Statistics 101B - Spring 2020
Homework 4 (Due Monday May 18 at 5:00pm)

** Please upload your homework on CCLE. Please write your full name, student ID number, and section number on your homework.
Must be PDF.
NO late work will be accepted!**
You do not need to write the questions in your homework. Please show your work!

1. Problem 6.22
2. Problem 6.17
   Add:
   (c) Identify the significant effects through constructing 95% confidence intervals. Do this first
      (i) if it is reasonable to assume that the interactions involving three or more factors are negligible, and then
      (ii) using Lenth's method.

6.13 An industrial engineer employed by a beverage bottler is interested in the effects of two different types of 32-ounce bottles on the time to deliver 12-bottle cases of the product. The two bottle types are glass and plastic. Two workers are used to perform a task consisting of moving 40 cases of the product 50 feet on a standard type of hand truck and stacking the cases in a display. Four replicates of a 2^2 factorial design are performed, and the times observed are listed in the following table. Analyze the data and draw appropriate conclusions. Analyze the residuals and comment on the model's adequacy.

	Worker			
Bottle Type	**1**		**2**	
Glass	5.12	4.89	6.65	6.24
	4.98	5.00	5.49	5.55
Plastic	4.95	4.43	5.28	4.91
	4.27	4.25	4.75	4.71

6.14 An article in the *AT&T Technical Journal* (March/April 1986, Vol. 65, pp. 39–50) describes the application of two-level factorial designs to integrated circuit manufacturing. A basic processing step is to grow an epitaxial layer on polished silicon wafers. The wafers mounted on a susceptor are positioned inside a bell jar, and chemical vapors are introduced. The susceptor is rotated, and heat is applied until the epitaxial layer is thick enough. An experiment was run using two factors: arsenic flow rate (A) and deposition time (B). Four replicates were run, and the epitaxial layer thickness was measured (μm). The data are shown in Table P6.1.

(a) Estimate the factor effects.

(b) Conduct an analysis of variance. Which factors are important?

(c) Write down a regression equation that could be used to predict epitaxial layer thickness over the region of arsenic flow rate and deposition time used in this experiment.

(d) Analyze the residuals. Are there any residuals that should cause concern?

(e) Discuss how you might deal with the potential outlier found in part (d).

6.15 *Continuation of Problem 6.14.* Use the regression model in part (c) of Problem 6.14 to generate a response surface contour plot for epitaxial layer thickness. Suppose it is critically important to obtain layer thickness of 14.5 μm. What settings of arsenic flow rate and decomposition time would you recommend?

6.16 *Continuation of Problem 6.15.* How would your answer to Problem 6.15 change if arsenic flow rate was more difficult to control in the process than the deposition time?

6.17 An experimenter has run a single replicate of a 2^4 design. The following effect estimates have been calculated:

$$\begin{array}{lll} A = 76.95 & AB = -51.32 & ABC = -2.82 \\ B = -67.52 & AC = 11.69 & ABD = -6.50 \\ C = -7.84 & AD = 9.78 & ACD = 10.20 \\ D = -18.73 & BC = 20.78 & BCD = -7.98 \\ & BD = 14.74 & ABCD = -6.25 \\ & CD = 1.27 & \end{array}$$

(a) Construct a normal probability plot of these effects.

(b) Identify a tentative model, based on the plot of the effects in part (a).

6.18 The effect estimates from a 2^4 factorial design are as follows: ABCD = −1.5138, ABC = −1.2661, ABD = −0.9852, ACD = −0.7566, BCD = −0.4842, CD = −0.0795, BD = −0.0793, AD = 0.5988, BC = 0.9216, AC = 1.1616, AB = 1.3266, D = 4.6744, C = 5.1458, B = 8.2469, and A = 12.7151. Are you comfortable with the conclusions that all main effects are active?

6.19 The effect estimates from a 2^4 factorial experiment are listed here. Are any of the effects significant? ABCD = −2.5251, BCD = 4.4054, ACD = −0.4932, ABD = −5.0842, ABC = −5.7696, CD = 4.6707, BD = −4.6620, BC = −0.7982, AD = −1.6564, AC = 1.1109, AB = −10.5229, D = −6.0275, C = −8.2045, B = −6.5304, and A = −0.7914.

