

## Choosing Wisely: Using the Appropriate Statistical Test for Trend in SAS

Christina Park, Jui-Ting Hsiung, Melissa Soohoo, and Elani Streja,  
University of California, Irvine

### ABSTRACT

Tests for trend are an informative and useful tool to examine whether means, medians or proportions of continuous or categorical variables increase or decrease across ordered groups. In clinical and epidemiological research, comparisons of baseline patient characteristics (e.g., demographic, clinical and laboratory data) across ordered levels of the categorized primary exposure are often examined with chi-square or analysis of variance (ANOVA) statistical tests. These latter tests identify the existence of differences in patient characteristics, yet provide little information on trends in the ordered groups. Trend tests provide additional insight into the pattern of the relationship between independent and dependent variables. Multiple methods are available in SAS to evaluate trends of continuous and categorical variables using PROC REG (simple linear regression) and PROC FREQ (Jonckheere-Terpstra, Cochran-Armitage and Cochran-Mantel-Haenszel tests) statements. However, choosing the appropriate statistical test can be a challenge. The choice of tests varies depending on the assumptions about the variable of interest including its type and distribution. Selecting an inappropriate test may lead to incorrect inferences about the trend of the variable across ordered exposure groups. This is important, especially when the results from trend tests may influence which variables are considered as covariates in models of adjustment. In this paper, we aim to (1) describe when to use specific statistical tests to evaluate trends in continuous or categorical variables across ordered groups, and (2) provide examples of SAS codes for trend tests and interpret the resulting output.

### INTRODUCTION

In epidemiological or clinical cohort research, descriptions of the study population under investigation are necessary to gain a better understanding of the hypothesis tested, results presented, and conclusions derived, as well as to make any inferences about potential biases in study results. Most research in this field present a description of patient characteristics in the first table of the manuscript, Table 1. This table provides information about the study population such as demographic data, comorbidities, and baseline laboratory measures. For many studies, this information is presented for the total population as a whole. Moreover, when the study examines differences in patient outcomes or clinical presentation according to a certain exposure, patient characteristics by exposure category are often presented as well. Statistical tests may be utilized to ascertain if patient characteristics differ across exposure groups.

When there are only two levels of the exposure category, identifying whether there are statistical differences between groups may be conducted with chi-square tests for categorical characteristics (e.g., presence of diseases), analysis of variance (t-test or ANOVA) for parametrically distributed continuous variables (e.g., age), or Mann-Whitney for non-parametric continuous variables (e.g., number of hospital visits in the past year). These statistical tests may also be relevant for nominal exposure categories such as insurance type. However, when examining characteristics across ordered exposure groups, these previously listed statistical tests would identify whether said characteristic in one of the exposure categories differed from the others, but these tests would not reveal which group was different or whether there was an increasing or decreasing trend across the ordered groups.

In this paper, we will present information on how to test for statistical trends across ordered groups of exposure category. We will show how to statistically address the question: Does the patient baseline characteristic increase or decrease across incrementally higher or lower levels of the primary exposure? We will also discuss how to select the appropriate statistical test depending on the type and distribution of variable examined.

An overview of the following statistical tests for trend and related SAS codes will be covered:

1. Linear Regression
2. Jonckheere-Terpstra Test
3. Cochran-Armitage Trend Test
4. Cochran-Mantel-Haenszel Test

More detailed explanations of these tests and SAS codes can be found in the SAS documentation, as well as books listed under the Recommended Reading section at the end of this paper.

Examples will be illustrated using data from the National Health and Nutrition Examination Survey (NHANES, 2009-2010), a study that assesses the health and nutritional status of the United States population. NHANES datasets are available for public use and can be downloaded from the Centers for Disease Control and Prevention (CDC) website (CDC/NCHS, 2009-2010). Although we will focus on trend tests in the context of clinical and epidemiological studies for this paper, trend tests can be also be used in other fields, such as economics, environmental, and public policy research.

