

**SAS Lab 3 For Statistics 514**

**Topics:**

Chapter 19. Two-Factor Analysis of Variance-Equal Sample Sizes

Chapter 20. Analysis of Factor Effects in Two-Factor Studies-Equal Sample Sizes

```

*Attendance 3, deer mice, residual etc, prozac;
data deermice;
  do temp = 1 to 4;
    do noise = 1 to 3;
      do mouse = 1 to 2;
        input roc @@;
          output;
        end;
      end;
    end;
  end;
datalines;
10.3 7.2 9.1 5.4 6.1 2.1
1.8 9.8 12.1 4.2 5.1 6.2
1.2 8.1 6.5 4.1 1.2 2.1
12.4 15.1 16.1 17.2 18.1 19.1
;
proc print data=deermice;
  title 'print of data';
  var temp noise mouse roc;
run;
proc means data=deermice noprint;
  title 'treatment means,temp and noise';
  var roc;
  by temp noise;
  output out=studyout mean=mn;
run;
proc print data=studyout;
  var temp noise mn;
run;
proc gplot data=studyout;
  title 'treatment means plot, deer mice';
  plot mn*temp=noise;
run;
proc gplot data=studyout;
  title 'treatment means plot, deer mice';
  plot mn*noise=temp;
run;
proc means data=deermice noprint;
  title 'treatment means,temp';
  var roc;
  by temp;
  output out=tempout mean=mn;
run;
proc print data=tempout;
  var temp mn;
run;
proc gplot data=tempout;
  title 'treatment means plot, deer mice, temp';
  plot mn*temp;
run;
proc sort data=deermice;
  by noise;
run;
proc means data=deermice noprint;
  title 'treatment means,noise';
  var roc;
  by noise;
  output out=noiseout mean=mn;
run;
proc print data=noiseout;
  var noise mn;
run;
proc gplot data=noiseout;
  title 'treatment means plot, deer mice, noise';
  plot mn*noise;
run;
proc glm data=deermice noprint;
  class temp noise;
  model roc = temp noise temp*noise;
  output out=deermiceout p=pred r=resid;
run;
proc sort data=deermiceout;
  by temp;
run;
proc print data=deermiceout;
  title 'two factor ANOVA, disk';
  var temp noise roc pred resid;
run;
proc gplot data=deermiceout;
  title 'residual plot';
  plot resid*pred;
run;
proc capability data=deermiceout noprint graphics;
  title 'normal probability plot for residuals, deer mice';
  probplot resid;
run;
proc sort data=deermiceout;
  by temp noise;
run;
proc timeplot data=deermiceout;
  title 'residual sequence plots';
  id temp noise;
  plot resid;
  by temp noise;
run;
QUIT;

```

```
*Attendance 3, deer mice, residual etc, prozac;
data deermice;
  do temp = 1 to 4;
    do noise = 1 to 3;
      do mouse = 1 to 2;
        input roc @@;
          output;
        end;
      end;
    end;
  end;
datalines;
10.3 7.2 9.1 5.4 6.1 2.1
1.8 9.8 12.1 4.2 5.1 6.2
1.2 8.1 6.5 4.1 1.2 2.1
12.4 15.1 16.1 17.2 18.1 19.1
;
proc glm data=deermice;
  title 'two factor ANOVA, roc';
  class temp noise;
  model roc = temp noise temp*noise;
  output out=deermiceout p=pred r=resid;
run;
QUIT;
```

```

*Attendance 3, regression of ANOVA, gasoline;
data gas;
  do catalyst = 1 to 4;
    do blend = 1 to 3;
      do case = 1 to 2;
        input yield @@;
          if catalyst = 1 then x1 = 1;
          if catalyst = 2 then x1 = -1;
          if blend = 1 then x2 = 1;
          if blend = 2 then x2 = 0;
          if blend = 3 then x2 = -1;
          if blend = 1 then x3 = 0;
          if blend = 2 then x3 = 1;
          if blend = 3 then x3 = -1;
          x12 = x1*x2;
          x13 = x1*x3;
          output;
        end;
      end;
    end;
  end;
datalines;
84 82 89 87 94 94
84 85 89 88 94 93
;
proc glm data=gas;
  title 'regression of two factor ANOVA, gas';
  model yield = x1 x2 x3 x12 x13;
  output out=gasout p=pred r=resid;
run;
proc print data=gasout;
  title 'response, predicted and residuals';
  var x1 x2 x3 x12 x13 yield pred resid;
run;
proc glm data=gas;
  title 'two factor ANOVA, gas';
  class catalyst blend;
  model yield = catalyst blend catalyst*blend;
  output out=gasout p=pred r=resid;
run;
proc print data=gasout;
  title 'response, predicted and residuals, ANOVA';
  var catalyst blend yield pred resid;
run;
QUIT;

