

## Arriving 'Normally' at t-test for a treatment combination

Naina Pandurangi, PharmaNet/i3, Mumbai, India

### ABSTRACT

In clinical trials involving more than one active treatment, often statistics in a treatment combination are analyzed. When the inferences (t-test p-value) of continuous data of a treatment combination are based on normality checks of the data, there is a slight probability of going wide of the mark if the uniqueness of raw data points is not maintained after combining the treatments or if one of the samples of interest has sparse data.

This paper discusses the two areas of caution and the resultant incorrect decisions/oversight that might occur:

- 1) Passing the input dataset for normality check: If not careful, one may make false normality decision.
- 2) Picking the t-test p-value based on equality of variances (standard F test) in case of lack of data: If not careful, one may overlook the t-test p-value.

**Note:** This paper only enforces the 'Look Before you leap' notion while conducting t-test depending upon the inference of normality of data using SAS<sup>®</sup>, however the theory also holds true if one meets any of the following two situations independently:

- 1) Checking for normality when a treatment combination is to be analyzed
- 2) t-test p-value to be conducted on lack of data

### INTRODUCTION

The two sample t-test is used to compare the means between samples and hence normality of data is the presumption that this test needs. Hence in order to derive the t-test results we first need to check the normality of the data and then conduct the t-test based on the inference from the normality check.

This paper illustrates a case where one of the two samples is a combination of two treatments and the second sample has scanty data.

To conduct the t-test in SAS<sup>®</sup> we would want to have the combination of two treatments as a separate treatment for convenience, along with the individual treatments in our dataset to be analyzed.

The example below discusses the impact of using this resultant dataset for the normality check and the misleading inference coupled with the chances of missing out the t-test p-value due to insufficient data in one of the two samples.

### EXAMPLE ILLUSTRATING THE CASE

Consider a case where we have a two sample t-test that requires the data to be normal. We need the t-test p-value of a treatment combination v/s another treatment, after verifying the normality of the data.

Consider the viral load results data of subjects with their respective treatments as shown in the following dataset called Data RAW, where TRT and RES are the treatment and result variables, respectively.

#### Data RAW:

	SUBJECT	TRT	RES
1	1	A	2.3
2	2	A	3.1
3	3	A	2.9
4	4	A	4.6
5	5	A	0.7
6	6	A	1.3
7	7	A	3.1
8	8	A	1.9
9	9	A	3.1
10	10	A	4.5
11	14	B	1.1
12	15	B	2.2
13	16	B	0.9
14	17	B	0.8
15	18	B	0.4
16	19	C	0.9
17	20	C	.

When we prepare the resultant visual specification for the Data RAW, the desired output looks like what is presented below.

**Visual specification of the desired output:**

Statistic	TRT A	TRT B	TRT AB	TRT C
N	XX	XX	XX	XX
Mean	XX.X	XX.X	XX.X	XX.X
SD	XX.XX	XX.XX	XX.XX	XX.XX
Median	XX.X	XX.X	XX.X	XX.X
Min, Max	XX.X, XX.X	XX.X, XX.X	XX.X, XX.X	XX.X, XX.X
t-test p-value: AB v/s C			0.xxx	

Here 'AB' refers to the treatment combination of treatments A and B.

**TYPICAL ANALYSIS**

Typical analysis involves creation of the appropriate dataset (step1), testing the normality of the data (step 2), and finally, having checked that the data is normal, conducting the two sample t test and providing inference based on p values (step 3). Each of these steps is detailed below for your quick reference.

**STEP 1:**

The common and SAS® friendly approach is to first shape the input data going into the procedures as follows:

```
data trt_comb;
  set raw;
  output;
  if trt in ('A' 'B') then do;
    trt='AB';
    output;
  end;
run;
```

**STEP 2: NORMALITY**

After getting the descriptive statistics, we proceed towards the Normality check:

```
ods output testsfornormality=normal_comb;
proc univariate data=trt_comb normal;
  var res;
run;
```

