

Analysis of Val Spicer's data involving %protein, temperature and %bound

Jason A. Osborne, Dept. of Statistics, NCSU

December, 2001

The design involved  $N = 36$  observations and two experimental factors: TEMP and PPROTEIN. There were 4 levels of the TEMP factor and 3 levels of the PPROTEIN factor for a total of  $4 \times 3 = 12$  treatment combinations. The response variable considered was PBOUND.

Of interest were effects of the two factors and the possible interaction between them.

Inspection of plots of the mean PBOUND by TEMP at each level of PPROTEIN (figure #1) indicates that PBOUND is nonlinear in TEMP. Mean PBOUND appears to be quadratic in TEMP. Additionally, the simple differences among levels of PPROTEIN within each level of TEMP are not the same across temperature, indicating a significant interaction between PPROTEIN and TEMP. Subsequent comparisons can be made by evaluating simple effects of PPROTEIN and/or TEMP while fixing the level of the other factor. Conservative (wider than necessary) 95% confidence intervals (using Scheffé's procedure) for all 66 pairwise differences among treatment means were obtained and are presented below, along with the ANOVA table. The reason they are conservative is that they allow for ALL of these comparisons, though interest can surely be confined to a smaller family than one including 66 comparisons. To see how to use this output, consider comparing the means for % protein values 2 ( $i = 1$ ) versus 6 ( $i = 3$ ) at temperature 75 ( $j = 4$ ) where  $y_{ijk}$  denotes the  $k^{th}$  rep at PPROTEIN level  $i$  and TEMP level  $j$ . The means are then  $\bar{y}_{14.} = 40.7$  and  $\bar{y}_{34.} = 75.3$ . They are numbered LSMEANS 4 and 12 respectively. The observed difference is  $-34.7$  with confidence limits  $-50.1$  and  $-19.2$ . Residual plots (figures #2,3) indicated that the data conform to the ANOVA assumptions of normality or homoscedasticity.

The SAS System 1  
The GLM Procedure

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	11	6757.666667	614.333333	41.73	<.0001
Error	24	353.333333	14.722222		
Corrected Total	35	7111.000000			

R-Square	Coeff Var	Root MSE	pbound Mean
0.950312	5.127334	3.836955	74.83333

Source	DF	Type III SS	Mean Square	F Value	Pr > F
pprotein	2	3167.166667	1583.583333	107.56	<.0001
temp	3	3168.555556	1056.185185	71.74	<.0001
pprotein*temp	6	421.944444	70.324074	4.78	0.0025

The GLM Procedure  
Least Squares Means

temp	pbound LSMEAN
45	76.1111111
55	86.0000000
65	77.2222222
75	60.0000000

pprotein	pbound LSMEAN
2	62.4166667
4	77.0000000
6	85.0833333

pprotein	temp	pbound LSMEAN	LSMEAN Number
2	45	66.3333333	1
2	55	77.6666667	2
2	65	65.0000000	3
2	75	40.6666667	4
4	45	79.6666667	5
4	55	87.3333333	6
4	65	77.0000000	7
4	75	64.0000000	8
6	45	82.3333333	9
6	55	93.0000000	10
6	65	89.6666667	11
6	75	75.3333333	12

Least Squares Means for Effect pprotein\*temp

i	j	Difference Between Means	Simultaneous 95% Confidence Limits for LSMean(i)-LSMean(j)	
1	2	-11.3333333	-26.801987	4.135320
1	3	1.3333333	-14.135320	16.801987
1	4	25.6666667	10.198013	41.135320
1	5	-13.3333333	-28.801987	2.135320
1	6	-21.0000000	-36.468654	-5.531346
1	7	-10.6666667	-26.135320	4.801987
1	8	2.3333333	-13.135320	17.801987
1	9	-16.0000000	-31.468654	-0.531346
1	10	-26.6666667	-42.135320	-11.198013
1	11	-23.3333333	-38.801987	-7.864680

