

Final analysis of Keener/Bashor data on  
Bacterial counts for washer-treated chickens

Jason A. Osborne, Dept. of Statistics, NCSU, October 2002

Bacterial counts were taken on chickens sampled from 4 different plants, one of which was visited in two different seasons. Within each plant, birds were sampled at various sites along a sequences of washers, with labels assigned in the table below:

Label	1	2	3	4
Site	Before	After	After	After
	1 <sup>st</sup> washer	3 <sup>rd</sup> washer	microbial rinse	Chill Tank

Not every site was sampled from every plant. Birds were sampled after the microbial rinse only at plants C and D. So, the total number of treatment combinations for a given day is  $17 = 3 \times 3 + 2 \times 4$ :

Plant	site label			
	1	2	3	4
A	x	x		x
B(spring)	x	x		x
B2(fall)	x	x		x
C	x	x	x	x
D	x	x	x	x

$n = 10$  birds were sampled for each of 3 days at of these plant  $\times$  site combinations for a total of  $10 \times 3 \times 17 = 510$  birds.

The response variable is bacterial count ( $y$ ). Using general linear models to analyze the data is made difficult by the problem of modality at 0. There are a substantial number of birds (106 out of 510) for which no bacteria were detected ( $y = 0$ ). (See table below.) The distribution of bacterial counts is therefore non-normal and no power transformation will fix this problem.

Bacterial Count	Frequency				Total
	Site 1	Site 2	Site 3	Site 4	
$y = 0$	19	22	17	48	106
$y > 0$	131	128	43	102	404
Total	150	150	60	150	

In addition to demonstrating the modality at 0, the table also provides some evidence of a site effect on bacterial counts, in the sense that the proportion of 0's differs significantly by site, ( $\chi^2 = 22.95, p < 0.0001$  on  $3df$ ). The frequency of birds with 0 counts is higher at sites 3 and 4 than at sites 1 and 2.

Rather than pursuing more complicated *zero-inflated* models for the analysis, DAY can be taken as the experimental unit, as opposed to bird. So, for each day, the average of the bacterial counts for the 10 birds will be taken as the response in a simplified analysis.

The statistical model for the log-bacterial count  $y_{ijk}$  on day  $k$  at site  $j$  in plant  $i$  is then

$$Y_{ijk} = \mu + \alpha_i + \beta_{j(i)} + E_{ijk}$$

where  $\mu$  denotes an overall average log-bacterial count,  $\alpha_i$  denote fixed plant effects,  $\beta_j(i)$  denote site effects, nested in plant Plots of bacterial count versus site by day and residual plots which appear at the end of this writeup reveal a clear problem with heteroscedasticity, where variability in bacterial counts increases with mean bacterial count. This is why the variance-stabilizing logarithmic transformation is used in the analysis. These log-bacterial counts are assumed to be independent and normally distributed (after logarithmic transformation.)

Inspection of the ANOVA table below reveals that almost all (98.9%) of the variability in mean daily log-bacterial counts is explained by this model with site effects nested in plants, with most of the variability coming from the site effect. As subsequent multiple comparisons

The GLM Procedure

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	16	102.9552590	6.4347037	277.05	<.0001
Error	51	1.1845249	0.0232260		
Corrected Total	67	104.1397839			

R-Square      Coeff Var      Root MSE      ly Mean  
0.988626      1.577503      0.152401      9.660884

Source	DF	Type I SS	Mean Square	F Value	Pr > F
Site(Plant)	16	102.9552590	6.4347037	277.05	<.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Site(Plant)	12	98.96170585	8.24680882	355.07	<.0001
Plant	4	3.99355315	0.99838829	42.99	<.0001

A table of the untransformed site means within plant is given below. Inspection of this table reveals the inhomogeneity of variances: the rightmost column exhibits a tendency to increase with the mean. Multiple comparisons between sites within a fixed plant and between plants for a fixed site are then carried out using the transformed data, with Bonferroni correction for multiplicity. In the subscripts we use the following notation to indicate whether or not two mean bacterial counts are significantly different:

- two site means in the same plant with the same lower case letter are not significantly different.
- two plant means from the same site with the same upper case letter are not significantly different.

