Statistics 5303
Spring 2016
Writing Assignment #3

A soufflé is a baked food that is supposed to puff up out of the ramekin (cooking pan) and have a light, fluffy texture. Soufflés are notoriously difficult to bake. In this experiment, we make 24 soufflés and measure their heights. They are made by varying the following factors. Eggs are either whites only or whole eggs. The ramekin is either buttered or unbuttered (not greased). The ramekin is placed on either the bottom or middle oven rack. We make three soufflés for each combination of factors (in a completely randomized fashion) and measure their heights.

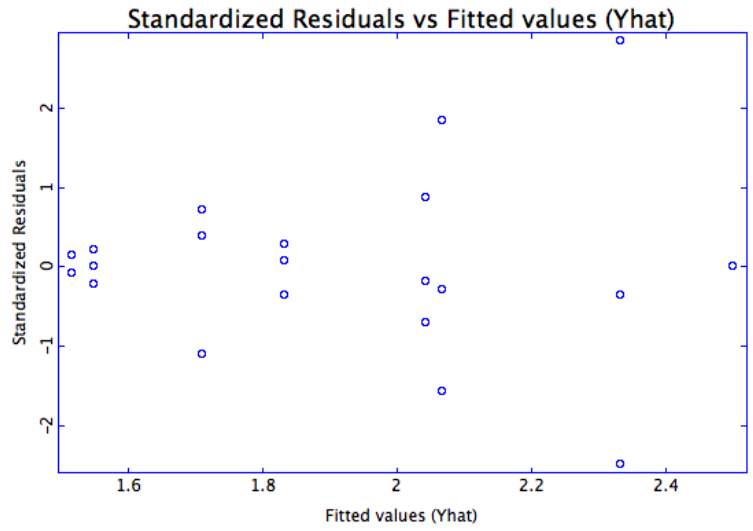
Below you will find a sketched analysis of the experiment. You will note that the analysis was not done with R; for example, what R would label "`Residual`" this package labels "`Error1`". The data for this experiment are in the file `souffle.dat.txt` on the class web page in `read.table(,header=TRUE)` format if you are interested.

You work for a restaurant chain, and your task is to write a **one-page** report on this experiment to your boss. Your boss is intelligent and understands science and the idea of experimentation, but your boss never took any statistics in college and thus knows relatively little about statistics.

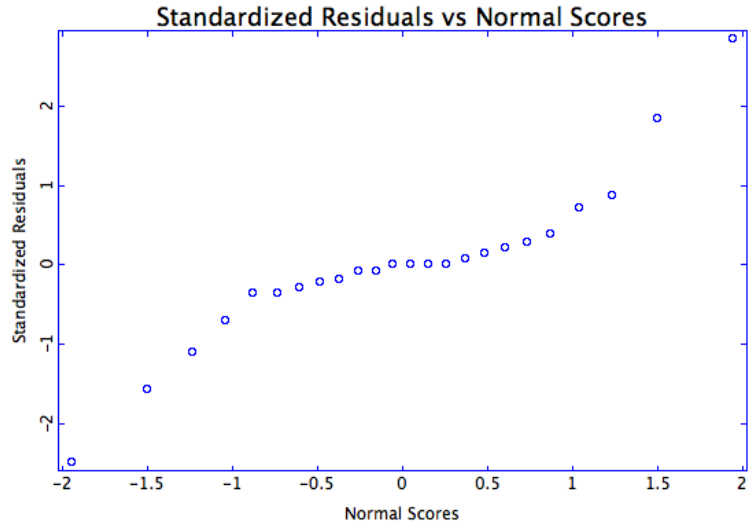
The experiment is a completely randomized design, with twenty-four units assigned to the factor/level combinations of eggs (whites or whole), butter (yes or no), and rack (bottom oven rack or middle). An initial analysis of variance follows:

	DF	SS	MS	F	P-value
CONSTANT	1	90.715	90.715	1105.82873	< 1e-08
egg	1	2.0417	2.0417	24.88826	0.00013385
rack	1	0.54	0.54	6.58269	0.020735
egg.rack	1	0.022817	0.022817	0.27814	0.60516
butter	1	0.045067	0.045067	0.54937	0.46931
egg.butter	1	0.00041667	0.00041667	0.00508	0.94407
rack.butter	1	0.020417	0.020417	0.24888	0.62465
egg.rack.butter	1	0.0010667	0.0010667	0.01300	0.91063
ERROR1	16	1.3125	0.082033		

Residuals from this ANOVA show strong evidence of non-constant variance.