## EVALUATING TYPE AND DISTRIBUTION OF THE DEPENDENT VARIABLE

When creating a descriptive table, patient characteristics are the dependent variables and ordinal exposure categories are the independent variables. It is important to note that patient characteristics may come in three types: (1) continuous—parametric, (2) continuous—non-parametric, or (3) categorical (binary and more than two levels). Before deciding on a test for trend, the type and distribution of the dependent variable should be examined and identified.

### CONTINUOUS VARIABLES

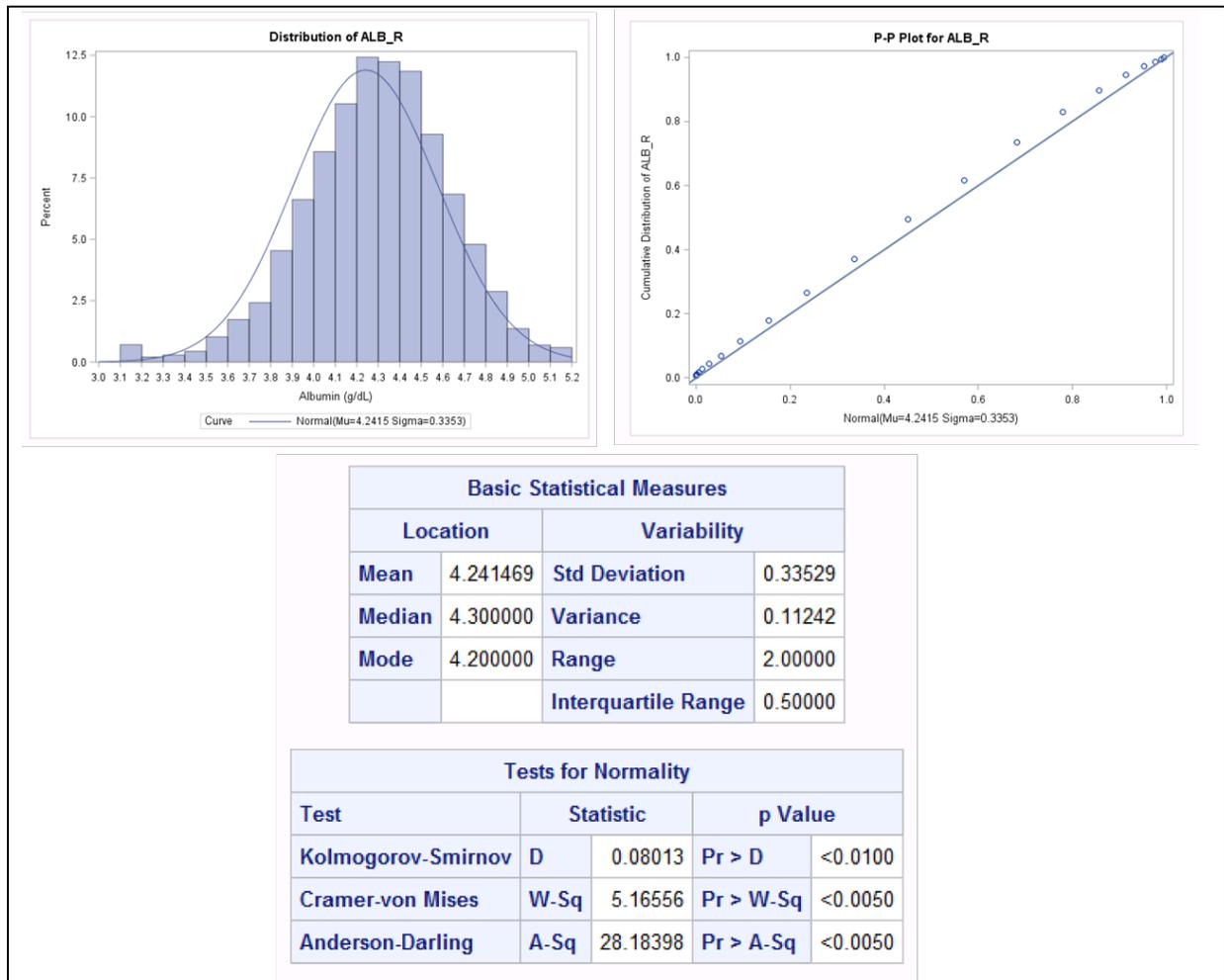
Visual and/or statistical tests can be used to assess whether the distribution of your continuous dependent variable is parametric or normal (bell-shaped curve). Examples of visual methods include creating histograms, boxplots, P-P (probability-probability) plots and Q-Q (quantile-quantile) plots. Additionally, statistical hypothesis tests such as Shapiro-Wilk, Kolmogorov-Smirnov and Anderson-Darling can be used to formally assess the normality of continuous data.

