true mean is **not zero**. i.e

$$H_0 : \mu = \mu_0 \text{ vs. } H_1 : \mu \neq \mu_0$$

   where $\mu_0 \neq 0$, then just subtract μ_0 from the tested variable in the DATA step.

2. One-sided test. Report the modified p-value from two-sided test as following:
   - $H_0 : \mu = \mu_0$ vs. $H_1 : \mu > \mu_0$, then one-sided probability is $p/2$ if $t > 0$, and is $1 - p/2$ if $t < 0$.
   - $H_0 : \mu = \mu_0$ vs. $H_1 : \mu < \mu_0$, then one-sided probability is $1 - p/2$ if $t > 0$, and is $p/2$ if $t < 0$.

One-sample t-test-Example

Example 4.1: Recall the corn data in the last homework. Now we are interested in the following questions:

- Is the (true) average rain different from 10?
- Is the (true) average yield different from 30?

One-sample t-test-Example Output

One Sample t-Test

The MEANS Procedure

Variable	N	Mean	Std Dev	t Value	Pr > |t|
yield1	19	3.1263158	3.4936199	3.90	0.0010
rain1	19	0.4947368	2.2621213	0.95	0.3531

One-sample t-test-Example Conclusion

- The hypotheses are (for rain variable)

$$H_0 : rain = 10 \text{ vs. } H_1 : rain \neq 10$$

The t-statistic for this test is **0.95** and the corresponding p-value is **0.3531**. So we **cannot reject** the null hypothesis at 5% level and conclude that the true average of rain is not different from 10.

- The hypotheses are (for yield variable)

$$H_0 : yield = 30 \text{ vs. } H_1 : yield \neq 30$$

The t-statistic for this test is **3.90** and the corresponding p-value is **0.0010**. So we **reject** the null hypothesis at 5% level and conclude that the true average of yield is statistically significantly different from 30.

One-sample t-test-Example Conclusion

- The hypotheses are (for rain variable)

$$H_0 : rain = 10 \text{ vs. } H_1 : rain \neq 10$$

  The t-statistic for this test is **0.95** and the corresponding p-value is **0.3531**. So we **cannot reject** the null hypothesis at 5% level and conclude that the true average of rain is not different from 10.

- The hypotheses are (for yield variable)

$$H_0 : yield = 30 \text{ vs. } H_1 : yield \neq 30$$

  The t-statistic for this test is **3.90** and the corresponding p-value is **0.0010**. So we **reject** the null hypothesis at 5% level and conclude that the true average of yield is statistically significantly different from 30.

Paired t-test

Paired t-tests: When the two samples are NOT independent.

- Create the differences as a new variable in the DATA step;
- Then apply PROC MEANS on this new variable (I.e. Perform a t-test on differences).

Paired t-test-Example

Example 4.2: Test whether the water qualities are same in spring and summer.

```
TITLE1 'Paired Sample t-Test';
TITLE2 'Seasonal Total Suspended Solid (TSS)';
/*TSS is water quality measurement */
DATA TSS;
INPUT Station Spring Summer @@;
Change = Summer - Spring; /*new variable of differences*
DATALINES;
1 80 82 2 73 71 3 70 95 4 60 69
5 88 100 6 84 71 7 60 69 8 37 60
9 91 95 10 98 99 11 52 65 12 78 83
13 40 60 14 79 86 15 59 62
;
PROC MEANS N MEAN STDERR T PRT;
VAR Change; RUN;
```

Paired t-test-Example Output

One Sample t-Test

The MEANS Procedure

Analysis Variable : Change				
N	Mean	Std Error	t Value	Pr > |t|
15	7.8666667	2.5613736	3.07	0.0083

Paired t-test-Example Conclusion

The hypotheses are

$$H_0 : summer = spring \text{ vs. } H_1 : summer \neq spring$$

or equivalently, to say

$$H_0 : change = 0 \text{ vs. } H_1 : change \neq 0$$

The t-statistic is **3.07** and the corresponding p-value of this test is **0.0083**, which is less than 0.05.So we **reject** the null hypothesis at 5% level and conclude that the TSS in summer and spring are different.

Two-sample t-test

PROC TTEST tests whether two population means are the same or not (two sided test). It will also produce an F test to test whether the variances from the two groups are the same or not. (H_0 : Variances are equal.)

- If the **variances are not equal**, use the two-sample t-test. On the printout the two-sample t-test statistic is found on the line beginning with 'Unequal'.
- If the **variances are equal**, use the pooled t-test. On the printout the two-sample t-test statistic is found on the line beginning with 'Equal'.