Residuals are also long tailed.



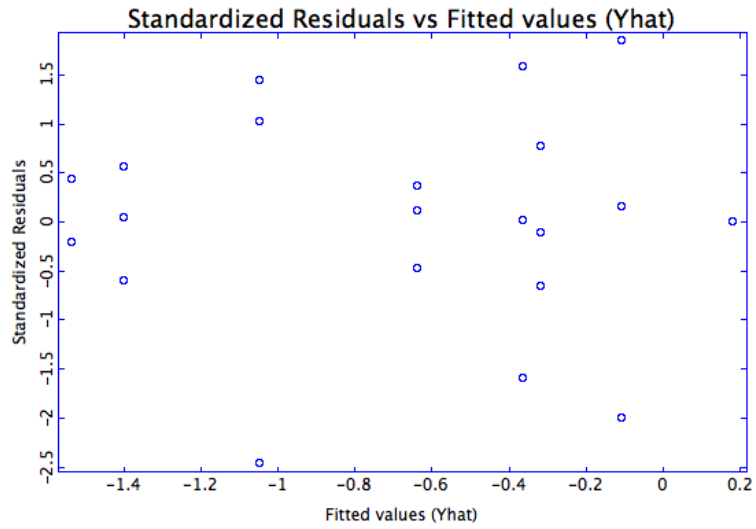
The heights range from 1.45 to 3, a ratio of just over 2. With such a small dynamic range, Box-Cox transformations will do very little to improve things, or put another way, it will take a very extreme power to make any difference. To correct for this, we define "excess height," which is height above 1.3. The thought is that all souffles will have a minimum height caused by the depth of batter used, even if they don't puff up at all. What we would really like is the amount above minimum. We don't know what that minimum is, but 1.3 is a bit below the lowest observed value and offers a reasonable modification.

Using excess height, Box-Cox analysis indicates that a log transformation will improve the stability of variance. The ANOVA for log excess height follows:

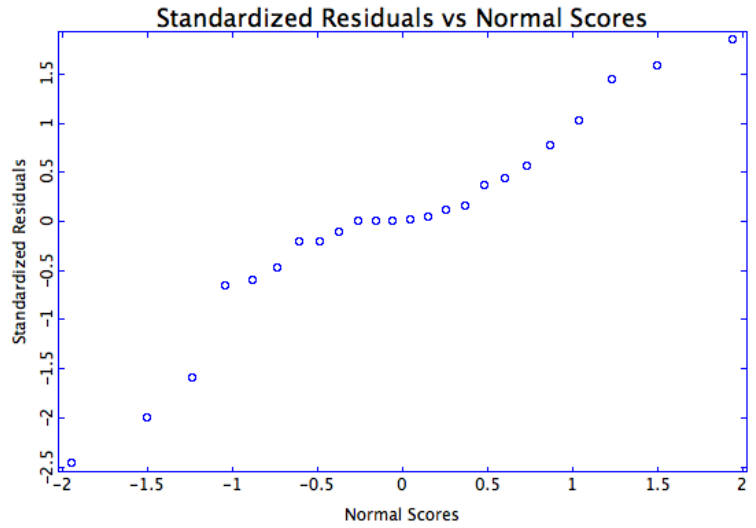
	DF	SS	MS	F	P-value
CONSTANT	1	10.222	10.222	56.88853	1.1804e-06
egg	1	6.0315	6.0315	33.56887	2.7424e-05
rack	1	1.517	1.517	8.44278	0.010319
egg.rack	1	0.0927	0.0927	0.51593	0.48294
butter	1	0.23236	0.23236	1.29320	0.2722
egg.butter	1	0.034134	0.034134	0.18997	0.66876

```
rack.butter          1          0.139          0.139       0.77359       0.39213
egg.rack.butter      1       0.001376       0.001376       0.00766       0.93135
ERROR1              16         2.8748        0.17968
```

On this scale, residuals have more constant variance



and look less long tailed (though still a little strange):



The only factors that matter are egg and rack. Using whites instead of whole eggs increases log excess height by 1.00 (95% CI .64 to 1.37). On the original scale, this corresponds to an increase in median excess height by a factor of $\exp(1) = 2.72$. Using the lower rack instead of the middle rack increases log excess height by .50 (95% CI .14 to .87). On the original scale, this corresponds to an increase in median excess height by a factor of $\exp(.5) = 1.65$. Taken together, we estimate the increase in median excess height obtained by using whites and lower rack to be 4.49 over using whole eggs and the middle rack.