Site	Plant	N	Mean	Std Dev
1	A	4	50470.0000 a	10002.3431
2	A	4	17683.3333 b	201.3841
4	A	4	13633.3333 b	2740.2352
1	B	4	53833.3333 a	169.9673
2	B	4	12900.0000 b	2404.1631
4	B	4	7133.3333 c	659.9663
1	B2	4	58583.3333 a	8344.8201
2	B2	4	12770.0000 b	952.5755
4	B2	4	8893.3333 b	1392.3681
1	C	4	77083.3333 a	12262.5672
2	C	4	41623.3333 b	7724.9221
3	C	4	3833.3333 c	590.6682
4	C	4	2633.3333 c	169.9673
1	D	4	80103.3333 a	5682.2844
2	D	4	46040.0000 b	4969.5942
3	D	4	2080.0000 c	327.8211
4	D	4	2463.3333 c	697.4876

Site	Plant	N	Mean	Std Dev
1	A	4	A 50470.0000	10002.3431
1	B	4	AB 53833.3333	169.9673
1	B2	4	B 58583.3333	8344.8201
1	C	4	B 77083.3333	12262.5672
1	D	4	B 80103.3333	5682.2844
2	A	4	A 17683.3333	201.3841
2	B	4	A 12900.0000	2404.1631
2	B2	4	A 12770.0000	952.5755
2	C	4	B 41623.3333	7724.9221
2	D	4	B 46040.0000	4969.5942
3	C	4	A 3833.3333	590.6682
3	D	4	B 2080.0000	327.8211
4	A	4	A 13633.3333	2740.2352
4	B	4	B 7133.3333	659.9663
4	B2	4	B 8893.3333	1392.3681
4	C	4	C 2633.3333	169.9673
4	D	4	C 2463.3333	697.4876

For insight into these pairwise comparisons, a matrix of Bonferroni adjusted  $p$ -values is included on the next two pages below, as is an ANOVA table:

	Site	Plant	ly LSMEAN	LSMEAN Number
	1	A	10.8124925	1
	2	A	9.7803290	2
	4	A	9.5035941	3
	1	B	10.8936444	4
	2	B	9.4528103	5
	4	B	8.8693624	6
	1	B2	10.9703397	7
	2	B2	9.4527281	8
	4	B2	9.0836526	9
	1	C	11.2434607	10
	2	C	10.6228060	11
	3	C	8.2417669	12
	4	C	7.8744029	13
	1	D	11.2892251	14
	2	D	10.7329925	15
	3	D	7.6304460	16
	4	D	7.7809729	17

Least Squares Means for effect Site(Plant)  
Pr > |t| for H0: LSMean(i)=LSMean(j)

i/j	1	2	3	4	5	6
1		<.0001	<.0001	1.0000	<.0001	<.0001
2	<.0001		1.0000	<.0001	0.5081	<.0001
3	<.0001	1.0000		<.0001	1.0000	<.0001
4	1.0000	<.0001	<.0001		<.0001	<.0001
5	<.0001	0.5081	1.0000	<.0001		0.0002
6	<.0001	<.0001	<.0001	<.0001	0.0002	
7	1.0000	<.0001	<.0001	1.0000	<.0001	<.0001
8	<.0001	0.5070	1.0000	<.0001	1.0000	0.0002
9	<.0001	<.0001	0.0388	<.0001	0.1659	1.0000
10	0.0280	<.0001	<.0001	0.2814	<.0001	<.0001
11	1.0000	<.0001	<.0001	1.0000	<.0001	<.0001
12	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
13	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
14	0.0069	<.0001	<.0001	0.0787	<.0001	<.0001
15	1.0000	<.0001	<.0001	1.0000	<.0001	<.0001
16	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
17	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001