#### Parametric

Although SAS codes such as the UNIVARIATE procedure with the HISTOGRAM statement and the BOXPLOT procedure can separately create figures illustrating the distribution of your continuous variable of interest, the CAPABILITY procedure provides a comprehensive view of multiple evaluations to assess normality simultaneously. A more detailed explanation of PROC CAPABILITY can be found in SAS/QC® 9.3: “Syntax: CAPABILITY Procedure”. In the following example, we examine the distribution of serum albumin in our dataset using PROC CAPABILITY:

```
proc capability data=chol9c normaltest;
  var alb_r;
  label alb_r="Albumin (g/dL)";
  histogram/normal endpoints=3.1 to 5.1 by 0.1;
  ppplot alb_r;
run;
```

Results from the PROC CAPABILITY statement for albumin are displayed in Output 1:



### Output 1. Partial Output from PROC CAPABILITY Statement for Albumin

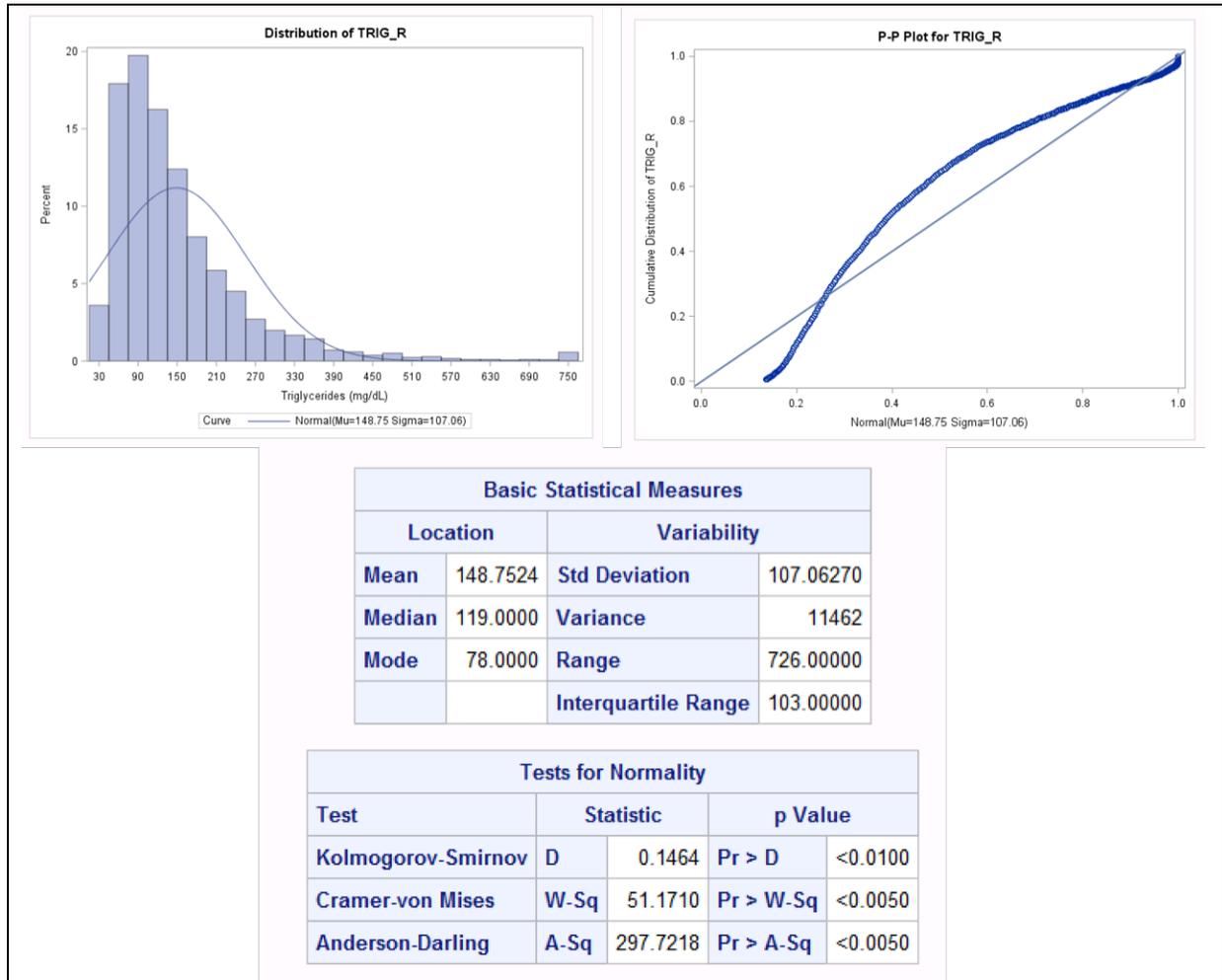
Output 1 shows a histogram (top left), P-P plot (top right) and tables of Basic Statistical Measures and Tests for Normality. Visually, both graphs show that the distribution of albumin is approximately normal. The histogram of albumin levels approximately follows a bell-shaped curve, while on the normal P-P plot, the data points lie roughly along a straight line. Furthermore, the mean and median albumin levels (4.2 and 4.3 g/dL, respectively) are close to each other. Statistical tests for normality (Kolmogorov-Smirnov, Cramer-von Mises and Anderson-Darling) suggest that the distribution of albumin is not normal (P-values <0.05), but this may be due to the fact that hypothesis tests are sensitive to sample size. Because of this sensitivity, we are prioritizing the evaluation of the illustrated distribution as to confirm that serum albumin is parametrically distributed.

