BIOS:5510:0001 Biostatistical Computing with R
Final Project

Points: 100 (30% of Final Grade)

Assigned: Monday, October 7, 2019
Due: Friday, October 18, 2019 by 5pm

Your mission is to develop an R version of the SAS TTEST procedure. Specifically, write an R function named `ttest` that has the definition given below. Statistics produced by the function should be computed according to the same formulas used by the SAS procedure (see SAS TTEST documentation for details). You may use any functionality provided by the default R system and may not use any external packages — no `library()` calls. You may use any functionality provided by the core R language, but not any external packages; i.e. no `library()` calls. Your function will be evaluated based on the following:

1. Functionality and robustness of code [25 pts].
2. Correctness of results [35 pts].
3. Formatting of output for clarity and readability [30 pts].
4. Efficiency of code [5 pts].
5. Consistency of code syntax with tidyverse style guide [5 pts].

Work on the project should be done independently and not discussed or shared with anyone other than the course instructor or TA.

Function Definition

Description

Performs t tests and computes confidence limits for samples from two independent groups.

Usage

```
ttest(x, class, alpha = 0.05, plots = FALSE)
```

Arguments

x a vector containing the numeric values for the two groups.

class a two-level character, factor, or numeric vector of the same length as **x** giving the corresponding group memberships.

alpha numeric (scalar); specifies that confidence intervals are to be $100(1-\alpha)\%$ confidence intervals. Must be between 0 and 1; the default value of 0.05 results in 95% confidence intervals.

plots logical; whether to produce distributional plots, including histograms, box-plots, interval, and Q-Q plots.

Example

```
ichomes$bsmtY <- as.numeric(ichomes$bsmt %in% c("1/2", "3/4", "Full"))
ttest(ichomes$sale.amount, ichomes$bsmtY, plots = TRUE)
```

The SAS System

The TTEST Procedure

Variable: sale_amount

bsmtY	N	Mean	Std Dev	Std Err	Minimum	Maximum
0	180	128472	76857.9	5728.6	38250.0	795000
1	597	195664	88761.9	3632.8	64000.0	815000
Diff (1-2)		-67192.1	86158.7	7326.3		

bsmtY	Method	Mean	95% CL Mean	
0		128472	117168	139776
1		195664	188529	202799
Diff (1-2)	**Pooled**	-67192.1	-81573.9	-52810.3
Diff (1-2)	**Satterthwaite**	-67192.1	-80535.4	-53848.7

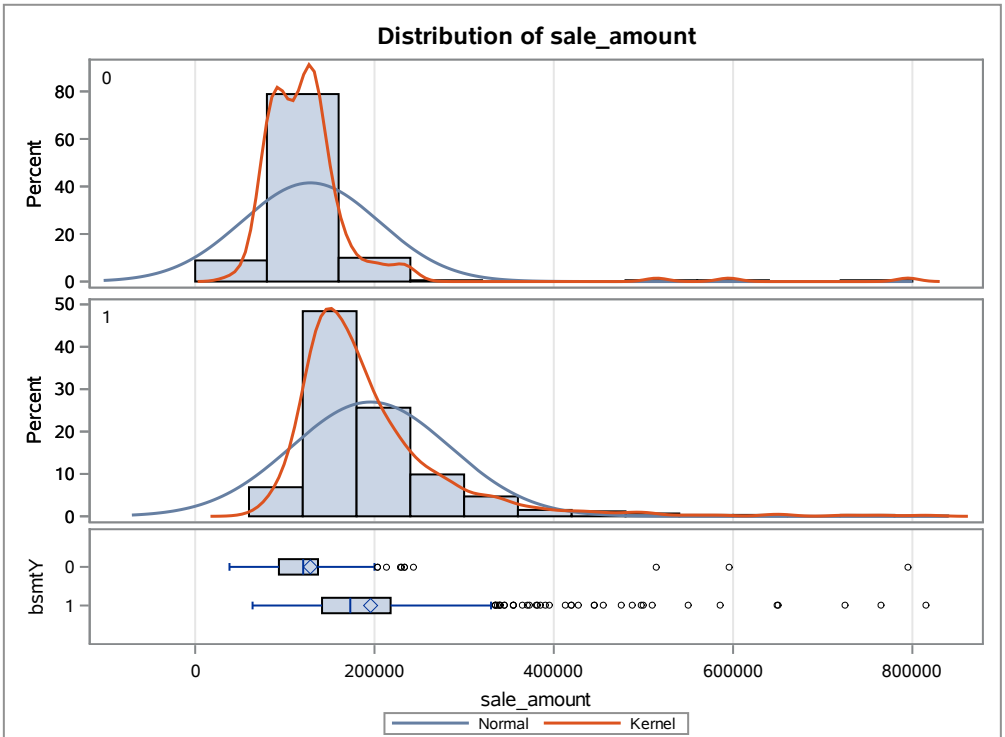
Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	775	-9.17	<.0001
Satterthwaite	Unequal	335.61	-9.91	<.0001

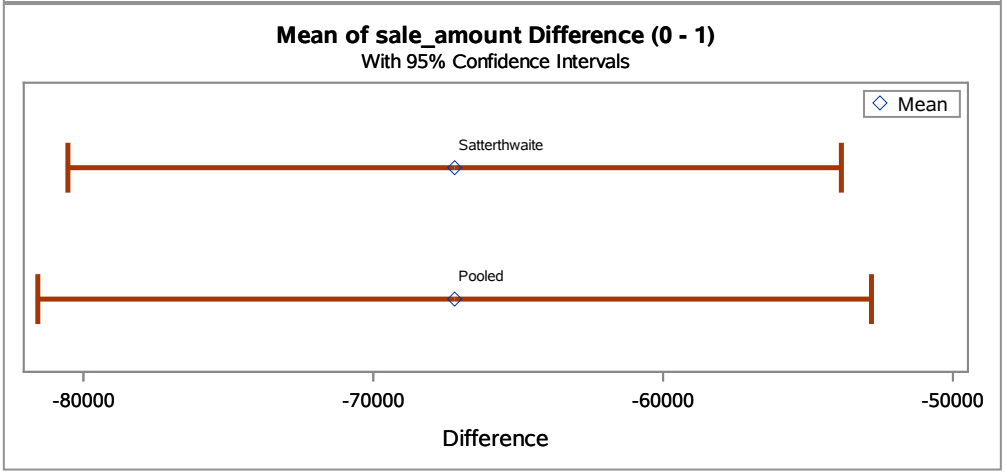
Equality of Variances				
Method	**Num DF**	**Den DF**	**F Value**	**Pr > F**
Folded F	596	179	1.33	0.0213

The SAS System

The TTEST Procedure

Variable: sale_amount





The SAS System

The TTEST Procedure

Variable: sale_amount