ENGR 599 Homework 5

Adrian Henle

```
using Distributions , PlutoUI
```

Exercise 6.1

Eliminating the center point from **Table 6.1** and **Fig. 6.1**, we are left with a 2^2 factorial design. Calculate the effect values for this factorial and compare them with the values of the coefficients in **Eq. 6.3**:

$$\hat{y} = 68 - 5.25x_1 + 4.25x_2$$

$$x_1 = -10.5$$

$$x_2 = 8.5$$

$$x_1x_2 = -0.5$$

The x_1 and x_2 effects are double the coefficients in **Eq. 6.3**. The interaction effect is very small by comparison.

Exercise 6.3

Evaluate the statistical significance of **Eq. 6.3** using an *F*-test and **Table 6.2**. Compare the ratio between the regression mean square and the residual mean square with the appropriate *F* value.

Table 6.2

Source of variation	Sum of squares	Degree of freedom	Mean square
Regression	182.50	2	91.25
Residual	5.50	4	1.38
Lack of fit	0.83	2	0.42
Pure error	4.67	2	2.34
Total	188.00	6	

% explained variation: 97.07
Maximum % explainable variation: 97.52

```
begin
    mean_square_regression = 91.25
    mean_square_residual = 1.38
end;
```

ratio = 66.12318840579711

```
ratio = mean_square_regression / mean_square_residual
```

F = 6.944271909999156

```
F = quantile(FDist(2,4), 0.95)
```

Exercise 6.5

Imagine that, in the C. *elegans* example, the researchers had preferred to take the concentration of glucose as the starting factor to determine the steepest ascent path, with an initial displacement of $25g/L$. Calculate the coordinates of the third point along the new path, and use **Eq. 6.5** to predict chitin yield.

Eq 6.5:

$$\hat{y} = 19.8 + 2x_1 + 5x_2 + 2.5x_3$$

Table 6.4
Levels of a 2^3 design with a center point, to study how chitin production by the fungus C. elegans is affected by the concentrations of three nutrients in the culture medium

	Factor	Level		
		–1	0	+1
G (x_1)	D-Glucose ($\mathrm{g\,L^{-1}}$)	20	40	60
A (x_2)	L-Asparagine ($\mathrm{g\,L^{-1}}$)	1	2	3
T (x_3)	Thiamine ($\mathrm{mg\,L^{-1}}$)	0.02	0.05	0.08

First point:

$$\Delta x_1 = \frac{2}{5}(+1) = +0.4$$

$$\Delta x_2 = +1$$

$$\Delta x_3 = \frac{2.5}{5}(+1) = +0.5$$

So,

$$G_n = \begin{cases} 25 & n = 1 \\ G_{n-1} + 8 * n & n > 1 \end{cases}$$

$$A_n = \begin{cases} 2 & n = 1 \\ A_{n-1} + n & n > 1 \end{cases}$$

$$T_n = \begin{cases} 0.05 & n = 1 \\ T_{n-1} + 1.5 * 10^{-3} n & n > 1 \end{cases}$$

```
begin
    # function generator
    f = (a, b, c) -> (n) -> n == 1 ? a : f(a, b, c)(n-1) + b*c
    # functions for params
    g = f(25, 0.4, 20)
    a = f(2, 1, 1)
    t = f(0.05, 0.5, 1.5e-3)
end;
```

```
enpoint =   (41.0, 4, 0.0515)
```

```
enpoint = (g(3), a(3), t(3))
```

Exercise 6.7

Use the data in **Table 6.8** to calculate a value showing that **Eq. 6.8** is statistically significant.

Table 6.8
Analysis of variance for fitting the $\hat{y} = b_0 + b_1x_1 + b_2x_2 + b_{11}x_1^2 + b_{22}x_2^2 + b_{12}x_1x_2$ model to the data in Table 6.7

Source of variation	Sum of squares	Degree of freedom	Mean square
Regression	144.15	5	28.83
Residual	2.76	5	0.55
Lack of fit	0.76	3	0.25
Pure error	2.00	2	1.00
Total	146.91	10	

% explained variation: 98.12
Maximum % explainable variation: 98.64

```
begin
    MSR = 28.83
    MSr = 0.55
    local F = quantile(FDist(5,5), 0.95)
end;
```

true

```
MSR/MSr > F
```

Exercise 6.9

Use the data in **Table 6.10** to calculate an experimental error estimate with more than 79 degrees of freedom.

Table 6.10
Analyses of variance for the fits of the linear and quadratic models to the Young's modulus data of Table 6.9. The values for the quadratic model are given in parentheses

Source of variation	Sum of squares	Degree of freedom	Mean square
Regression	37.34 (*43.23*)	3 (*9*)	12.45 (*4.80*)
Residual	8.44 (*2.55*)	102 (*96*)	0.088 (*0.028*)
Lack of fit	6.76 (*0.87*)	23 (*17*)	0.29 (*0.051*)
Pure error	1.68	79	0.023
Total	45.78	105	

% explained variation: 81.56 (94.43)
Maximum % explainable variation: 96.33

The residual mean square has 96 degrees of freedom. Taking the root of its error as an estimate of pure error gives $\sqrt{MS_r} = 0.167$