

SAS Lab 10 For Statistics 514

Topics:

Chapter 28. Nested Designs, Subsampling, and Partially Nested Designs

```
*Attendance 10,28.3,nested design, test scores;
DATA TESTSCORES;
  INPUT SCORES COURSE $ INSTRUCTOR $ TEST $;
DATALINES;
69 1 1 1
75 1 1 2
46 1 1 3
71 1 2 1
80 1 2 2
77 1 2 3
88 2 1 1
90 2 1 2
91 2 1 3
92 2 2 1
85 2 2 2
95 2 2 3
51 3 1 1
65 3 1 2
57 3 1 3
49 3 2 1
71 3 2 2
63 3 2 3
;
PROC GLM DATA=TESTSCORES;
  TITLE 'nested, test scores';
  CLASS course instructor;
  MODEL scores = course instructor(course);
RUN;
PROC NESTED DATA=testscores AOV;
  title 'nested, within operators, test scores';
  CLASS course instructor;
  BY course;
RUN;
QUIT;
```

```

*Attendance 10, 28.4, nested design, test scores;
DATA TESTSCORES;
  INPUT SCORES COURSE $ INSTRUCTOR $ TEST $;
DATALINES;
69 1 1 1
75 1 1 2
46 1 1 3
71 1 2 1
80 1 2 2
77 1 2 3
88 2 1 1
90 2 1 2
91 2 1 3
92 2 2 1
85 2 2 2
95 2 2 3
51 3 1 1
65 3 1 2
57 3 1 3
49 3 2 1
71 3 2 2
63 3 2 3
;
PROC GLM DATA=TESTSCORES noprint;
  TITLE 'nested, test scores';
  CLASS course instructor;
  MODEL scores = course instructor(course);
  OUTPUT OUT=testscoresout PREDICTED=PRED RESIDUALS=RESID;
RUN;
PROC PRINT DATA=testscoresout;
  TITLE 'residuals printout, nested, scores';
  VAR course instructor test PRED RESID;
RUN;
PROC GPLOT DATA=testscoresout;
  TITLE 'residual vs fitted, nested, scores';
  PLOT RESID*PRED;
RUN;
proc capability data=testscoresout noprint graphics;
  title 'normal probability plot for residuals';
  probplot resid;
run;
PROC CHART DATA=testscoresout;
  TITLE 'vbar, nested, test scores';
  VBAR RESID / GROUP = course;
RUN;
QUIT;

```

```

*Attendance 10, 28.5, nested design, test scores, effects;
DATA TESTSCORES;
  INPUT SCORES COURSE $ INSTRUCTOR $ TEST $;
DATALINES;
69  1  1  1
75  1  1  2
46  1  1  3
71  1  2  1
80  1  2  2
77  1  2  3
88  2  1  1
90  2  1  2
91  2  1  3
92  2  2  1
85  2  2  2
95  2  2  3
51  3  1  1
65  3  1  2
57  3  1  3
49  3  2  1
71  3  2  2
63 3 2 3
;
PROC GLM DATA=TESTSCORES noprint;
  TITLE 'nested, test scores';
  CLASS course instructor;
  MODEL scores = course instructor(course);
  OUTPUT OUT=testscoresout PREDICTED=PRED RESIDUALS=RESID;
RUN;
PROC MEANS DATA=testscores;
  title 'mean number of cases produced by course/instructor';
  VAR scores;
  CLASS course instructor;
  output out=testscoresmeans mean=mntests;
RUN;
PROC MEANS DATA=testscores;
  title 'mean score of tests for each course';
  VAR scores;
  class course;
RUN;
QUIT;