Two-sample t-test- PROC TTEST

Syntax PROC TTEST

```
PROC TTEST DATA = dataset;
CLASS variables; /* Identifier of two groups */
/* which has only two values (numeric or character) */
VAR variables;
RUN;
```

Two-sample t-test- Example

Example 4.3 Example 4.1 cont.

```
TITLE1 'Modified Dataset Corn';
DATA corn;
SET corndata;
IF rain<=9.7 THEN catrain='drought';
ELSE catrain='normal';
KEEP yield catrain;
;
RUN;
TITLE1 'Two Sample t-Test';
PROC TTEST data=corn;
CLASS catrain;
VAR yield;
RUN;
```

Two-sample t-test- Example Output

Two Sample t-Test

The TTEST Procedure
Variable: yield

catrain	N	Mean	Std Dev	Std Err	Minimum	Maximum
drought	9	31.9556	3.2110	1.0703	26.8000	35.5000
normal	10	34.1800	3.5555	1.1244	26.8000	38.3000
Diff (1-2)		-2.2244	3.3977	1.5612		

catrain	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
drought		31.9556	29.4874	34.4237	3.2110	2.1689	6.1515
normal		34.1800	31.6365	36.7235	3.5555	2.4456	6.4910
Diff (1-2)	Pooled	-2.2244	-5.5182	1.0693	3.3977	2.5496	5.0937
Diff (1-2)	Satterthwaite	-2.2244	-5.4996	1.0507			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	17	-1.42	0.1723
Satterthwaite	Unequal	16.998	-1.43	0.1700

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	9	8	1.23	0.7844

Two-sample t-test- Example Conclusion (mean)

Under the assumption of unequal/equal variances, the TTEST procedure displays results using Satterthwaite's method/by pooling the group variances.

- **t Value** : an approximate t statistic for testing the null hypothesis that the means of the two groups are equal
- **DF** : approximate degrees of freedom
- $Pr > |t|$:probability of a greater absolute value of t under the null hypothesis. This is the two-tailed significance probability. The one-tailed probability is computed the same way as in a one-sample t test.

T-test: $H_0 : \mu_{drought} = \mu_{mormal}$ vs $H_1 : \mu_{drought} \neq \mu_{normal}$. **Don't reject**

Two-sample t-test- Example Conclusion (mean)

Under the assumption of unequal/equal variances, the TTEST procedure displays results using Satterthwaite's method/by pooling the group variances.

t Value : an approximate t statistic for testing the null hypothesis that the means of the two groups are equal

DF : approximate degrees of freedom

$Pr > |t|$:probability of a greater absolute value of t under the null hypothesis. This is the two-tailed significance probability. The one-tailed probability is computed the same way as in a one-sample t test.

T-test: $H_0 : \mu_{drought} = \mu_{mormal}$ vs $H_1 : \mu_{drought} \neq \mu_{normal}$. **Don't reject**

Two-sample t-test- Example Conclusion (variance)

PROC TTEST then gives the results of the test of equality of variances:

F Value : the F' (folded) statistic

Num DF and Den DF : numerator and denominator degrees of freedom in each group

Pr>F : probability of a greater F' value. This is the two-tailed significance probability.

F-test: $H_0 : \sigma_1^2 = \sigma_2^2$ vs. $H_1 : \sigma_1^2 \neq \sigma_2^2$, $p = 0.7844$. Therefore we **do not reject** the null hypothesis at $\alpha = 0.05$ level, i.e. regardless of the amount of rain, there is no statistical evidence that there is a difference in the variability of the yield of corn. Therefore we preform a pooled t-test.

Two-sample t-test- Example Conclusion (variance)

PROC TTEST then gives the results of the test of equality of variances:

F Value : the F' (folded) statistic

Num DF and Den DF : numerator and denominator degrees of freedom in each group

Pr>F : probability of a greater F' value. This is the two-tailed significance probability.

F-test: $H_0 : \sigma_1^2 = \sigma_2^2$ vs. $H_1 : \sigma_1^2 \neq \sigma_2^2$, $p = 0.7844$. Therefore we **do not reject** the null hypothesis at $\alpha = 0.05$ level, i.e. regardless of the amount of rain, there is no statistical evidence that there is a difference in the variability of the yield of corn. Therefore we preform a pooled t-test.

Two-sample t-test- Simulation Example

Example 4.4: simulated two samples
Simulation: sample 1$\sim N(20, 100)$, sample 2$\sim N(51, 100)$
Simulation: sample 1$\sim N(20, 100)$, sample 2$\sim N(51, 400)$
Code: ▸ Example 4.4
Note: You need to decide whether equal or unequal variances are appropriate for your data, then report the proper p-value.