Adjustment for Multiple Comparisons: Bonferroni

i/j	7	8	9	10	11	12
1	1.0000	<.0001	<.0001	0.0280	1.0000	<.0001
2	<.0001	0.5070	<.0001	<.0001	<.0001	<.0001
3	<.0001	1.0000	0.0388	<.0001	<.0001	<.0001
4	1.0000	<.0001	<.0001	0.2814	1.0000	<.0001
5	<.0001	1.0000	0.1659	<.0001	<.0001	<.0001
6	<.0001	0.0002	1.0000	<.0001	<.0001	<.0001
7		<.0001	<.0001	1.0000	0.2993	<.0001
8	<.0001		0.1662	<.0001	<.0001	<.0001
9	<.0001	0.1662		<.0001	<.0001	<.0001
10	1.0000	<.0001	<.0001		<.0001	<.0001
11	0.2993	<.0001	<.0001	<.0001		<.0001
12	<.0001	<.0001	<.0001	<.0001	<.0001	
13	<.0001	<.0001	<.0001	<.0001	<.0001	0.1743
14	0.6351	<.0001	<.0001	1.0000	<.0001	<.0001
15	1.0000	<.0001	<.0001	0.0024	1.0000	<.0001
16	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
17	<.0001	<.0001	<.0001	<.0001	<.0001	0.0114

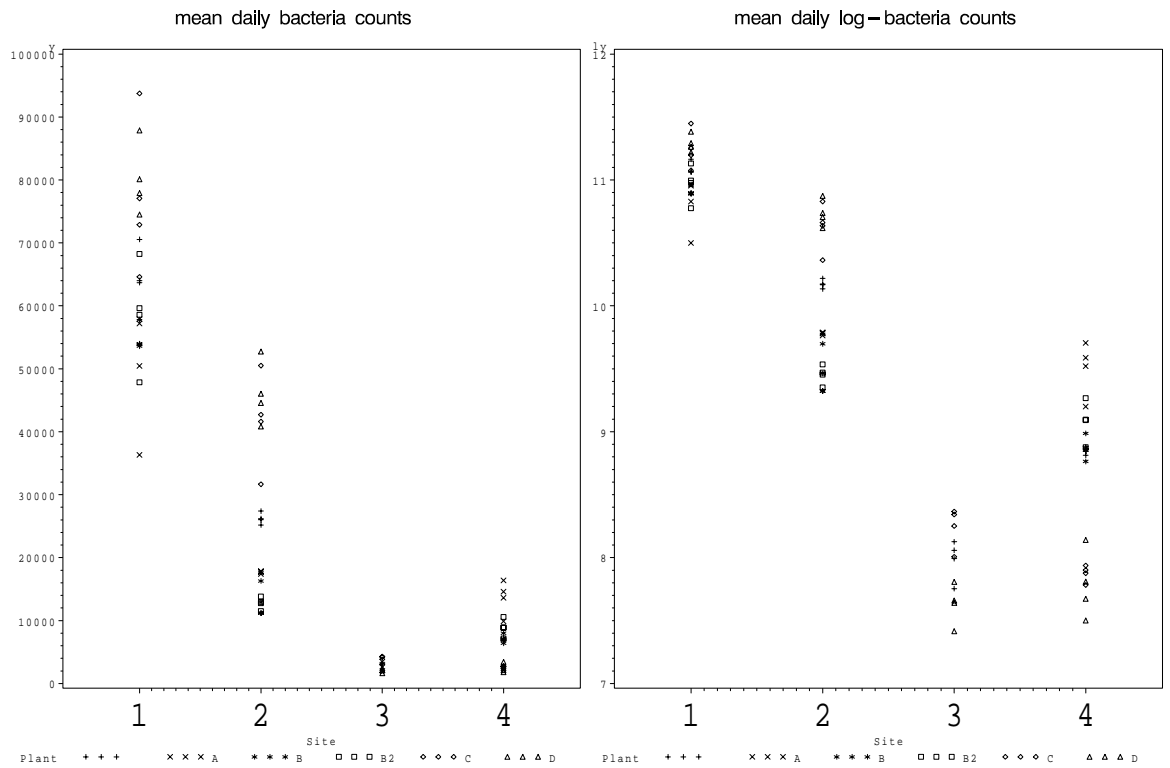
Least Squares Means for effect Site(Plant)

Pr > |t| for H0: LSMean(i)=LSMean(j)