```

```

*Attendance 10,28.6,nested design, test scores;
*regression;
DATA TESTSCORES;
  INPUT SCORES COURSE $ INSTRUCTOR $ TEST $ X1 X2 X3 X4 X5 X6;
DATALINES;
69 1 1 1 1 0 1 0 0 0
75 1 1 2 1 0 1 0 0 0
46 1 1 3 1 0 1 0 0 0
71 1 2 1 1 0 -1 0 0 0
80 1 2 2 1 0 -1 0 0 0
88 2 1 1 0 1 0 1 0 0
90 2 1 2 0 1 0 1 0 0
91 2 1 3 0 1 0 1 0 0
92 2 2 1 0 1 0 -1 0 0
51 3 1 1 -1 -1 0 0 1 0
65 3 1 2 -1 -1 0 0 1 0
57 3 1 3 -1 -1 0 0 1 0
49 3 2 1 -1 -1 0 0 0 1
71 3 2 2 -1 -1 0 0 0 1
63 3 2 3 -1 -1 0 0 0 1
43 3 3 1 -1 -1 0 0 -1 -1
78 3 3 2 -1 -1 0 0 -1 -1
;
PROC REG DATA=testscores COVOUT OUTEST = COVEST;
  TITLE 'nested full model,missing obs';
  MODEL scores = X1 X2 X3 X4 X5 X6;
RUN;
PROC REG DATA=testscores;
  TITLE 'nested reduced (course) model,missing obs';
  MODEL scores = X3 X4 X5 X6;
RUN;
PROC REG DATA=testscores;
  TITLE 'nested reduced (instructor) model,missing obs';
  MODEL scores = X1 X2;
RUN;
QUIT;

```

```
*Attendance 10, 28.7, subsampling design, test scores;
DATA TESTSCORES;
  INPUT SCORES COURSE $ INSTRUCTOR $ TEST $;
DATALINES;
69 1 1 1
75 1 1 2
46 1 1 3
71 1 2 1
80 1 2 2
77 1 2 3
88 2 1 1
90 2 1 2
91 2 1 3
92 2 2 1
85 2 2 2
95 2 2 3
51 3 1 1
65 3 1 2
57 3 1 3
49 3 2 1
71 3 2 2
63 3 2 3
;
PROC GLM DATA=TESTSCORES;
  TITLE 'subsampling, test scores';
  CLASS course instructor;
  MODEL scores = course instructor(course);
RUN;
QUIT;
```

```

*Practice quiz 5, 28.4, bottle, nested, residuals;
DATA BOTTLING;
  INPUT CASES MACHINE $ OPERATOR $ DAY $;
  DATALINES;
65 1 1 1 1
58 1 1 1 2
63 1 1 1 3
57 1 1 1 4
66 1 1 1 5
68 1 2 1 1
62 1 2 2 2
75 1 2 2 3
64 1 2 2 4
70 1 2 2 5
56 1 3 1 1
65 1 3 2 2
58 1 3 3 3
70 1 3 3 4
64 1 3 3 5
45 1 4 1 1
56 1 4 2 2
54 1 4 3 3
48 1 4 4 4
60 1 4 4 5
74 2 1 1 1
81 2 1 1 2
76 2 1 1 3
80 2 1 1 4
68 2 1 1 5
69 2 2 1 1
76 2 2 2 2
80 2 2 2 3
78 2 2 2 4
73 2 2 2 5
52 2 3 1 1
56 2 3 2 2
62 2 3 3 3
58 2 3 3 4
51 2 3 3 5
73 2 4 1 1
78 2 4 2 2
83 2 4 3 3
75 2 4 4 4
76 2 4 4 5
69 3 1 1 1
83 3 1 1 2
74 3 1 1 3
78 3 1 1 4
80 3 1 1 5
63 3 2 1 1
70 3 2 2 2
72 3 2 2 3
68 3 2 2 4
75 3 2 2 5
81 3 3 1 1
72 3 3 2 2
73 3 3 3 3
76 3 3 3 4
70 3 3 3 5
67 3 4 1 1
79 3 4 2 2
73 3 4 3 3
77 3 4 4 4
71 3 4 4 5
;
PROC GLM DATA=BOTTLING noprint;
  TITLE '28-4(a), nested, bottle';
  CLASS MACHINE OPERATOR;
  MODEL CASES = MACHINE OPERATOR(MACHINE);
  OUTPUT OUT=BOTTLINGout PREDICTED=CASESPRED RESIDUALS=RESID;
RUN;
PROC PRINT DATA=BOTTLINGout;
  TITLE '28-4(a), residuals printout, nested, bottle';
  VAR MACHINE OPERATOR CASES CASESPRED RESID;
RUN;
PROC GPLOT DATA=BOTTLINGout;
  TITLE '28-4(a), residual vs fitted, nested, bottle';
  PLOT RESID*CASESPRED;
RUN;
proc capability data=bottlingout noprint graphics;
  title '28(a), normal probability plot for residuals';
  probplot resid;
run;
PROC CHART DATA=BOTTLINGout;
  TITLE '28-4(b), vbar, nested, bottle';
  VBAR RESID / GROUP = MACHINE;
RUN;
QUIT;