```

```
*Attendance 3, deer mice, pool interaction with error;
data deermice;
  do temp = 1 to 4;
    do noise = 1 to 3;
      do mouse = 1 to 2;
        input roc @@;
          output;
        end;
      end;
    end;
  end;
datalines;
10.3 7.2 9.1 5.4 6.1 2.1
1.8 9.8 12.1 4.2 5.1 6.2
1.2 8.1 6.5 4.1 1.2 2.1
12.4 15.1 16.1 17.2 18.1 19.1
;
proc glm data=deermice;
  title 'two factor ANOVA, roc';
  class temp noise;
  model roc = temp noise;
  output out=deermiceout p=pred r=resid;
run;
QUIT;
```

```
*Attendance 3, deer mice, residual etc;
data deermice;
    do temp = 1 to 4;
        do noise = 1 to 3;
            do mouse = 1 to 2;
                input roc @@;
                    output;
                end;
            end;
        end;
    end;
datalines;
10.3 7.2 9.1 5.4 6.1 2.1
1.8 9.8 12.1 4.2 5.1 6.2
1.2 8.1 6.5 4.1 1.2 2.1
12.4 15.1 16.1 17.2 18.1 19.1
;
proc glm data=deermice;
    title 'two factor ANOVA, roc';
    class temp noise;
    model roc = temp noise temp*noise;
    output out=deermiceout p=pred r=resid;
run;
QUIT;
```

```
*Attendance 3, deer mice, residual etc;
data deermice;
  do temp = 1 to 4;
    do noise = 1 to 3;
      do mouse = 1 to 2;
        input roc @@;
          output;
        end;
      end;
    end;
  end;
datalines;
10.3 7.2 9.1 5.4 6.1 2.1
1.8 9.8 12.1 4.2 5.1 6.2
1.2 8.1 6.5 4.1 1.2 2.1
12.4 15.1 16.1 17.2 18.1 19.1
;
proc sort data=deermice;
  by temp noise;
run;
proc means data=deermice noprint;
  title 'treatment means';
  var roc;
  by temp noise;
  output out=deerout mean=mn;
run;
proc print data=deerout;
  var temp noise mn;
run;
proc glm data=deermice;
  title 'two factor ANOVA, roc';
  class temp noise;
  model roc = temp noise temp*noise;
  output out=deermiceout p=pred r=resid;
run;
QUIT;
```

```
*Homework 2, 19.5, Treatment mean plots;
data study;
  do factora = 1 to 2;
    do factorb = 1 to 4;
      input mn @;
          lnmn = log(mn);
          output;
        end;
    end;
  datalines;
250 265 268 269
288 273 270 269
;
proc print data=study;
  title '19.5, data, study';
  var factora factorb mn lnmn;
run;
proc sort data=study;
  by factorb;
run;
proc means data=study;
  title '19.5(a), treatment means plot, study';
  var mn;
  by factorb;
  output out=studyout mean=mnfactor;
run;
proc gplot data=study;
  title '19.5(b), treatment means plot, study';
  plot mn*factorb=factora;
run;
proc gplot data=study;
  title '19.5(c), treatment means plot, log trans';
  plot lnmn*factorb=factora;
run;
quit;
```



```

*Homework 2, 19.16 residual etc, DISK DRIVE SERVICE;
DATA DISK;
    INPUT MINUTES TECH $ MAKE $ CASE @@;
DATALINES;
62 1 1 1
48 1 1 2
63 1 1 3
57 1 1 4
69 1 1 5
57 1 2 1
45 1 2 2
39 1 2 3
54 1 2 4
44 1 2 5
59 1 3 1
53 1 3 2
67 1 3 3
66 1 3 4
47 1 3 5
51 2 1 1
57 2 1 2
45 2 1 3
50 2 1 4
39 2 1 5
61 2 2 1
58 2 2 2
70 2 2 3
66 2 2 4
51 2 2 5
55 2 3 1
58 2 3 2
50 2 3 3
69 2 3 4
49 2 3 5
59 3 1 1
65 3 1 2
55 3 1 3
52 3 1 4
70 3 1 5
58 3 2 1
63 3 2 2
70 3 2 3
53 3 2 4
60 3 2 5
47 3 3 1
56 3 3 2
51 3 3 3
44 3 3 4
50 3 3 5
;
proc sort data=disk;
    by tech make;
run;
proc means data=disk;
    title '19.16(a), treatment means';
    var minutes;
    by tech make;
    output out=studyout mean=mn;
run;
proc print data=diskout;
    var tech make minutes mn;
run;
proc glm data=disk;
    class tech make;
    model minutes = tech make tech*make;
    output out=diskout p=pred r=resid;
run;
proc print data=diskout;
    title '19-16(a),(b), two factor ANOVA, disk';
    var tech make minutes pred resid;
run;
proc gplot data=diskout;
    title '19.16(c), residual plot';
    plot resid*pred;
run;
proc capability data=diskout noprint graphics;
    title '19.16(d) normal probability plot for residuals, disk';
    probplot resid;
run;
proc corr data=diskout;
    title '19.16(d) correlation of normal prob plot';
    var resid percentile;
run;
proc timeplot data=diskout;
    title '19.16(e), residual sequence plots';
    id tech make;
    plot resid;
    by tech make;
run;
QUIT;