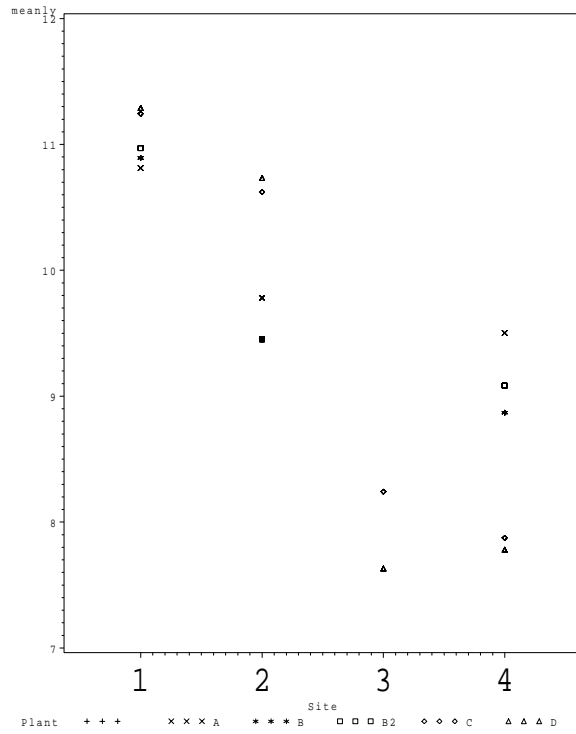
Dependent Variable: ly

i/j	13	14	15	16	17
1	<.0001	0.0069	1.0000	<.0001	<.0001
2	<.0001	<.0001	<.0001	<.0001	<.0001
3	<.0001	<.0001	<.0001	<.0001	<.0001
4	<.0001	0.0787	1.0000	<.0001	<.0001
5	<.0001	<.0001	<.0001	<.0001	<.0001
6	<.0001	<.0001	<.0001	<.0001	<.0001
7	<.0001	0.6351	1.0000	<.0001	<.0001
8	<.0001	<.0001	<.0001	<.0001	<.0001
9	<.0001	<.0001	<.0001	<.0001	<.0001
10	<.0001	1.0000	0.0024	<.0001	<.0001
11	<.0001	<.0001	1.0000	<.0001	<.0001
12	0.1743	<.0001	<.0001	<.0001	0.0114
13		<.0001	<.0001	1.0000	1.0000
14	<.0001		0.0006	<.0001	<.0001
15	<.0001	0.0006		<.0001	<.0001
16	1.0000	<.0001	<.0001		1.0000
17	1.0000	<.0001	<.0001	1.0000	

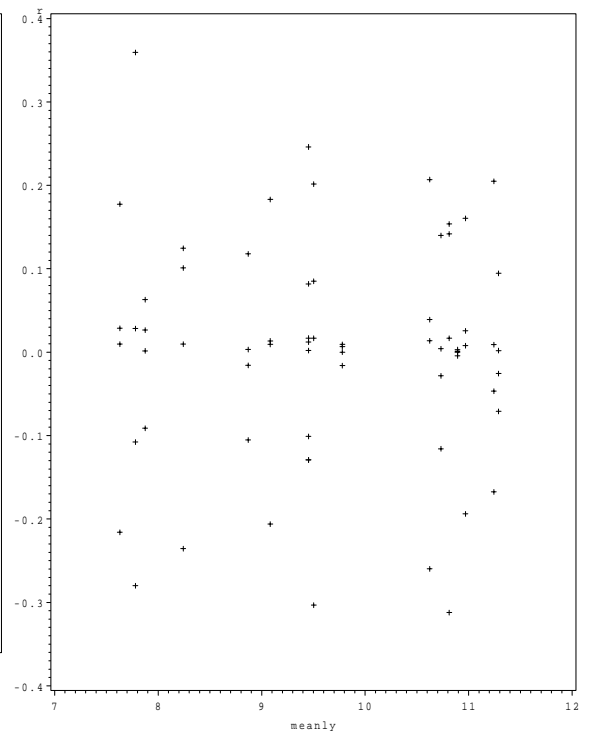
Conclusions: certainly, there is strong evidence ( $F = 355, df = 12, 51, p < 0.0001$ ) that bacterial counts decrease as birds go through the washer systems. There is also strong evidence ( $F = 43, df = 4, 51, p < 0.0001$ ) of plant-to-plant variability. However, pairwise comparisons among sites within each plant indicate that there were no significant differences between sites 3 and 4 in either of the two plants where these measurements were taken. Plant B was visited in two different seasons. When site means are compared across plants, no significant differences between plants B and B2 were found.



mean log-bacteria counts



residuals (log-transform) versus predicted



residuals (no transform) versus predicted