```

```

*Practice Quiz 5, 28.5 bottle, nested, inference;
DATA BOTTLING;
  INPUT CASES MACHINE $ OPERATOR $ DAY $;
  DATALINES;
65 1 1 1 1
58 1 1 1 2
63 1 1 1 3
57 1 1 1 4
66 1 1 1 5
68 1 2 1 1
62 1 2 2 2
75 1 2 3 3
64 1 2 4 4
70 1 2 5 5
56 1 3 1 1
65 1 3 2 2
58 1 3 3 3
70 1 3 4 4
64 1 3 5 5
45 1 4 1 1
56 1 4 2 2
54 1 4 3 3
48 1 4 4 4
60 1 4 5 5
74 2 1 1 1
81 2 1 2 2
76 2 1 3 3
80 2 1 4 4
68 2 1 5 5
69 2 2 1 1
76 2 2 2 2
80 2 2 3 3
78 2 2 4 4
73 2 2 5 5
52 2 3 1 1
56 2 3 2 2
62 2 3 3 3
58 2 3 4 4
51 2 3 5 5
73 2 4 1 1
78 2 4 2 2
83 2 4 3 3
75 2 4 4 4
76 2 4 5 5
69 3 1 1 1
83 3 1 2 2
74 3 1 3 3
78 3 1 4 4
80 3 1 5 5
63 3 2 1 1
70 3 2 2 2
72 3 2 3 3
68 3 2 4 4
75 3 2 5 5
81 3 3 1 1
72 3 3 2 2
73 3 3 3 3
76 3 3 4 4
70 3 3 5 5
67 3 4 1 1
79 3 4 2 2
73 3 4 3 3
77 3 4 4 4
71 3 4 5 5
;
PROC MEANS DATA=BOTTLING;
  title '28-5(b), mean number of cases produced by machine/operator';
  VAR CASES;
  CLASS MACHINE OPERATOR;
  output out=bottlingmeans mean=mncases;
RUN;
PROC MEANS DATA=BOTTLING;
  title '28-5(b), mean number of cases produced by machine';
  VAR CASES;
  class machine;
RUN;
PROC MEANS DATA=BOTTLING;
  title '28-5(b), mean number of cases produced by operator';
  VAR cases;
  class operator;
RUN;
PROC GPLOT DATA=bottlingmeans;
  TITLE '28-5(b) treatment means plot';
  PLOT mncases*operator=machine;
RUN;
PROC GPLOT DATA=bottlingmeans;
  TITLE '28-5(b) treatment means plot';
  PLOT mncases*machine=operator;
RUN;
PROC GLM DATA=BOTTLING;
  title '28-5(c), bottle, nested, tests';
  CLASS MACHINE OPERATOR;
  MODEL CASES = MACHINE OPERATOR(MACHINE);
  OUTPUT OUT=BOTTLINGout PREDICTED=CASESPRED RESIDUALS=RESID;
RUN;
PROC NESTED DATA=BOTTLING AOV;
  title '28-5(f), nested, within operators';
  CLASS MACHINE OPERATOR;
  BY MACHINE;
RUN;
QUIT;

