

# STAT 5110/6110: SAS Programming and Applications

## 1-C. Commonly-Used SAS Procedures

Peng Zeng

Department of Mathematics and Statistics

Auburn University

# Basic SAS Procedures

Some commonly-used SAS procedures

- `proc print`: show values of a data set
- `proc contents`: show information of a data set
- `proc means`: calculate descriptive statistics
- `proc freq`: construct frequency tables
- `proc sort`: sort observations

# PROC Step

The PROC (procedure) step analyzes data, produces output, or manages SAS files.

- input: usually a SAS data set
- output: a report or an updated SAS data set.

For example, you can use PROC steps to do the following:

- create a report that lists the data
- analyze data
- create a summary report
- produce plots and charts

# SAS Procedures

Each SAS procedure processes datasets in a different way.

```
proc whatever data = SAS-Dataset more-proc-options;  
  /* more SAS statements */  
run;
```

- The `data =` option specifies the name of the dataset you are working with.
- When `data =` option is omitted, SAS assumes that you are working with the most recent dataset.
- It is strongly recommended that you always include `data =` option, especially when you are working with several datasets simultaneously.

# Proc Print

`proc print` is used to show the contents of a dataset.

```
proc print data = SAS-Dataset;  
run;
```

You can also specify the names of variables that you want to display.

```
proc print data = SAS-Dataset;  
  var var1 var2;  
run;
```

The existing SAS output of `proc print` will not change automatically after you modify a SAS dataset!!

# Proc Contents

`proc contents` shows the summary information of a SAS dataset.

```
proc contents data = SAS-Dataset;  
run;
```

The output includes

- name of the data set
- number of observations
- number of variables
- for each variable
  - variable name
  - type (numeric or character)
  - length
  - format
  - informat

# Proc Means

`proc means` calculates descriptive statistics for numeric variables. The following code computes summary statistics for `var1` and `var2`, separately.

```
proc means data = SAS-Dataset;  
  var var1 var2;  
run;
```

By default, SAS outputs the number of observations, mean, standard deviation, minimum and maximum.

# Proc Freq

`proc freq` produces summary information for categorical variables. The following code constructs frequency tables for `var1` and `var2`, separately.

```
proc freq data = SAS-Dataset;  
  table var1 var2;  
run;
```

By default, SAS outputs the frequency/percentage, cumulative frequency/percentage of each level of the categorical variable.



## Proc Sort

`proc sort` sorts the observations in a dataset by some variables in either ascending (default) or descending order.

```
proc sort data = SAS-Dataset;  
  by var1 var2;  
run;
```

`descending` reverses the sort order for the variable that **immediately** follows in the statement so that observations are sorted from the largest value to the smallest value.

```
proc sort data = SAS-Dataset;  
  by descending var1 var2;  
  /* descending for var1, ascending for var2 */  
run;
```

- You need to use `proc print` to check the modified data sets.

## Example: Body Measures in NHANES

The data file (`bodymeasure.csv`) are extracted from NHANES and contains some demographic information and variables regarding body measures. The variables are

SEQN	age	edukid	eduadult
incomehouse	race	marital	gender
weight	recumbent	height	leg
armlength	armcircum	waist	triceps
subscapular			

Question:

- How many variables and how many observations in this data?
- Find the mean and standard deviation of height and weights.
- What is the percentage of females?
- What is the percentage of married subjects?

## In-Class Exercise

- Create a SAS data set from `blood-pressure.csv`.
- Use `proc print` to check the values in the SAS data set and compare them with the raw data
- Use `proc contents` to check the contents of the data set and identify the number of observations and variables and the list of variables
- Find the male patient with the largest systolic blood pressure.
- Find the female patient with the smallest diastolic blood pressure.
- What is the percent of the male patients?
- What is the mean and standard deviation of systolic blood pressure?