**OUTPUT from Normality check on combination data:**

```

The UNIVARIATE Procedure
Variable: res

Moments
N          31      Sum weights          31
Mean      2.1516129  Sum Observations    66.7
Std Deviation 1.31754088  Variance            1.73591398
Skewness   0.46644875  Kurtosis            -0.8462594
Uncorrected SS 195.59  Corrected SS        52.0774194
Coeff Variation 61.2350335  Std Error Mean      0.23663733

Basic Statistical Measures

Location          Variability
Mean      2.151613  Std Deviation      1.31754
Median    2.200000  Variance            1.73591
Mode      3.100000  Range              4.20000
                          Interquartile Range  2.20000

Tests for Location: Mu0=0

Test          -Statistic-      -----p Value-----
Student's t   t    9.092449  Pr > |t|    <.0001
Sign          M     15.5  Pr >= |M|   <.0001
Signed Rank   S     248  Pr >= |S|   <.0001

Tests for Normality

Test          --Statistic--      -----p Value-----
Shapiro-wilk  W     0.90736  Pr < W     0.0111
Kolmogorov-Smirnov  D     0.160333  Pr > D     0.0414
Cramer-von Mises  W-Sq  0.132567  Pr > W-Sq  0.0401
Anderson-Darling  A-Sq  0.940072  Pr > A-Sq  0.0165
    
```

As per the output ('Tests for Normality' section, the p-values are <0.05), the data does **not** seem to be normal. Usually either one of the 4 tests or a combination is of these is referred for the inference, here we see all the 4 p-values are less than 0.05, so we can consider the data to be **non** normal.

**GAFFE:**

If **TRT\_COMB** is the data we use to check normality we certainly were not Normal while doing so! And this is proved by the numbers in the output.

It is clearly visible in the descriptive statistics section (N=31) that each subject under treatments A or B is counted twice than once as output in the data. (N=31 because # (A) =10, # (B) =5, #(C) =2, # (.)=1, treatments A and B counted twice contributing 30 to N, and treatment C contributing 1 to N, totalling 31, not adding the missing value corresponding to RES variable in the treatment C).

**What should be the right approach to correct this duplication of count resulting in to incorrect inference?**

We propose the following approach to resolve this issue. As one would have observed during the previous step, the outputs were generated from the combination data. We propose to replicate the same steps on the raw data which would provide accurate N that is 16 in this case.

Now, let's see the output from the Normality check on the **RAW** data.

**OUTPUT from Normality check on RAW data:**

```

The UNIVARIATE Procedure
Variable: res

Moments
N              16      Sum weights          16
Mean           2.1125   Sum Observations  33.8
Std Deviation  1.33660016  Variance          1.7865
Skewness       0.54674126 Kurtosis          -0.7313227
Uncorrected SS 98.2    Corrected SS      26.7975
Coeff variation 63.2710137 Std Error Mean    0.33415004

Basic Statistical Measures

Location              Variability
Mean      2.112500    Std Deviation  1.33660
Median    2.050000    Variance      1.78650
Mode      3.100000    Range         4.20000
                          Interquartile Range  2.20000

Tests for Location: Mu0=0

Test              -Statistic-      -----p Value-----
Student's t      t      6.32201      Pr > |t|      <.0001
Sign             M      8              Pr >= |M|     <.0001
Signed Rank      S      68             Pr >= |S|     <.0001

Tests for Normality

Test              --Statistic--      -----p Value-----
Shapiro-wilk     W      0.912668      Pr < W      0.1285
Kolmogorov-Smirnov D      0.165868      Pr > D      >0.1500
Cramer-von Mises W-Sq  0.074414      Pr > W-Sq   0.2339
Anderson-Darling A-Sq  0.516016      Pr > A-Sq   0.1679
    
```

As we can see from the 'Tests for Normality' section, the data is Normal (p-value for all 4 tests is >0.05), thus indicating that the analysis based on combination data can be misleading as far as the test of normality is concerned.

**Rule1:** We need to make sure that Normality is checked on raw data only.

**STEP 3: T-TEST TWO SAMPLE**

Having conducted the normality check on the raw data, we propose to conduct two sample t-test to compare the means between treatment combination AB and treatment C.