```



```

*Practice Quiz 5, 28.6 bottle, nested, effects;
DATA BOTTLING;
  INPUT CASES MACHINE $ OPERATOR $ DAY $;
DATALINES;
65 1 1 1 1
58 1 1 1 2
63 1 1 1 3
57 1 1 1 4
66 1 1 1 5
68 1 2 1 1
62 1 2 2 2
75 1 2 2 3
64 1 2 2 4
70 1 2 2 5
56 1 3 1 1
65 1 3 2 2
58 1 3 3 3
70 1 3 4 4
64 1 3 5 5
45 1 4 1 1
56 1 4 2 2
54 1 4 3 3
48 1 4 4 4
60 1 4 5 5
74 2 1 1 1
81 2 1 2 2
76 2 1 3 3
80 2 1 4 4
68 2 1 5 5
69 2 2 1 1
76 2 2 2 2
80 2 2 3 3
78 2 2 4 4
73 2 2 5 5
52 2 3 1 1
56 2 3 2 2
62 2 3 3 3
58 2 3 4 4
51 2 3 5 5
73 2 4 1 1
78 2 4 2 2
83 2 4 3 3
75 2 4 4 4
76 2 4 5 5
69 3 1 1 1
83 3 1 2 2
74 3 1 3 3
78 3 1 4 4
80 3 1 5 5
63 3 2 1 1
70 3 2 2 2
72 3 2 3 3
68 3 2 4 4
75 3 2 5 5
81 3 3 1 1
72 3 3 2 2
73 3 3 3 3
76 3 3 4 4
70 3 3 5 5
67 3 4 1 1
79 3 4 2 2
73 3 4 3 3
77 3 4 4 4
71 3 4 5 5
;
PROC GLM DATA=BOTTLING;
  TITLE '28-6(a,b,c), nested, bottle';
  CLASS MACHINE OPERATOR;
  MODEL CASES = MACHINE OPERATOR(MACHINE);
  OUTPUT OUT=BOTTLING2 PREDICTED=CASESPRED RESIDUALS=RESID;
RUN;
PROC MEANS DATA=BOTTLING;
  title '28-6, mean number of cases produced by machine/operator';
  VAR CASES;
  CLASS MACHINE OPERATOR;
  output out=bottlingmeans mean=mncases;
RUN;
PROC MEANS DATA=BOTTLING;
  title '28-6(a), mean number of cases produced by machine';
  VAR CASES;
  class machine;
RUN;
QUIT;

```

```

*Practice Quiz 5, 28.7 bottle, nested, mixed;
DATA BOTTLING;
  INPUT CASES MACHINE $ OPERATOR $ DAY $;
  DATALINES;
65 1 1 1 1
58 1 1 1 2
63 1 1 1 3
57 1 1 1 4
66 1 1 1 5
68 1 2 1 1
62 1 2 2 2
75 1 2 2 3
64 1 2 2 4
70 1 2 2 5
56 1 3 1 1
65 1 3 2 2
58 1 3 3 3
70 1 3 3 4
64 1 3 3 5
45 1 4 1 1
56 1 4 2 2
54 1 4 3 3
48 1 4 4 4
60 1 4 4 5
74 2 1 1 1
81 2 1 1 2
76 2 1 1 3
80 2 1 1 4
68 2 1 1 5
69 2 2 1 1
76 2 2 2 2
80 2 2 2 3
78 2 2 2 4
73 2 2 2 5
52 2 3 1 1
56 2 3 2 2
62 2 3 3 3
58 2 3 3 4
51 2 3 3 5
73 2 4 1 1
78 2 4 2 2
83 2 4 3 3
75 2 4 4 4
76 2 4 4 5
69 3 1 1 1
83 3 1 1 2
74 3 1 1 3
78 3 1 1 4
80 3 1 1 5
63 3 2 1 1
70 3 2 2 2
72 3 2 2 3
68 3 2 2 4
75 3 2 2 5
81 3 3 1 1
72 3 3 2 2
73 3 3 3 3
76 3 3 3 4
70 3 3 3 5
67 3 4 1 1
79 3 4 2 2
73 3 4 3 3
77 3 4 4 4
71 3 4 4 5
;
PROC GLM DATA=BOTTLING;
  TITLE '28-7(b,c,d,e,f), nested, mixed';
  CLASS MACHINE OPERATOR;
  MODEL CASES = MACHINE OPERATOR(MACHINE);
RUN;
proc anova data=bottling;
  title '28-7(g) modified levene test, variance constant?';
  class operator;
  model cases = operator;
  means operator / hovtest = levene (type=abs);
run;
QUIT;