One-way ANOVA

- **Model**

  Suppose we have K groups and group j has p_j observations,

  $$y_{ij} = \mu + a_i + e_{ij}, i = 1, ..., K, j = 1, ..., p_j$$

  where $e_{ij} \sim_{i.i.d} N(0, \sigma^2)$, μ is common mean and a_i is group effect.
- One-way ANOVA is a simple extension of two-sample t-test. In a two-sample t-test, we are interested in whether two groups have the same mean, while in one-way anova, we are testing whether two or more groups have equal means.

One-way ANOVA

- **Model**

  Suppose we have K groups and group j has p_j observations,

  $$y_{ij} = \mu + a_i + e_{ij}, i = 1, ..., K, j = 1, ..., p_j$$

  where $e_{ij} \sim_{i.i.d} N(0, \sigma^2)$, μ is common mean and a_i is group effect.
- One-way ANOVA is a simple extension of two-sample t-test. In a two-sample t-test, we are interested in whether two groups have the same mean, while in one-way anova, we are testing whether two or more groups have equal means.
- The variable that separates the data into these groups is referred to as a factor. It can be either character or numeric, and can have two or more levels.
- Another way to describe: a one-way anova tests whether a certain factor is statistically significant in a linear model. I.e. One-way ANOVA tests

  $$H_0 : a_1 = a_2 =, ..., = a_K \text{ vs } H_1 : \text{at least one } a_i \neq 0$$

One-way ANOVA

- **Model**
  Suppose we have K groups and group j has p_j observations,
  $$y_{ij} = \mu + a_i + e_{ij}, i = 1, ..., K, j = 1, ..., p_j$$
  where $e_{ij} \sim_{i.i.d} N(0, \sigma^2)$, μ is common mean and a_i is group effect.
- One-way ANOVA is a simple extension of two-sample t-test. In a two-sample t-test, we are interested in whether two groups have the same mean, while in one-way anova, we are testing whether two or more groups have equal means.
- The variable that separates the data into these groups is referred to as a factor. It can be either character or numeric, and can have two or more levels.
- Another way to describe: a one-way anova tests whether a certain factor is statistically significant in a linear model. I.e. One-way ANOVA tests
  $$H_0 : a_1 = a_2 =, ..., = a_K \text{ vs } H_1 : \text{at least one } a_i \neq 0$$

One-way ANOVA PROC GLM and PROC ANOVA

PROC GLM and PROC ANOVA share the same form.

Syntax GLM

PROC GLM DATA=dataset;
BY variables;
CLASS variables;
MODEL dep-var = indep-var(s);
MEANS effects < /*options* >;
RUN;

Syntax ANOVA

PROC ANOVA DATA=dataset;
BY variables;
CLASS variables;
MODEL dep-var = indep-var(s);
MEANS effects < /*options* >;
RUN;

One-way ANOVA-Difference for PROC GLM and ANOVA

Table: Differences between PROC GLM and PROC ANOVA

Procedures	for balanced data?	for unbalanced data?	Efficiency for balanced data
GLM	YES	YES	LOW
ANOVA	YES	NO	HIGH

One-way ANOVA- MEANS Option

Table: MEANS Option List

option	description
alpha=.01	Significance level used in multiple comparisons. Default is .05.
bon	Request Bonferroni t-test of differences between means.
cldiff	Request confidence intervals for all pairwise differences between means.
duncan	Request Duncan's multiple comparisons.
lines	List the means in descending order, indicating those means not significantly different with a line segment beside them.
lsd	Request pairwise t-tests, which is equivalent to Fisher's least-significant -difference test when cell sizes are equal.
scheffe	Request Scheffe's multiple comparisons.
snk	Requests the student-Newman-Keuls multiple range test.
tukey	Request Tukey's studentized range test.

One-way ANOVA-Example

Example 4.5 Taillite Example:
Researchers want to test whether the average braking time of drivers follows different types of trucks equipped with center high-mounted stop lamp (CHMSL) are the same or not
Code: ▸ Example 4.5

Conclusion: Output shows that the p-value is 0.0109, this means we **reject** the null hypothesis that the means for the different type of vehicles are the same. I.e. There is enough statistical evidence to suggest that the average response time are different for each vehicle types.

One-way ANOVA-Example

Example 4.5 Taillite Example:
Researchers want to test whether the average braking time of drivers follows different types of trucks equipped with center high-mounted stop lamp (CHMSL) are the same or not
Code: Example 4.5

Conclusion: Output shows that the p-value is 0.0109, this means we **reject** the null hypothesis that the means for the different type of vehicles are the same. I.e. There is enough statistical evidence to suggest that the average response time are different for each vehicle types.

