

STAT 430

Problem 6-2, p.197 (5th Edition)

```
=====
ssh xyz@glue.umd.edu , tap sas913 (old sas82), sas
https://www.statlab.umd.edu/sasdoc/sashtml/onldoc.htm
```

Data on reading speed by two methods: 'CODY' and 'SMITH'.  
Read the data, and perform t-test and Wilcoxon rank-sum test.

"LENGTH" is needed. It is a FORMAT statement. It tells SAS how to read the METHOD: 'CODY' or 'SMITH' . Note SMITH has length 5. We can replace 5 by 12 etc. but if use "LENGTH METHOD \$ 1;" then the name of the method is chopped to one letter: C and S. If we use "input METHOD \$;" then only get the CODY data. So, "LENGTH METHOD \$ 5;" is not an input statement. It is a format statement.

```
options ps=45 ls=70;
***Problem 6-2;
DATA READING;
    LENGTH METHOD $ 5;
    DO METHOD = 'CODY','SMITH';
        DO SUBJ = 1 TO 14;
            INPUT SPEED @;
            OUTPUT;
        END;
    END;
DATALINES;
500 450 505 404 555 567 588 577 566 644 511 522 543 578
355 388 440 600 510 501 502 489 499 489 515 520 520 480
;

PROC PRINT DATA=READING; <--To see the data use PROC PRINT:
RUN;
```

Obs	METHOD	SUBJ	SPEED
1	CODY	1	500
2	CODY	2	450
3	CODY	3	505
4	CODY	4	404
5	CODY	5	555
6	CODY	6	567
7	CODY	7	588
8	CODY	8	577
9	CODY	9	566
10	CODY	10	644
11	CODY	11	511
12	CODY	12	522
13	CODY	13	543
14	CODY	14	578
15	SMITH	1	355
16	SMITH	2	388
17	SMITH	3	440
18	SMITH	4	600
19	SMITH	5	510
20	SMITH	6	501
21	SMITH	7	502
22	SMITH	8	489
23	SMITH	9	499
24	SMITH	10	489
25	SMITH	11	515
26	SMITH	12	520
27	SMITH	13	520
28	SMITH	14	480

```

PROC TTEST DATA=READING;
  TITLE "T-test Comparing Reading Speed";
  CLASS METHOD;
  VAR SPEED;
RUN;

```

The TTEST Procedure

Statistics

Variable	METHOD	N	Lower CL		Upper CL	Lower CL
			Mean	Mean	Mean	Std Dev
SPEED	CODY	14	501.35	536.43	571.51	44.042
SPEED	SMITH	14	451.83	486.29	520.75	43.268
SPEED	Diff (1-2)		3.3568	50.143	96.929	47.424

Statistics

Variable	METHOD	Std Dev	Upper CL		Minimum	Maximum
			Std Dev	Std Err		
SPEED	CODY	60.751	97.873	16.236	404	644
SPEED	SMITH	59.684	96.154	15.951	355	600
SPEED	Diff (1-2)	60.22	82.528	22.761		

T-Tests

Variable	Method	Variances	DF	t Value	Pr >  t
SPEED	Pooled	Equal	26	2.20	0.0367
SPEED	Satterthwaite	Unequal	26	2.20	0.0367

Note: p-val for one-sided test  $Pr > t$  is  $0.0367/2 = 0.01835$   
 The alternative is  $H_1: CODY > SMITH$  since  $536.43 > 486.29!!!$

Equality of Variances

Variable	Method	Num DF	Den DF	F Value	Pr > F
SPEED	Folded F	13	13	1.04	0.9500

```

PROC NPAR1WAY DATA=READING WILCOXON;
  TITLE "Wilcoxon Rank Sum Test";
  CLASS METHOD;
  VAR SPEED;
  EXACT WILCOXON; <--If we don't include this, we get same results.
RUN;

```

Wilcoxon Rank Sum Test

The NPAR1WAY Procedure

Wilcoxon Scores (Rank Sums) for Variable SPEED  
Classified by Variable METHOD

METHOD	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
CODY	14	254.0	203.0	21.757927	18.142857
SMITH	14	152.0	203.0	21.757927	10.857143

Average scores were used for ties.

Wilcoxon Two-Sample Test

```

Statistic (S)                254.0000

Normal Approximation
Z                             2.3210
One-Sided Pr > Z              0.0101
Two-Sided Pr > |Z|           0.0203 <-- 2*0.0101

t Approximation
One-Sided Pr > Z              0.0140
Two-Sided Pr > |Z|           0.0281 <-- 2*0.0140

Exact Test
One-Sided Pr >= S             0.0090
Two-Sided Pr >= |S - Mean|    0.0179 <-- 2*0.0090

```

Z includes a continuity correction of 0.5.

Kruskal-Wallis Test

Chi-Square	5.4942
DF	1
Pr > Chi-Square	0.0191