```

```

*Practice Quiz 5, 28.8 bottle, nested, random;
DATA BOTTLING;
  INPUT CASES MACHINE $ OPERATOR $ DAY $;
  DATALINES;
65 1 1 1
58 1 1 2
63 1 1 3
57 1 1 4
66 1 1 5
68 1 2 1
62 1 2 2
75 1 2 3
64 1 2 4
70 1 2 5
56 1 3 1
65 1 3 2
58 1 3 3
70 1 3 4
64 1 3 5
45 1 4 1
56 1 4 2
54 1 4 3
48 1 4 4
60 1 4 5
74 2 1 1
81 2 1 2
76 2 1 3
80 2 1 4
68 2 1 5
69 2 2 1
76 2 2 2
80 2 2 3
78 2 2 4
73 2 2 5
52 2 3 1
56 2 3 2
62 2 3 3
58 2 3 4
51 2 3 5
73 2 4 1
78 2 4 2
83 2 4 3
75 2 4 4
76 2 4 5
69 3 1 1
83 3 1 2
74 3 1 3
78 3 1 4
80 3 1 5
63 3 2 1
70 3 2 2
72 3 2 3
68 3 2 4
75 3 2 5
81 3 3 1
72 3 3 2
73 3 3 3
76 3 3 4
70 3 3 5
67 3 4 1
79 3 4 2
73 3 4 3
77 3 4 4
71 3 4 5
;
PROC GLM DATA=BOTTLING;
  TITLE '28-8(b,c,d,e), nested, random';
  CLASS MACHINE OPERATOR;
  MODEL CASES = MACHINE OPERATOR(MACHINE);
RUN;
QUIT;

```

```

*Practice Quiz 5, 28.14 control, nested, regression;
DATA INTERNAL;
  INPUT DATA REGION $ TEAM $ MONTH $ X1 X2 X3 X4 X5 X6;
DATALINES;
151.6 1 1 1 1 1 0 1 0 0 0
141.2 1 1 2 1 0 1 0 0 0
149.4 1 1 3 1 0 1 0 0 0
143.2 1 2 1 1 0 0 1 0 0
139.4 1 2 2 1 0 0 1 0 0
131.4 1 3 1 1 0 -1 -1 0 0
136.0 1 3 2 1 0 -1 -1 0 0
163.8 2 1 1 0 1 0 0 1 0
154.2 2 1 2 0 1 0 0 1 0
151.6 2 2 1 0 1 0 0 -1 0
157.0 3 1 1 -1 -1 0 0 0 1
147.2 3 1 2 -1 -1 0 0 0 1
160.0 3 2 1 -1 -1 0 0 0 -1
151.6 3 2 2 -1 -1 0 0 0 -1
;
PROC REG DATA=INTERNAL;
  TITLE '28.14(a), nested, internal control, regression, missing obs';
  MODEL DATA = X1 X2 X3 X4 X5 X6;
  OUTPUT OUT=INTERNALout PREDICTED=DATAPRED RESIDUAL=RESID;
RUN;
PROC PRINT DATA=INTERNALout;
  TITLE '28.14(b), residuals, internal control';
  VAR DATA REGION TEAM MONTH DATAPRED RESID;
RUN;
PROC GPLOT DATA=INTERNALout;
  TITLE '28-14(b), residual vs fitted, internal control';
  PLOT RESID*DATAPRED;
RUN;
proc capability data=internalout noprint graphics;
  title '28-14(b), normal probability plot for residuals';
  probplot resid;
run;
QUIT;