### Non-Parametric

We will similarly use PROC CAPABILITY on all continuous variables of interest in our dataset. In this next example, we show the distribution of serum triglycerides in our data with PROC CAPABILITY:

```
proc capability data=chol9c normaltest;
  var trig_r;
  label trig_r="Triglycerides (mg/dL)";
  histogram/normal endpoints=31 to 757;
  ppplot trig_r;
run;
```

Results from the PROC CAPABILITY statement for triglycerides are displayed in Output 2:



### Output 2. Partial Output from PROC CAPABILITY Statement for Triglycerides

For the distribution of triglyceride levels, the histogram (top left), P-P plot (top right) and Basic Statistical Measures and Tests for Normality tables in Output 2 all show that the assumption of normality has likely been violated. We see that the histogram of triglyceride levels is right skewed and the data points do not fall along a straight line in the P-P plot. Additionally, the mean and median triglyceride levels (149 and 119 mg/dL, respectively) are not near each other. Results from all three tests for normality are significant (P-values <0.05), which suggest that the distribution of triglycerides is not normal. Given these results, we conclude that serum triglyceride levels follow a non-parametric distribution.

### CATEGORICAL VARIABLES

For data that are categorical, the FREQ procedure can be used to identify the number of levels that the dependent variable of interest contains. Here, we run a PROC FREQ on female and race categorical variables by including both variables in the TABLE statement:

```
proc freq data=chol9c;
  table female race;
run;
```

Results from the PROC FREQ statement for female and race are displayed in Output 3:

FEMALE	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	2901	48.76	2901	48.76
1	3048	51.24	5949	100.00

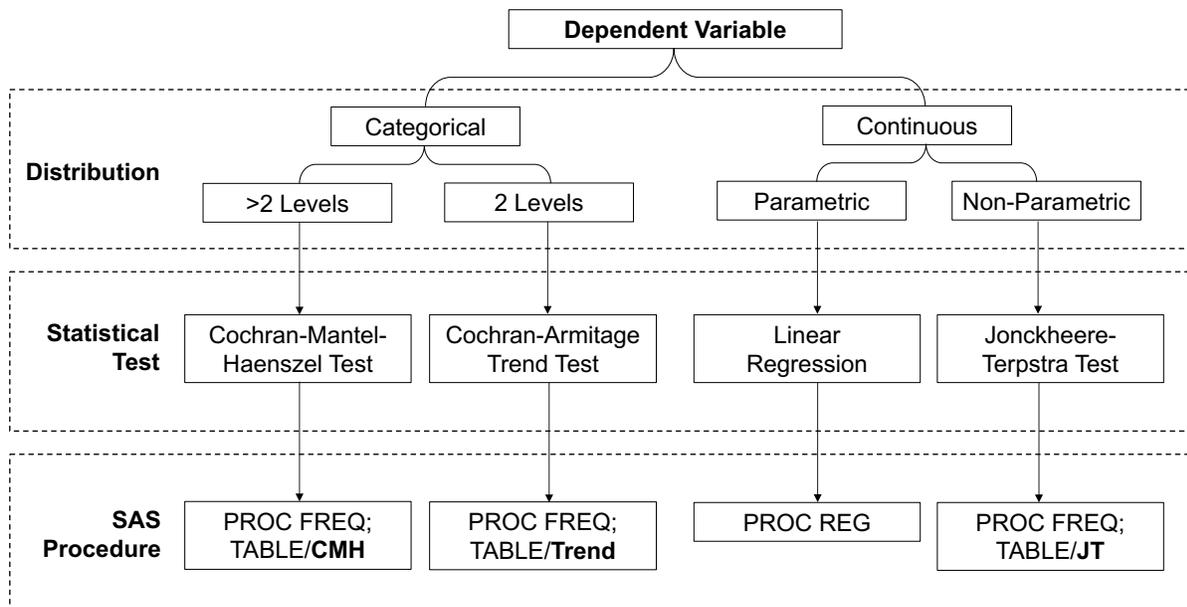
RACE	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Mexican American	1123	18.88	1123	18.88
Other Hispanic	610	10.25	1733	29.13
Non-Hispanic White	2859	48.06	4592	77.19