The necessary SAS® codes may be written as follows;

```

ods output equality=equality ttests=ttests;
proc ttest data=trt_comb;
  class trt;
  var res;
  where trt in ('AB' 'C');
run;
    
```

**OUTPUT from TTEST procedure:**

The TTEST Procedure

Variable: RES

trt	N	Mean	Std Dev	Std Err	Minimum	Maximum
AB	15	2.1933	1.3424	0.3466	0.4000	4.6000
C	1	0.9000	.	.	0.9000	0.9000
Diff (1-2)		1.2933	1.3424	1.3864		

  

trt	Method	Mean	95% CL Mean	Std Dev	95% CL Std Dev
AB		2.1933	1.4499 2.9367	1.3424	0.9828 2.1171
C		0.9000	.	.	.
Diff (1-2)	Pooled	1.2933	-1.6803 4.2670	1.3424	0.9828 2.1171
Diff (1-2)	Satterthwaite	1.2933	.	.	.

  

Method	Variances	DF	t Value	Pr >  t
Pooled	Equal	14	0.93	0.3667
Satterthwaite	Unequal	0	.	.

Equality of Variances

Method	Num DF	Den DF	F Value	Pr > F
Folded F	14	0	.	.

Conventionally, we go for t-test based on the F test of equality of variances as following:

```
data ttest;
    merge equality ttests;
    if probf>=0.05 and variances='Equal' then tpval=put(probt, 5.3);
    else if probf<0.05 and variances='Unequal' then tpval=put(probt, 5.3);
run;
```

**Merged dataset:**

	Variable	Method	NumDF	DenDF	FValue	ProbF	Variances	tValue	DF	Probt	tpval
1	RES	Pooled	14	0	.	.	Equal	0.93	14	0.3667	
2	RES	Satterth	.	.	.	.	Unequal	.	0	.	

**GAFFE:**

Following this approach we don't get any t-test p-value (TPVAL is missing in the merged dataset above) as it is clear from the output that the F-test yields no value due to lack of data in the treatment group C i.e. PROBF is missing. In fact, we can not get this t test p-value using conventional approach as shown in the resultant table above.

**How do we get correct t-test p-value information despite the above challenge?**

To get the correct t-test p-value in such cases, we propose to use the following code. Please note that the t-test value is generated in the system, however, it was not captured in a right way in the situation where F-value is missing. The following code only suggests how to capture this t-test value and bridge the gap.

```
data ttest;
  merge equality ttests;
  if probf>=0.05 and variances='Equal' then tpval=put(probt, 5.3);
  else if probf<0.05 and variances='Unequal' then tpval=put(probt, 5.3);
  else if probf=. And probt=. then tpval=put(probt, 5.3);
run;
```

**Rule 2:** Modify SAS® code such that even if F-test p-value is missing, t-test value gets captured since system generates it by default.

**Note:** As seen in the output above, t-test yields the p-values for both Equal and Unequal Variance. However the two sample t-test basically assumes equality of variances. This is tested using F and if there is strong evidence of unequal variance (F p-value<0.05), we pick the T p-value for Unequal variance i.e. Satterthwaite or if F p-value>=0.05 then the T p-value for Equal variance i.e. Pooled is picked.

Now, in case of lack of data, F-test results are missing i.e. there is no evidence of unequal variance and hence we would see the t-test p-value for equal variance only i.e. the Pooled one (as also seen in the output above)

## CONCLUSION:

**Normality check:** Checking normality of a data is actually checking the distribution of the data points and whether they are symmetric about the mean. Hence this has to be checked on raw data **only** and not the data where we are manually introducing new data points.

**t-test p-value:** Ideally we pick the t-test p-value based on F-test p-value but in case F-test yields missing statistics we should not ignore the t-test p-value. We need to check the T-test data and pick the non missing T-test p-value correctly.

As mentioned earlier in the abstract section, the points discussed above can be applicable in independent scenarios as well. In any case, we need to carefully look at the data before passing them through the SAS<sup>®</sup> procedures and then look at the corresponding output before passing conclusions!

## ACKNOWLEDGMENTS

I thank all my colleagues who have helped me in understanding these statistical concepts and reviewed and provided useful comments on my paper. I am grateful to Dr. Prashant Kirkire, country manager, PharmaNet/i3, India, who has provided valuable suggestions and inputs to my paper and my colleagues Sandeep Sawant and Neha Mohan who have truly inspired me to make this happen.

## CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Name: Naina Pandurangi  
Enterprise: PharmaNet/i3  
Address: 7<sup>th</sup> floor, Corporate Center, opp. to VITS hotel, Andheri-Kurla road, Andheri (E) – 400059.  
City, State ZIP: Mumbai, MH, India  
Work Phone: +91-22-30554020  
E-mail: nainap9@gmail.com

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.