

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

Jason A. Osborne, Dept. of Statistics, NCSU

December, 2001

The design involved $N = 48$ observations and two experimental factors: TEMP and PPROTEIN. There were 4 levels of the TEMP factor and 4 levels of the PPROTEIN factor. The response variable considered was PFREE. For all 12 observations made in the experiment where the PPROTEIN factor was zero, PFREE=100 was observed without any variation. It is clear that PFREE=100 when PPROTEIN=0. All of the other 3×4 treatment combinations led to some experimental error. Because this violates the underlying assumption of heterogeneity of variance, the PPROTEIN=0 observations were excluded from the analysis which follows. Of interest were effects of the two factors and the possible interaction between them.

Inspection of plots of the mean PFREE by TEMP at each level of PPROTEIN (figure #1) indicates that PFREE is nonlinear in TEMP. Mean PFREE 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. The interactions appear to be qualitative instead of quantitative. That is, though the curves in the interaction plot do not intersect at all, they just have slightly different rates of change. It may be possible to obtain an additive model by simple logarithmic or square root transformations, but this approach was not considered here. Conservative (wider than necessary) 95% confidence intervals (using Tukey's procedure) for all 66 pairwise differences among treatment means were obtained and are presented below, along with the ANOVA table. Residual plots (figures #2,3) indicated that the data conform to the ANOVA assumptions of normality or homoscedasticity.

The SAS System
The GLM Procedure

1

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	11	6775.555556	615.959596	42.16	<.0001
Error	24	350.666667	14.611111		
Corrected Total	35	7126.222222			

Source	DF	Type III SS	Mean Square	F Value	Pr > F
pprotein	2	3158.722222	1579.361111	108.09	<.0001
temp	3	3193.555556	1064.518519	72.86	<.0001
pprotein*temp	6	423.277778	70.546296	4.83	0.0023

The GLM Procedure
Least Squares Means

pprotein	pfree LSMEAN	Standard Error
2	62.4166667	1.1034458
4	76.8333333	1.1034458
6	85.0833333	1.1034458

temp	pfree LSMEAN	Standard Error
45	76.1111111	1.2741494
55	86.0000000	1.2741494
65	77.1111111	1.2741494
75	59.8888889	1.2741494

pprotein	temp	pfree 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	76.6666667	7
4	75	63.6666667	8
6	45	82.3333333	9
6	55	93.0000000	10
6	65	89.6666667	11
6	75	75.3333333	12

Simultaneous 95% confidence intervals for differences
Adjustment for Multiple Comparisons: Tukey

i	j	Difference Between Means	Simultaneous 95% Confidence Limits for LSMean(i)-LSMean(j)	
1	2	-11.333333	-22.586567	-0.080100
1	3	1.333333	-9.919900	12.586567
1	4	25.666667	14.413433	36.919900
1	5	-13.333333	-24.586567	-2.080100
1	6	-21.000000	-32.253234	-9.746766
1	7	-10.333333	-21.586567	0.919900
1	8	2.666667	-8.586567	13.919900
1	9	-16.000000	-27.253234	-4.746766
1	10	-26.666667	-37.919900	-15.413433

1	11	-23.333333	-34.586567	-12.080100
1	12	-9.000000	-20.253234	2.253234
2	3	12.666667	1.413433	23.919900
2	4	37.000000	25.746766	48.253234
2	5	-2.000000	-13.253234	9.253234
2	6	-9.666667	-20.919900	1.586567
2	7	1.000000	-10.253234	12.253234
2	8	14.000000	2.746766	25.253234
2	9	-4.666667	-15.919900	6.586567
2	10	-15.333333	-26.586567	-4.080100
2	11	-12.000000	-23.253234	-0.746766
2	12	2.333333	-8.919900	13.586567
3	4	24.333333	13.080100	35.586567
3	5	-14.666667	-25.919900	-3.413433
3	6	-22.333333	-33.586567	-11.080100
3	7	-11.666667	-22.919900	-0.413433
3	8	1.333333	-9.919900	12.586567
3	9	-17.333333	-28.586567	-6.080100
3	10	-28.000000	-39.253234	-16.746766
3	11	-24.666667	-35.919900	-13.413433
3	12	-10.333333	-21.586567	0.919900
4	5	-39.000000	-50.253234	-27.746766
4	6	-46.666667	-57.919900	-35.413433
4	7	-36.000000	-47.253234	-24.746766
4	8	-23.000000	-34.253234	-11.746766
4	9	-41.666667	-52.919900	-30.413433
4	10	-52.333333	-63.586567	-41.080100
4	11	-49.000000	-60.253234	-37.746766
4	12	-34.666667	-45.919900	-23.413433
5	6	-7.666667	-18.919900	3.586567
5	7	3.000000	-8.253234	14.253234
5	8	16.000000	4.746766	27.253234
5	9	-2.666667	-13.919900	8.586567
5	10	-13.333333	-24.586567	-2.080100
5	11	-10.000000	-21.253234	1.253234
5	12	4.333333	-6.919900	15.586567
6	7	10.666667	-0.586567	21.919900
6	8	23.666667	12.413433	34.919900
6	9	5.000000	-6.253234	16.253234
6	10	-5.666667	-16.919900	5.586567
6	11	-2.333333	-13.586567	8.919900
6	12	12.000000	0.746766	23.253234
7	8	13.000000	1.746766	24.253234
7	9	-5.666667	-16.919900	5.586567
7	10	-16.333333	-27.586567	-5.080100
7	11	-13.000000	-24.253234	-1.746766
7	12	1.333333	-9.919900	12.586567
8	9	-18.666667	-29.919900	-7.413433
8	10	-29.333333	-40.586567	-18.080100
8	11	-26.000000	-37.253234	-14.746766