Non-Hispanic Black	1028	17.28	5620	94.47
Other Race-Including Multi-Racial	329	5.53	5949	100.00

**Output 3. Output from PROC FREQ Statement for Female and Race**

As we see in Output 3, female has two levels, while race has five levels. The statistical tests that we will use to evaluate test for trend for categorical variables in our dataset will differ for those that are binary (female) or more than 2 levels (race).

### CHOOSING THE APPROPRIATE TEST FOR TREND

Now that we have determined the type and distribution of each variable of interest, we will identify the appropriate statistical test for trend for the variable type. A flow chart representing the decision sequences for selecting the appropriate statistical test for trend and accompanying SAS function for a given dependent variable is presented in Figure 1:



**Figure 1. Flowchart for Selecting the Appropriate Statistical Test for Trend and Corresponding SAS Procedure Depending on the Type and Distribution of the Dependent Variable of Interest**

Examples of SAS procedures with discussion on the interpretation of the output are provided in the next section.

## **CONTINUOUS VARIABLES**

### **Parametric**

For continuous parametric variables, a simple linear regression can be used to assess if the distribution of the means of the dependent variable increase or decrease linearly across the ordered categorical exposure group. The main assumptions of this model include that the dependent variable and errors of the linear regression model are normally distributed.

### **Non-Parametric**

Alternatively, the Jonckheere-Terpstra test is a non-parametric test for continuous variables that can detect a trend between a non-normally distributed dependent variable and an ordered independent variable. The Jonckheere-Terpstra test is used to test the null hypothesis that there is no difference in the medians among the ordered groups, or the alternative hypothesis that there is an ordered difference in the medians of the dependent variable.