1	12	-9.000000	-24.468654	6.468654
2	3	12.666667	-2.801987	28.135320
2	4	37.000000	21.531346	52.468654
2	5	-2.000000	-17.468654	13.468654
2	6	-9.666667	-25.135320	5.801987
2	7	0.666667	-14.801987	16.135320
2	8	13.666667	-1.801987	29.135320
2	9	-4.666667	-20.135320	10.801987
2	10	-15.333333	-30.801987	0.135320
2	11	-12.000000	-27.468654	3.468654
2	12	2.333333	-13.135320	17.801987
3	4	24.333333	8.864680	39.801987
3	5	-14.666667	-30.135320	0.801987
3	6	-22.333333	-37.801987	-6.864680
3	7	-12.000000	-27.468654	3.468654
3	8	1.000000	-14.468654	16.468654
3	9	-17.333333	-32.801987	-1.864680
3	10	-28.000000	-43.468654	-12.531346
3	11	-24.666667	-40.135320	-9.198013
3	12	-10.333333	-25.801987	5.135320
4	5	-39.000000	-54.468654	-23.531346
4	6	-46.666667	-62.135320	-31.198013
4	7	-36.333333	-51.801987	-20.864680
4	8	-23.333333	-38.801987	-7.864680
4	9	-41.666667	-57.135320	-26.198013
4	10	-52.333333	-67.801987	-36.864680
4	11	-49.000000	-64.468654	-33.531346
4	12	-34.666667	-50.135320	-19.198013
5	6	-7.666667	-23.135320	7.801987
5	7	2.666667	-12.801987	18.135320
5	8	15.666667	0.198013	31.135320
5	9	-2.666667	-18.135320	12.801987
5	10	-13.333333	-28.801987	2.135320
5	11	-10.000000	-25.468654	5.468654
5	12	4.333333	-11.135320	19.801987
6	7	10.333333	-5.135320	25.801987
6	8	23.333333	7.864680	38.801987
6	9	5.000000	-10.468654	20.468654
6	10	-5.666667	-21.135320	9.801987
6	11	-2.333333	-17.801987	13.135320
6	12	12.000000	-3.468654	27.468654
7	8	13.000000	-2.468654	28.468654
7	9	-5.333333	-20.801987	10.135320
7	10	-16.000000	-31.468654	-0.531346
7	11	-12.666667	-28.135320	2.801987
7	12	1.666667	-13.801987	17.135320
8	9	-18.333333	-33.801987	-2.864680
8	10	-29.000000	-44.468654	-13.531346
8	11	-25.666667	-41.135320	-10.198013
8	12	-11.333333	-26.801987	4.135320

9	10	-10.666667	-26.135320	4.801987
9	11	-7.333333	-22.801987	8.135320
9	12	7.000000	-8.468654	22.468654
10	11	3.333333	-12.135320	18.801987
10	12	17.666667	2.198013	33.135320
11	12	14.333333	-1.135320	29.801987

figure #1

Plot of p\*temp. Symbol is value of pprotein.