8	12	-11.666667	-22.919900	-0.413433
9	10	-10.666667	-21.919900	0.586567
9	11	-7.333333	-18.586567	3.919900
9	12	7.000000	-4.253234	18.253234
10	11	3.333333	-7.919900	14.586567
10	12	17.666667	6.413433	28.919900
11	12	14.333333	3.080100	25.586567

p-values for pairwise comparisons
Pr > |t| for H0: LSMean(i)=LSMean(j)

i/j	1	2	3	4	5	6
1		0.0473	1.0000	<.0001	0.0111	<.0001
2	0.0473		0.0182	<.0001	0.9999	0.1409
3	1.0000	0.0182		<.0001	0.0040	<.0001
4	<.0001	<.0001	<.0001		<.0001	<.0001
5	0.0111	0.9999	0.0040	<.0001		0.4093
6	<.0001	0.1409	<.0001	<.0001	0.4093	
7	0.0926	1.0000	0.0375	<.0001	0.9973	0.0744
8	0.9990	0.0067	1.0000	<.0001	0.0014	<.0001
9	0.0014	0.9279	0.0005	<.0001	0.9990	0.8921
10	<.0001	0.0024	<.0001	<.0001	0.0111	0.7948
11	<.0001	0.0296	<.0001	<.0001	0.1146	0.9997
12	0.2083	0.9997	0.0926	<.0001	0.9549	0.0296
i/j	7	8	9	10	11	12
1	0.0926	0.9990	0.0014	<.0001	<.0001	0.2083
2	1.0000	0.0067	0.9279	0.0024	0.0296	0.9997
3	0.0375	1.0000	0.0005	<.0001	<.0001	0.0926
4	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
5	0.9973	0.0014	0.9990	0.0111	0.1146	0.9549
6	0.0744	<.0001	0.8921	0.7948	0.9997	0.0296
7		0.0142	0.7948	0.0011	0.0142	1.0000
8	0.0142		0.0002	<.0001	<.0001	0.0375
9	0.7948	0.0002		0.0744	0.4719	0.5376
10	0.0011	<.0001	0.0744		0.9935	0.0004
11	0.0142	<.0001	0.4719	0.9935		0.0052
12	1.0000	0.0375	0.5376	0.0004	0.0052	

Figure #1
Mean by temperature and pprotein
Plot of p*temp. Symbol is value of pprotein.