```

```

*Practice Quiz 5, 28.15 control, nested, regression, tests;
DATA INTERNAL;
  INPUT DATA REGION $ TEAM $ MONTH $ X1 X2 X3 X4 X5 X6;
DATALINES;
151.6 1 1 1 1 1 0 1 0 0 0
141.2 1 1 2 1 0 1 0 0 0
149.4 1 1 3 1 0 1 0 0 0
143.2 1 2 1 1 0 0 1 0 0
139.4 1 2 2 1 0 0 1 0 0
131.4 1 3 1 1 0 -1 -1 0 0
136.0 1 3 2 1 0 -1 -1 0 0
163.8 2 1 1 0 1 0 0 1 0
154.2 2 1 2 0 1 0 0 1 0
151.6 2 2 1 0 1 0 0 -1 0
157.0 3 1 1 -1 -1 0 0 0 1
147.2 3 1 2 -1 -1 0 0 0 1
160.0 3 2 1 -1 -1 0 0 0 -1
151.6 3 2 2 -1 -1 0 0 0 -1
;
PROC REG DATA=INTERNAL COVOUT OUTEST = COVEST;
  TITLE '28.15(a) nested full model, missing obs';
  MODEL DATA = X1 X2 X3 X4 X5 X6;
RUN;
PROC REG DATA=INTERNAL;
  TITLE '28.15(a) nested reduced (region) model, missing obs';
  MODEL DATA = X3 X4 X5 X6;
RUN;
PROC REG DATA=INTERNAL;
  TITLE '28.15(b) nested reduced (team) model, missing obs';
  MODEL DATA = X1 X2;
RUN;
PROC MEANS DATA=INTERNAL;
  TITLE '28.15(c), means, by region, for contrast';
  VAR DATA;
  CLASS REGION;
RUN;
PROC PRINT
  DATA = COVEST;
  title '28.15(c) variances for contrast';
RUN;
QUIT;

```

```

*Practice Quiz 5, 28.17 color, subsample, residuals;
data questionnaire;
  input response color lot week;
  DATALINES;
28 1 1 1 1
32 1 1 1 2
26 1 2 1 1
23 1 2 2 2
31 1 3 1 1
29 1 3 2 2
27 1 4 1 1
24 1 4 2 2
35 1 5 1 1
37 1 5 2 2
34 2 1 1 1
33 2 1 2 2
29 2 2 1 1
27 2 2 2 2
25 2 3 1 1
22 2 3 2 2
31 2 4 1 1
34 2 4 2 2
29 2 5 1 1
25 2 5 2 2
31 3 1 1 1
35 3 1 2 2
25 3 2 1 1
28 3 2 2 2
27 3 3 1 1
25 3 3 2 2
29 3 4 1 1
25 3 4 2 2
28 3 5 1 1
31 3 5 2 2
;
PROC GLM DATA=questionnaire noprint;
  TITLE '28-17(a), subsample, color';
  CLASS color lot;
  MODEL response = color lot(color);
  OUTPUT OUT=questionnaireout PREDICTED=PRED RESIDUALS=RESID;
RUN;
PROC PRINT DATA=questionnaireout;
  TITLE '28-17(a), residuals printout, subsample, color';
  VAR color lot week response PRED RESID;
RUN;
PROC GPLOT DATA=questionnaireout;
  TITLE '28-17(a), residual vs fitted, subsample, color';
  PLOT RESID*PRED;
RUN;
proc capability data=questionnaireout noprint graphics;
  title '28-17(a), normal probability plot for residuals';
  probplot resid;
run;
proc anova data=questionnaire;
  title '28-17(b) modified levene test, variance constant?';
  class lot;
  model response = lot;
  means lot / hovtest = levene (type=abs);
run;
QUIT;

```

```

*Practice Quiz 5, 28.18 color, subsample, inference;
data questionnaire;
  input response color lot week;
  DATALINES;
28 1 1 1
32 1 1 2
26 1 2 1
23 1 2 2
31 1 3 1
29 1 3 2
27 1 4 1
24 1 4 2
35 1 5 1
37 1 5 2
34 2 1 1
33 2 1 2
29 2 2 1
27 2 2 2
25 2 3 1
22 2 3 2
31 2 4 1
34 2 4 2
29 2 5 1
25 2 5 2
31 3 1 1
35 3 1 2
25 3 2 1
28 3 2 2
27 3 3 1
25 3 3 2
29 3 4 1
25 3 4 2
28 3 5 1
31 3 5 2
;
PROC GLM DATA=questionnaire;
  TITLE '28-18(a), subsample, color';
  CLASS color lot;
  MODEL response = color lot(color);
  OUTPUT OUT=questionnaireout PREDICTED=PRED RESIDUALS=RESID;
RUN;
PROC MEANS DATA=questionnaire;
  TITLE '28.18(d), means, by color';
  VAR response;
  CLASS color;
RUN;
QUIT;

```