```

```

*Homework 2, 19.17 TWO FACTOR ANOVA, DISK;
DATA DISK;
  INPUT MINUTES TECH $ MAKE $ CASE @@;
DATALINES;
62 1 1 1
48 1 1 2
63 1 1 3
57 1 1 4
69 1 1 5
57 1 2 1
45 1 2 2
39 1 2 3
54 1 2 4
44 1 2 5
59 1 3 1
53 1 3 2
67 1 3 3
66 1 3 4
47 1 3 5
51 2 1 1
57 2 1 2
45 2 1 3
50 2 1 4
39 2 1 5
61 2 2 1
58 2 2 2
70 2 2 3
66 2 2 4
51 2 2 5
55 2 3 1
58 2 3 2
50 2 3 3
69 2 3 4
49 2 3 5
59 3 1 1
65 3 1 2
55 3 1 3
52 3 1 4
70 3 1 5
58 3 2 1
63 3 2 2
70 3 2 3
53 3 2 4
60 3 2 5
47 3 3 1
56 3 3 2
51 3 3 3
44 3 3 4
50 3 3 5
;
PROC MEANS DATA=DISK noprint;
  VAR MINUTES;
  BY TECH MAKE;
  OUTPUT OUT=DISKout MEAN=MN;
RUN;
PROC GPLOT DATA=DISKout;
  TITLE '19.17(a) treatment mean plots, disk';
  PLOT MN*MAKE=tech;
RUN;
PROC ANOVA DATA=DISK;
  TITLE '19.17(b)(c)(d), two factor ANOVA, disk';
  CLASS TECH MAKE;
  MODEL MINUTES = TECH MAKE TECH*MAKE;
RUN;
QUIT;