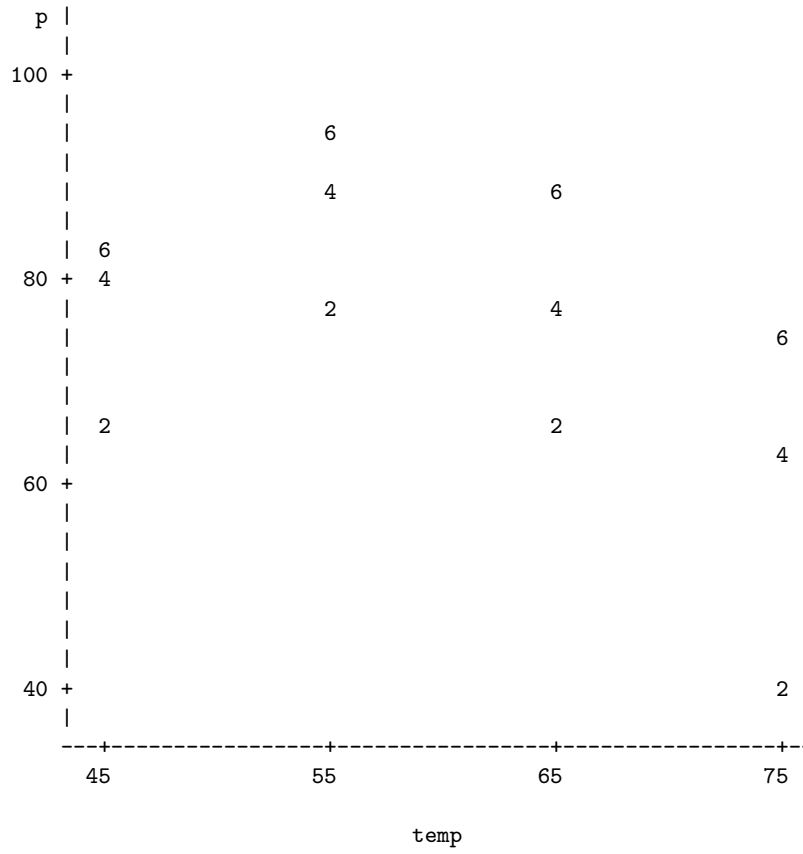


Figure #2
Residual by predicted

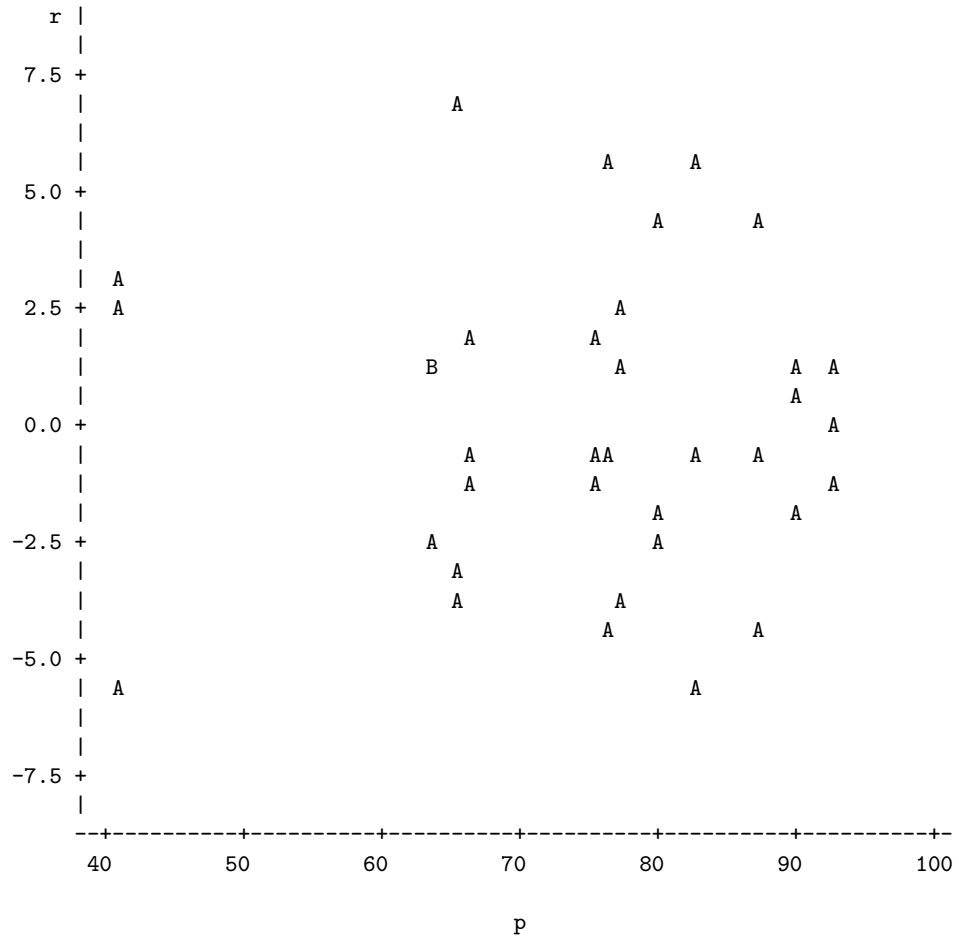


Figure #3
 Normal and stem and leaf plots of residuals

Stem Leaf	#	Boxplot
7 0	1	
6		
5 37	2	
4 37	2	
3 3	1	
2 33	2	
1 0333377	7	+-----+
0 03	2	+
-0 73333	5	*-----*
-1 77330	5	
-2 77	2	+-----+
-3 70	2	
-4 730	3	
-5 73	2	

-----+-----+-----+

The UNIVARIATE Procedure
 Variable: r