One-way ANOVA-Example

Example 4.6 Compare PROC GLM procedure with PROC TTEST procedure. Take out observations from the dataset taillite which correspond to level 1 and 2 in vehtype.
Code: ▸ Example 4.6

Conclusion: Output shows p-value from GLM F-test is 0.5553. **Do not reject**. p-value from t-test is 0.5553. **Do not reject**.

One-way ANOVA-Example

Example 4.6 Compare PROC GLM procedure with PROC TTEST procedure. Take out observations from the dataset taillite which correspond to level 1 and 2 in vehtype.
Code: Example 4.6

Conclusion: Output shows p-value from GLM F-test is 0.5553. **Do not reject**. p-value from t-test is 0.5553. **Do not reject**.

Two-way ANOVA

- **Model**
  Suppose we have 2 main factors, row factor and column factor, for observation k with row factor i and column factor j:

  $$y_{ijk} = \mu + a_i + b_j + (ab_{ij}) + e_{ijk}$$

  where $e_{ijk} \sim_{i.i.d} N(0, \sigma^2)$, μ is common mean and a_i is row effect and b_j is column effect with a possible interaction term ab_{ij}.
- In two-way ANOVA, there exist two factors (main effects) and a possible interaction term in the model.
- A hypothesis test is associated with each term. The null hypothesis could be either that there is no main effect or no interaction effect.

Two-way ANOVA

- **Model**

  Suppose we have 2 main factors, row factor and column factor, for observation k with row factor i and column factor j:

$$y_{ijk} = \mu + a_i + b_j + (ab_{ij}) + e_{ijk}$$

  where $e_{ijk} \sim_{i.i.d} N(0, \sigma^2)$, μ is common mean and a_i is row effect and b_j is column effect with a possible interaction term ab_{ij}.
- In two-way ANOVA, there exist two factors (main effects) and a possible interaction term in the model.
- A hypothesis test is associated with each term. The null hypothesis could be either that there is no main effect or no interaction effect.
- The code is the same as for a one-way ANOVA; the only change necessary is in the MODEL statement.

Two-way ANOVA

- **Model**
  Suppose we have 2 main factors, row factor and column factor, for observation k with row factor i and column factor j:

$$y_{ijk} = \mu + a_i + b_j + (ab_{ij}) + e_{ijk}$$

  where $e_{ijk} \sim_{i.i.d} N(0, \sigma^2)$, μ is common mean and a_i is row effect and b_j is column effect with a possible interaction term ab_{ij}.
- In two-way ANOVA, there exist two factors (main effects) and a possible interaction term in the model.
- A hypothesis test is associated with each term. The null hypothesis could be either that there is no main effect or no interaction effect.
- The code is the same as for a one-way ANOVA; the only change necessary is in the MODEL statement.

Two-way ANOVA-Syntax

Syntax

```
MODEL y = a b a*b ;
MODEL y = a | b; /* Do all crossing of factors. */
```

Two-way ANOVA-Example

Example 4.7 These data were taken from data provided by Scott McClelland from a study comparing the Hybrid III dummy to humans. Dummy and human skulls were hit by side impacts and stiffness measurements were taken in two locations. The response variable is stiffness. Is the average stiffness of dummies and humans different? Is the average stiffness for type of impactor different?

Code: Example 4.7

Conclusion Interaction term (species*impactor) is not significant (p-value=0.4358). species is not significant (p-value=0.0723). Impactor is significant (p-value=0.0001). Therefore, the average stiffness measure for humans and dummies do not differ, but the average stiffness measure for types of impactor do differ.

Two-way ANOVA-Example

Example 4.7 These data were taken from data provided by Scott McClelland from a study comparing the Hybrid III dummy to humans. Dummy and human skulls were hit by side impacts and stiffness measurements were taken in two locations. The response variable is stiffness. Is the average stiffness of dummies and humans different? Is the average stiffness for type of impactor different?
Code: Example 4.7

Conclusion Interaction term (species*impactor) is not significant (p-value=0.4358). species is not significant (p-value=0.0723). Impactor is significant (p-value=0.0001). Therefore, the average stiffness measure for humans and dummies do not differ, but the average stiffness measure for types of impactor do differ.

Two-way ANOVA-Model Checking

In ANOVA, the error terms are assumed to be independent and normally distributed with mean 0 and constant variance.

Example 4.8 Model Checking

Code: Example 4.8