## **CATEGORICAL VARIABLES**

### **Binary**

The Cochran-Armitage trend test can be used to assess whether a trend is present between a binary (two levels, 0/1) dependent variable and an independent variable that has been categorized into more than two ordered categories. The Cochran-Armitage trend test is designed to test the null hypothesis that there is no ordered differences in the distribution (proportions) of the dependent variable across ordered categories.

### **More than Two Levels**

Alternatively, the Cochran-Mantel-Haenszel statistic is useful for evaluating the overall ordered differences in proportions of dependent categorical variables of more than two levels across ordered categories. The Cochran-Mantel-Haenszel test tests the null hypothesis that the proportions of the dependent variable are the same among the ordered exposure categories.

## **APPLICATION**

In this section, we will show how to evaluate the presence of a trend and estimate P-values for trend for each variable. In interpreting the test for trend output, a two-sided P-value  $<0.05$  will be considered statistically significant. Using a cohort of 5,949 adults with available high-density lipoprotein (HDL) measurements in NHANES 2009-2010, we will apply SAS procedures to describe patient characteristics and test for trends across four ordered categories of HDL level exposure:  $<40$ ,  $40-<50$ ,  $50-<60$  and  $\geq 60$  mg/dL. PROC MEANS and PROC FREQ were used to describe dependent variables with means  $\pm$  standard deviations (continuous—parametric), medians with interquartile ranges (continuous—non-parametric) and proportions (categorical) for the total population and within each HDL category, as appropriate for the variable type.

Summary estimates are provided in Table 1:

Variable*	Total Population	HDL (mg/dL)				P-value for trend†
		<40	40-<50	50-<60	≥60	
<b>N (%)</b>	5,949	1,278 (21)	1,611 (27)	1,345 (23)	1,715 (29)	
Female, %	51	31	43	57	69	<0.0001
Race/Ethnicity, %						<0.0001
Non-Hispanic White	48	50	46	47	49	
Non-Hispanic Black	17	13	17	19	20	
Mexican American	19	22	20	18	15	
Other Hispanic	10	10	11	10	10	
Other Race-Including Multi-Racial	6	5	5	6	6	
Albumin (g/dL)	4.24 ± 0.34	4.23 ± 0.32	4.25 ± 0.34	4.24 ± 0.35	4.24 ± 0.33	0.37
Triglycerides (mg/dL)	119 (79, 182)	196 (131, 297)	135 (93, 196)	107 (75, 154)	85 (61, 119)	<0.0001
*Data are presented as mean ± SD, median (interquartile range) or percentage, where appropriate. Percentages are rounded to the nearest whole number.						
†P-values for trend were calculated with the use of linear regression, Jonckheere-Terpstra, Cochran-Armitage or Cochran-Mantel-Haenszel tests, where appropriate.						

**Table 1. Baseline Characteristics of 5,949 NHANES Adults by HDL Levels**

### EXAMPLE 1. CONTINUOUS – PARAMETRIC

From Table 1, we note that the means and standard deviations for serum albumin are nearly identical for each category of HDL. We will now test whether according to statistical evaluation if there are no differences in the albumin means across the ordered HDL groups. Simple linear regression can be applied using the REG procedure to assess the linear association between the dependent variable serum albumin (alb\_r) and independent ordered categorical variable HDL (hdl\_cat). It is important to note that the independent ordered categorical variable (hdl\_cat) is on the right side of the equal sign in the MODEL statement, as we are modeling how the dependent variable (alb\_r) changes for each unit increase in the ordered categories of HDL:

```
proc reg data=chol9c;
  model alb_r=hdl_cat;
run;
```

Results from the PROC REG statement for albumin are displayed in Output 4:

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr >  t
Intercept	1	4.23242	0.01097	385.94	<.0001
HDL_CAT	1	0.00350	0.00389	0.90	0.3686

