

STAT 430, Fall 2010
Solution of Problem 6.4, p. 197 (5th Edition)

=====

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

- a. Matched pair comparisons within each group.
- b. Use of WHERE between groups.

```
OPTION PS=35 LS=70;
```

We generate 3 groups of data A,B,C by DO LOOP:

```
DATA QUEST6_4;  
DO GROUP = 'A', 'B', 'C';  
DO I= 1 TO 10;  
X= ROUND(RANNOR(135)*10+300+5*(GROUP= 'A')-7*(GROUP EQ 'C'));  
Y= ROUND(RANUNI(135)*100 +X);  
D = X-Y; <-----I added this to do t-test on D WITHIN each group.  
OUTPUT;          However, Prob. 6.4 does not ask for this.  
END;  
END;  
DROP I;  
RUN;
```

MOTE: (GROUP= 'A')=1 if GROUP=A, and 0 otherwise.

To see the data use PROC PRINT:

```
PROC PRINT DATA=QUEST6_4;  
BY GROUP;  
VAR X Y D;  
RUN;
```

----- GROUP=A -----

| Obs | X | Y | D |
|-----|-----|-----|-----|
| 1 | 307 | 405 | -98 |
| 2 | 281 | 366 | -85 |
| 3 | 308 | 380 | -72 |
| 4 | 304 | 376 | -72 |
| 5 | 321 | 341 | -20 |
| 6 | 312 | 326 | -14 |
| 7 | 300 | 391 | -91 |
| 8 | 298 | 376 | -78 |
| 9 | 320 | 386 | -66 |
| 10 | 300 | 390 | -90 |

----- GROUP=B -----

| Obs | X | Y | D |
|-----|-----|-----|-----|
| 11 | 288 | 359 | -71 |
| 12 | 290 | 303 | -13 |
| 13 | 301 | 375 | -74 |
| 14 | 291 | 351 | -60 |
| 15 | 307 | 363 | -56 |
| 16 | 297 | 327 | -30 |
| 17 | 273 | 350 | -77 |
| 18 | 296 | 377 | -81 |
| 19 | 316 | 376 | -60 |

----- GROUP=C -----

| Obs | X | Y | D |
|-----|-----|-----|-----|
| 21 | 286 | 293 | -7 |
| 22 | 276 | 297 | -21 |
| 23 | 293 | 387 | -94 |
| 24 | 290 | 291 | -1 |

| | | | |
|----|-----|-----|-----|
| 25 | 290 | 330 | -40 |
| 26 | 276 | 374 | -98 |
| 27 | 287 | 377 | -90 |
| 28 | 293 | 322 | -29 |
| 29 | 299 | 379 | -80 |
| 30 | 294 | 331 | -37 |

a. Matched pair comparisons within each group.

Since X,Y are dependent, to test $\text{mean}(x)=\text{mean}(y)$ WITHIN each group, we apply t-test to the difference data D:

```
PROC TTEST DATA=QUEST6_4;
  by group;
  VAR D;
  RUN;
```

----- GROUP=A -----

The TTEST Procedure

Statistics

| Variable | N | Lower CL Mean | Mean | Upper CL Mean | Lower CL Std Dev | Std Dev |
|----------|----|------------------|-------|------------------|---------------------|---------|
| D | 10 | -89.33 | -68.6 | -47.87 | 19.928 | 28.972 |

Statistics

| Variable | Upper CL Std Dev | Std Err | Minimum | Maximum |
|----------|---------------------|---------|---------|---------|
| D | 52.892 | 9.1618 | -98 | -14 |

T-Tests

| Variable | DF | t Value | Pr > t |
|----------|----|---------|---------|
| D | 9 | -7.49 | <.0001 |

----- GROUP=B -----

The TTEST Procedure

Statistics

| Variable | N | Lower CL Mean | Mean | Upper CL Mean | Lower CL Std Dev | Std Dev |
|----------|----|------------------|-------|------------------|---------------------|---------|
| D | 10 | -74.24 | -58.8 | -43.36 | 14.848 | 21.586 |