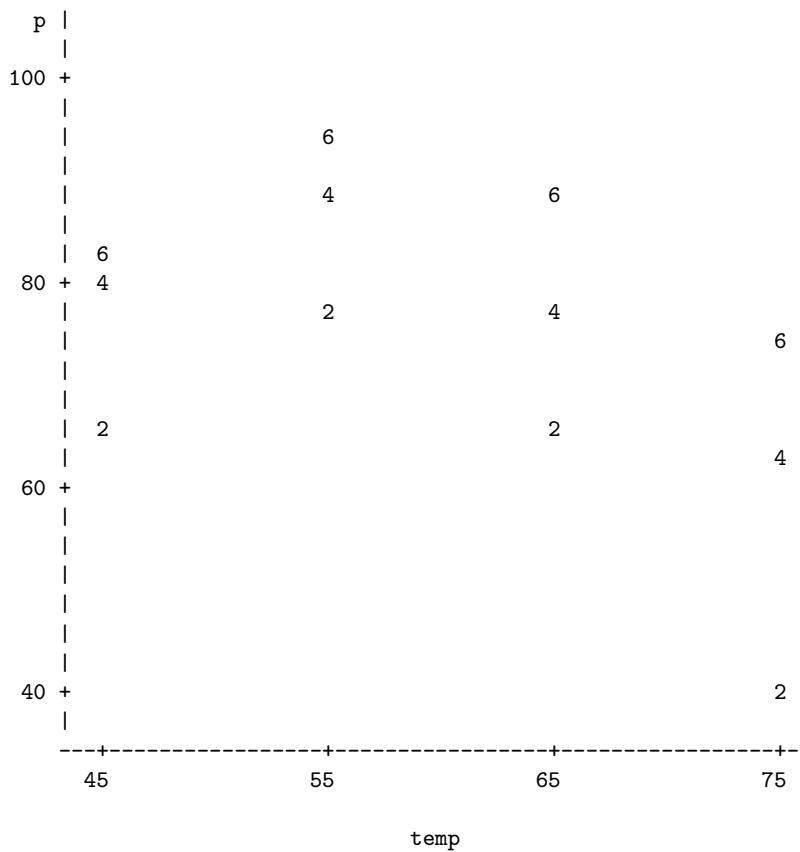


figure #2

Plot of  $r \cdot p$ . Legend: A = 1 obs, B = 2 obs, etc.

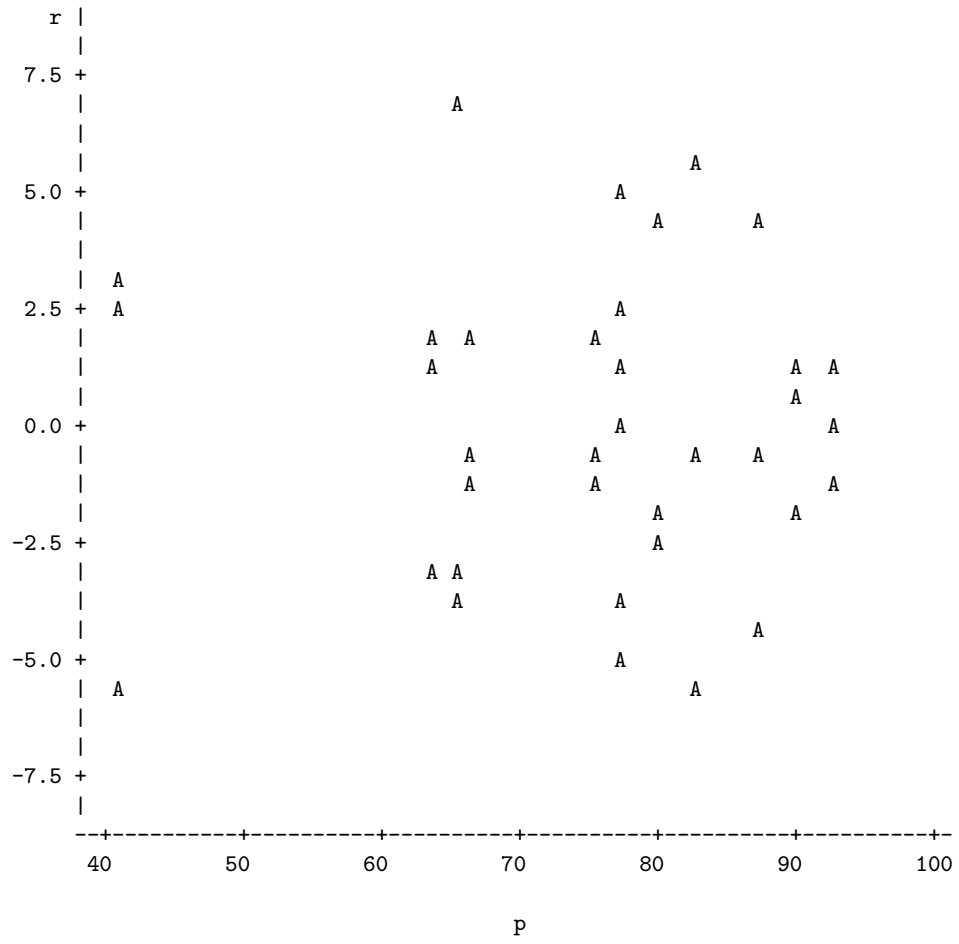


figure #3  
The UNIVARIATE Procedure  
Variable: r