■ **TABLE P6.1**
The 2^2 Design for Problem 6.14

		Replicate					Factor Levels	
A	***B***	**I**	**II**	**III**	**IV**		**Low (−)**	**High (+)**
−	−	14.037	16.165	13.972	13.907	*A*	55%	59%
+	−	13.880	13.860	14.032	13.914			
−	+	14.821	14.757	14.843	14.878	*B*	Short	Long
+	+	14.888	14.921	14.415	14.932		(10 min)	(15 min)

SS **6.20** Consider a variation of the bottle filling experiment from Example 5.3. Suppose that only two levels of carbonation are used so that the experiment is a 2^3 factorial design with two replicates. The data are shown in Table P6.2.

(a) Analyze the data from this experiment. Which factors significantly affect fill height deviation?

(b) Analyze the residuals from this experiment. Are there any indications of model inadequacy?

(c) Obtain a model for predicting fill height deviation in terms of the important process variables. Use this model to construct contour plots to assist in interpreting the results of the experiment.

(d) In part (a), you probably noticed that there was an interaction term that was borderline significant. If you did not include the interaction term in your model, include it now and repeat the analysis. What difference did this make? If you elected to include the interaction term in part (a), remove it and repeat the analysis. What difference does the interaction term make?

SS **6.21** I am always interested in improving my golf scores. Since a typical golfer uses the putter for about 35–45 percent of his or her strokes, it seems reasonable that improving one's putting is a logical and perhaps simple way to improve a golf score ("The man who can putt is a match for any man."—Willie Parks, 1864–1925, two-time winner of the British Open). An experiment was conducted to study the effects of four factors on putting accuracy. The design factors are length of putt, type of putter, breaking putt versus straight putt, and level versus downhill putt. The response variable is distance from the ball to the center of the cup after the ball comes to rest. One golfer performs the experiment, a 2^4 factorial design with seven replicates was used, and all putts are made in random order. The results are shown in Table P6.3.

(a) Analyze the data from this experiment. Which factors significantly affect putting performance?

(b) Analyze the residuals from this experiment. Are there any indications of model inadequacy?

6.22 A company markets its products by direct mail. An experiment was conducted to study the effects of three factors on the customer response rate for a particular product. The three factors are A = type of mail used (3rd class, 1st class), B = type of descriptive brochure (color, black-and-white), and C = offered price ($19.95, $24.95). The mailings are made to two groups of 8000 randomly selected customers, with 1000 customers in each group receiving each treatment combination. Each group of customers is considered as a replicate. The response variable is the number of orders placed. The experimental data are shown in Table P6.4.

(a) Analyze the data from this experiment. Which factors significantly affect the customer response rate?

(b) Analyze the residuals from this experiment. Are there any indications of model inadequacy?

(c) What would you recommend to the company?

6.23 Consider the single replicate of the 2^4 design in Example 6.2. Suppose that we had arbitrarily decided to analyze the data assuming that all three- and four-factor interactions were negligible. Conduct this analysis and compare your results with those obtained in the example. Do you think that it is a good idea to arbitrarily assume interactions to be negligible even if they are relatively high-order ones?

6.24 An experiment was run in a semiconductor fabrication plant in an effort to increase yield. Five factors, each at two levels, were studied. The factors (and levels) were A = aperture setting (small, large), B = exposure time (20% below nominal, 20% above nominal), C = development time (30 and 45 s), D = mask dimension (small, large), and E = etch time SS

■ **TABLE P6.4**

The Direct Mail Experiment from Problem 6.22

	Coded Factors			Number of Orders		Factor Levels		
Run	*A*	*B*	*C*	Replicate 1	Replicate 2		Low (−1)	High (+1)
1	−	−	−	50	54	*A* (class)	3rd	1st
2	+	−	−	44	42	*B* (type)	BW	Color
3	−	+	−	46	48	*C* ($)	$19.95	$24.95
4	+	+	−	42	43			
5	−	−	+	49	46			
6	+	−	+	48	45			
7	−	+	+	47	48			
8	+	+	+	56	54			