Statistics

| Variable | Upper CL Std Dev | Std Err | Minimum | Maximum |
|----------|---------------------|---------|---------|---------|
| D | 39.408 | 6.8261 | -81 | -13 |

T-Tests

| Variable | DF | t Value | Pr > t |
|----------|----|---------|---------|
| D | 9 | -8.61 | <.0001 |

----- GROUP=C -----

The TTEST Procedure

Statistics

| Variable | N | Lower CL Mean | Mean | Upper CL Mean | Lower CL Std Dev | Std Dev |
|----------|---|------------------|------|------------------|---------------------|---------|
|----------|---|------------------|------|------------------|---------------------|---------|

| | | | | | | |
|---|----|-------|-------|-----|--------|--------|
| D | 10 | -76.4 | -49.7 | -23 | 25.675 | 37.328 |
|---|----|-------|-------|-----|--------|--------|

Statistics

| Variable | Upper CL Std Dev | Std Err | Minimum | Maximum |
|----------|---------------------|---------|---------|---------|
| D | 68.146 | 11.804 | -98 | -1 |

T-Tests

| Variable | DF | t Value | Pr > t |
|----------|----|---------|---------|
| D | 9 | -4.21 | 0.0023 |

=====

b. Use of WHERE between groups.

Apparently, Problem 6.4 is to compare X in groups A and C, then also compare Y in groups A and C. For this we perform two t-tests: $H_0: \text{mean}_A(x) = \text{mean}_C(x)$, and $H_0: \text{mean}_A(y) = \text{mean}_C(y)$ BETWEEN groups, using WHERE as follows:

```
PROC TTEST DATA=QUEST6_4;
  TITLE "Comparing Groups A and C";
  WHERE GROUP IN ('A' 'C');
  CLASS GROUP;
  VAR X Y;
RUN;
```

Comparing Groups A and C

08:20 Friday, October 2, 2009

The TTEST Procedure

Statistics

| Variable | GROUP | N | Lower CL | | Upper CL | Lower CL |
|----------|------------|----|----------|-------|----------|----------|
| | | | Mean | Mean | Mean | Std Dev |
| X | A | 10 | 296.77 | 305.1 | 313.43 | 8.0047 |
| X | C | 10 | 283.03 | 288.4 | 293.77 | 5.1595 |
| X | Diff (1-2) | | 7.5014 | 16.7 | 25.899 | 7.3977 |
| Y | A | 10 | 356.59 | 373.7 | 390.81 | 16.454 |
| Y | C | 10 | 310.72 | 338.1 | 365.48 | 26.325 |
| Y | Diff (1-2) | | 5.6149 | 35.6 | 65.585 | 24.115 |

Statistics

| Variable | GROUP | Std Dev | Upper CL | | Std Err | Minimum | Maximum |
|----------|------------|---------|----------|---------|---------|---------|---------|
| | | | Std Dev | Std Dev | | | |
| X | A | 11.638 | 21.246 | 3.6801 | 281 | 321 | |
| X | C | 7.5011 | 13.694 | 2.3721 | 276 | 299 | |
| X | Diff (1-2) | 9.7903 | 14.478 | 4.3784 | | | |
| Y | A | 23.921 | 43.671 | 7.5646 | 326 | 405 | |
| Y | C | 38.272 | 69.87 | 12.103 | 291 | 387 | |
| Y | Diff (1-2) | 31.914 | 47.195 | 14.272 | | | |

The TTEST Procedure

T-Tests

| Variable | Method | Variances | DF | t Value | Pr > t |
|----------|---------------|-----------|------|---------|---------|
| X | Pooled | Equal | 18 | 3.81 | 0.0013 |
| X | Satterthwaite | Unequal | 15.4 | 3.81 | 0.0016 |
| Y | Pooled | Equal | 18 | 2.49 | 0.0226 |
| Y | Satterthwaite | Unequal | 15.1 | 2.49 | 0.0247 |

Equality of Variances

| Variable | Method | Num DF | Den DF | F Value | Pr > F |
|----------|----------|--------|--------|---------|--------|
| X | Folded F | 9 | 9 | 2.41 | 0.2068 |
| Y | Folded F | 9 | 9 | 2.56 | 0.1777 |