**Output 4. Partial Output from the PROC REG Statement for Albumin**

In Output 4, the Parameter Estimates table provides the degrees of freedom (DF), parameter estimate, standard error, t Value and P-value for the intercept and HDL category. The parameter estimate (or slope) of HDL category is 0.004. This estimate is interpreted as: for every unit increase in HDL category albumin increases by 0.004 g/dL. The P-value of the slope is 0.37 which is not statistically significant, indicating that our data do not differ from the null hypothesis for this test which is that the mean albumin levels are the same for each HDL category (as seen in Table 1) or that the slope of the line across HDL category is not significantly different from 0 or is flat. This can be further illustrated by examining the fitted plotted line with the residuals from this statistical test.

## EXAMPLE 2. CONTINUOUS – NON-PARAMETRIC

For triglycerides (trig\_r), we observed in Table 1 that the median levels of triglycerides with increasing HDL (from left to right) decrease across HDL ordered categories (hdl\_cat). Of note, the data may also be interpreted by reading across columns right to left or across decreasing HDL in Table 1, which shows that the median triglyceride levels increase with decreasing HDL. Nonetheless, for this non-parametric continuous variable, the PROC FREQ with JT option in the TABLE statement can be used to generate a P-value for test for trend. For more details about the PROC FREQ statement with JT option, see Base SAS 9.4®: “Jonckheere-Terpstra Test”. In the TABLE statement, we need to indicate our variables of interest as a cross tabulation. Following the “/” symbol, we specify the JT and NOPRINT (suppresses contingency table) options:

```
proc freq data=chol9c;  
  table trig_r*hdl_cat/jt noprint;  
run;
```

Results from the PROC FREQ statement for triglycerides are displayed in Output 5:

Jonckheere-Terpstra Test	
Statistic	6001624.500
Z	-37.9421
One-sided Pr < Z	<.0001
Two-sided Pr >  Z	<.0001

### Output 5. Output from PROC FREQ Statement with JT Option for Triglycerides

As shown in Output 5, the Jonckheere-Terpstra Test table gives us the test statistic, z statistic, one-sided and two-sided P-values for triglycerides. Here, the two-sided P-value is <0.0001, which tells us that there is a significant trend between triglycerides and HDL category. We added this P-value to our Table 1, and coupled with our observation of the median values, we determine that triglyceride levels decrease as HDL category increases.

## EXAMPLE 3. CATEGORICAL – BINARY

When we examine the female characteristic in Table 1, we see that the proportion of females across ordered HDL groups increases from 31 to 69% across increasing HDL category. Since female is a binary categorical variable, we use the PROC FREQ statement with the TREND option to examine the trend in proportion of females across HDL categories. This code will provide analysis according to the Cochran-Armitage Trend test. Additional details about the PROC FREQ statement with the TREND option can be found in Base SAS 9.4® Example 3.8: “Cochran-Armitage Trend Test”. In the TABLE statement, we need to include a cross tabulation of female and HDL category (hdl\_cat) and also specify the TREND and NOPRINT options:

```
proc freq data=chol9c;  
  table female*hdl_cat/trend noprint;  
run;
```

Results from the PROC FREQ statement with TREND option for female are displayed in Output 6:

Cochran-Armitage Trend Test	
Statistic (Z)	-21.9670
One-sided Pr < Z	<.0001
Two-sided Pr >  Z	<.0001

**Output 6. Output from PROC FREQ Statement with TREND Option for Female**

The Cochran-Armitage Trend Test table in Output 6 contains the z statistic, one-sided and two-sided P-values. As we can see from the output showing a two-sided P-value of <0.0001, we determine that there is a statistically significant relationship between females and HDL categories. By our observation of how the proportions change across HDL groups, we can state that the proportion of females significantly increases with increasing HDL category.