```

```

*Homework 2, 19.19 TWO FACTOR ANOVA;
*KIDNEY FAILURE HOSPITALIZATION;
DATA KIDNEY;
  INPUT DAYS DURATION $ WEIGHTGAIN $ CASE @@;
  logdays = log10(days + 1);
DATALINES;
0 1 1 1
2 1 1 2
1 1 1 3
3 1 1 4
0 1 1 5
2 1 1 6
0 1 1 7
5 1 1 8
6 1 1 9
8 1 1 10
2 1 2 1
4 1 2 2
7 1 2 3
12 1 2 4
15 1 2 5
4 1 2 6
3 1 2 7
1 1 2 8
5 1 2 9
20 1 2 10
15 1 3 1
10 1 3 2
8 1 3 3
5 1 3 4
25 1 3 5
16 1 3 6
7 1 3 7
30 1 3 8
3 1 3 9
27 1 3 10
0 2 1 1
1 2 1 2
1 2 1 3
0 2 1 4
4 2 1 5
2 2 1 6
7 2 1 7
4 2 1 8
0 2 1 9
3 2 1 10
5 2 2 1
3 2 2 2
2 2 2 3
0 2 2 4
1 2 2 5
1 2 2 6
3 2 2 7
6 2 2 8
7 2 2 9
9 2 2 10
10 2 3 1
8 2 3 2
12 2 3 3
3 2 3 4
7 2 3 5
15 2 3 6
4 2 3 7
9 2 3 8
6 2 3 9
1 2 3 10
;
PROC MEANS DATA=KIDNEY noprint;
  VAR logDAYS;
  BY DURATION WEIGHTGAIN;
  OUTPUT OUT=KIDNEYout MEAN=MN;
RUN;
PROC GPLOT DATA=KIDNEYout;
  TITLE '19-19(a), treatment means plot, kidney';
  PLOT MN*WEIGHTGAIN=duration;
RUN;
PROC ANOVA DATA=KIDNEY;
  TITLE '19.19(b)(c)(d), TWO FACTOR ANOVA, kidney';
  CLASS DURATION WEIGHTGAIN;
  MODEL logDAYS = DURATION WEIGHTGAIN DURATION*WEIGHTGAIN;
RUN;
QUIT;

```

```

*Homework 2, 20.8 CI factors, disk;
DATA DISK;
  INPUT MINUTES TECH $ MAKE $ CASE @@;
  DATALINES;
62 1 1 1
48 1 1 2
63 1 1 3
57 1 1 4
69 1 1 5
57 1 2 1
45 1 2 2
39 1 2 3
54 1 2 4
44 1 2 5
59 1 3 1
53 1 3 2
67 1 3 3
66 1 3 4
47 1 3 5
51 2 1 1
57 2 1 2
45 2 1 3
50 2 1 4
39 2 1 5
61 2 2 1
58 2 2 2
70 2 2 3
66 2 2 4
51 2 2 5
55 2 3 1
58 2 3 2
50 2 3 3
69 2 3 4
49 2 3 5
59 3 1 1
65 3 1 2
55 3 1 3
52 3 1 4
70 3 1 5
58 3 2 1
63 3 2 2
70 3 2 3
53 3 2 4
60 3 2 5
47 3 3 1
56 3 3 2
51 3 3 3
44 3 3 4
50 3 3 5
;
PROC ANOVA DATA=DISK;
  TITLE '20.8, MSE for two factor ANOVA, disk';
  CLASS TECH MAKE;
  MODEL MINUTES = TECH MAKE TECH*MAKE;
RUN;
PROC MEANS DATA=DISK noprint;
  VAR MINUTES;
  BY TECH MAKE;
  OUTPUT OUT=DISKout MEAN=MNMIN;
RUN;
DATA DISKout;
  SET DISKout;
  RECPMIN = 1/MNMIN;
  LOGMIN = LOG10(MNMIN);
RUN;
proc print data=diskout;
  TITLE '20.8(e), original and transformed means, disk';
  var mnmin recpmin logmin;
run;
PROC GPLOT DATA=DISKout;
  TITLE '20.8(e), treatment means plot, disk, trans data';
  PLOT MNMIN*MAKE=TECH;
  PLOT RECPMIN*MAKE=TECH;
  PLOT LOGMIN*MAKE=TECH;
RUN;
QUIT;

```