**EXAMPLE 4. CATEGORICAL – MORE THAN TWO LEVELS**

Lastly, for race, in Table 1 we observe that the proportion of Non-Hispanic Black patients increases, while the proportion of Mexican Americans decreases with increasing HDL category (hdl\_cat). Since race is a categorical variable with more than two levels, we will use the PROC FREQ statement with the CMH option. You can find more details about the PROC FREQ statement with CMH option in Base SAS 9.4® Example 3.7: “Cochran-Mantel-Haenszel Statistics”. Estimates from the Cochran-Mantel-Haenszel test can be obtained with the following code:

```
proc freq data=chol9c;
  table race*hdl_cat/cmh noprint;
run;
```

Results from the PROC FREQ statement with CMH option for race are displayed in Output 7:

Cochran-Mantel-Haenszel Statistics (Based on Table Scores)				
Statistic	Alternative Hypothesis	DF	Value	Prob
1	Nonzero Correlation	1	35.0227	<.0001
2	Row Mean Scores Differ	4	41.8054	<.0001
3	General Association	12	50.2400	<.0001

**Output 7. Output from PROC FREQ Statement with CMH Option for Race**

Output 7 shows the Cochran-Mantel-Haenszel Statistics table of three different alternative hypotheses: nonzero correlation, row mean scores differ and general association. Information on degrees of freedom (DF), test statistics (value) and two-sided P-values (prob) are also shown. Since we are interested in a potential trend between race across HDL categories, we read the P-value for “General Association”. The two-sided P-value of <0.0001 for General Association tells us that there is an overall statistically significant association between race and HDL categories. Some analysts may prefer to dummy code each level of this multilevel categorical variable to examine if there are significant trends for a specific response type. In that case, the Cochran-Armitage Trend test would be used as described above since the dummy coded variable would thereby be binary.

## CONSIDERATIONS

In statistical hypothesis testing, we typically turn to P-values to help us decide whether our results are statistically significant or not. However, we should approach our interpretation of P-values with caution, especially as they pertain to results of trend tests. P-values are particularly sensitive to sample size. In large studies such as our example study cohort, the P-values of the statistical tests for trend are usually significant (P-value <0.05). It is thus important to review all available information and consider whether the trends are also clinically meaningful.

Finally, the guidelines presented in this paper are not meant to be steadfast rules. The intent of this paper is to provide guidance on choosing the appropriate tests for trend and understanding the resulting output.

## CONCLUSION

Tests for trend offer another option to evaluate the presence of increasing or decreasing differences in the distribution of dependent variables across independent ordinal categories. In SAS, we can easily perform trend tests for continuous and categorical variables using simple procedures as described.

## REFERENCES

Centers for Disease Control and Prevention (CDC). National Center for Health Statistics (NCHS). National Health and Nutrition Examination Survey Data. Hyattsville, MD: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, 2009-2010. Available at <https://wwwn.cdc.gov/nchs/nhanes/ContinuousNhanes/Default.aspx?BeginYear=2009>. Accessed April 29, 2019.

SAS Institute Inc. 2011. SAS/QC® 9.3 User's Guide. Cary, NC: SAS Institute Inc.

SAS Institute Inc. 2016. Base SAS® 9.4 Procedures Guide: Statistical Procedures, Fifth Edition. Cary, NC: SAS Institute Inc.

## ACKNOWLEDGMENTS

The opinions expressed in this paper are those of the authors and do not represent the views of the University of California, Irvine.

## RECOMMENDED READING

- Rosner, B. 2011. *Fundamentals of Biostatistics*. 7th ed. Boston, MA: Brooks/Cole, Cengage Learning.
- Baldi, B. and Moore D.S. 2014. *The Practice of Statistics in the Life Sciences*. 3rd ed. New York, NY: W.H. Freeman and Company.

## CONTACT INFORMATION

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

Christina Park  
University of California, Irvine  
[parkc5@hs.uci.edu](mailto:parkc5@hs.uci.edu)  
[cpark5@uw.edu](mailto:cpark5@uw.edu)

Elani Streja  
University of California, Irvine  
[estreja@hs.uci.edu](mailto:estreja@hs.uci.edu)

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.