Reference: Chapter 6 of Montgomery(8e) **Maghsoodloo**

DOE (or DOX) FOR BASE-2 BALANCED FACTORIALS

The notation 2^k is used to denote a factorial experiment involving *k* factors (A, B, C, D, ..., K) each at 2 levels. For example, the notation 2^4 implies that we have 4 factors (A, B, C, D) each at 2 levels (low = 0 or –1, and high = 1 or +1). Since $2^4 = 16$, then we have 16 FLCs, the corresponding OA will have 16 rows, and the exponent 4 implies we can write 4 columns arbitrarily, one column per factor. Further, there is also a specific notation that is applicable only and only to FLCs of base-2 designs. For the case of 2^4 factorial, the symbol "*ac*" represents the FLC where factors A and C are at their high levels while factors B and D are at their low levels. In other words, the presence of a small letter indicates the corresponding factor is at high level while the absence of a small letter indicates the factor is at its low level. When all factors are at their low levels, then the symbol (1) is used to represent the FLC (0, 0, ..., 0) = (–1, –1, –1, –1, ..., –1). Further, the notation "*b*" implies that factors A, C, and D are at their low levels while B is at its high level, etc. Table 6.10 of Example 6.2, on page 257 of Montgomery(8e), illustrates this concept very well. In Table 6.11 at the bottom of page 258, the 4 columns pertaining to the 4 main factors (A, B, C, & D) were written completely arbitrarily as illustrated atop the next page, while the interaction columns were obtained by simple multiplication of the factors involved in the interaction. Table 6.10 of Montgomery(8e), reproduced atop the next page, shows that the total effect (or contrast) of the 2nd-order (or 3-way) interaction ACD is computed as follows: Contrast (ACD) = $-(1) + a - b + ab + c - ac + bc - abc + d - ad + bd - abd - cd + acd - bcd + abcd$. For the data of the Example 6.2 on page 257 of Montgomery's 8th edition, we obtain the Contrast (ACD) = – 45 + 71 – 48 + 65 + 68 – 60 + 80 – 65 + 43 – 100 + 45 –104 – 75 + 86 – 70 + 96 = –13.0. Since there are 8 pair-wise comparisons (with n = 1 observations per FLC in this example), then the average effect of ACD interaction is computed from $\overline{\text{ACD}} = \dfrac{\text{Contrast (ACD)}}{8\text{n}} = \dfrac{\text{Contrast}}{\text{n}\,2^{k-1}}$ = –13.00 / 8 = – 1.625,

which agrees with Table 6.12 atop page 258 of Montgomery's 8th edition. For practice, you should write the contrasts for BD and ABCD interaction effects using Table 6.10 of

One full replicate of a 2^4 factorial (Table 6.10 of Montgomery(8e) on his page 257)

D	C	B	A	FLC	AC=A×C	BD = B×D	A×B×C×D	ACD	y_{ijkL}
-	-	-	-	(1)	+	+	+	-	45
-	-	-	+	a	-	+	-	+	71
-	-	+	-	b	+	–	-	-	48
-	-	+	+	ab	-	–	+	+	65
-	+	-	-	c	-	+	-	+	68
-	+	-	+	ac	+	+	+	-	60
-	+	+	-	bc	-	–	+	+	80
-	+	+	+	abc	+	–	-	-	65
+	-	-	-	d	+	–	-	+	43
+	-	-	+	ad	-	–	+	-	100
+	-	+	-	bd	+	+	+	+	45
+	-	+	+	abd	-	+	-	-	104
+	+	-	-	cd	-	–	+	-	75
+	+	-	+	acd	+	–	-	+	86
+	+	+	-	bcd	-	+	-	-	70
+	+	+	+	abcd	+	+	+	+	96

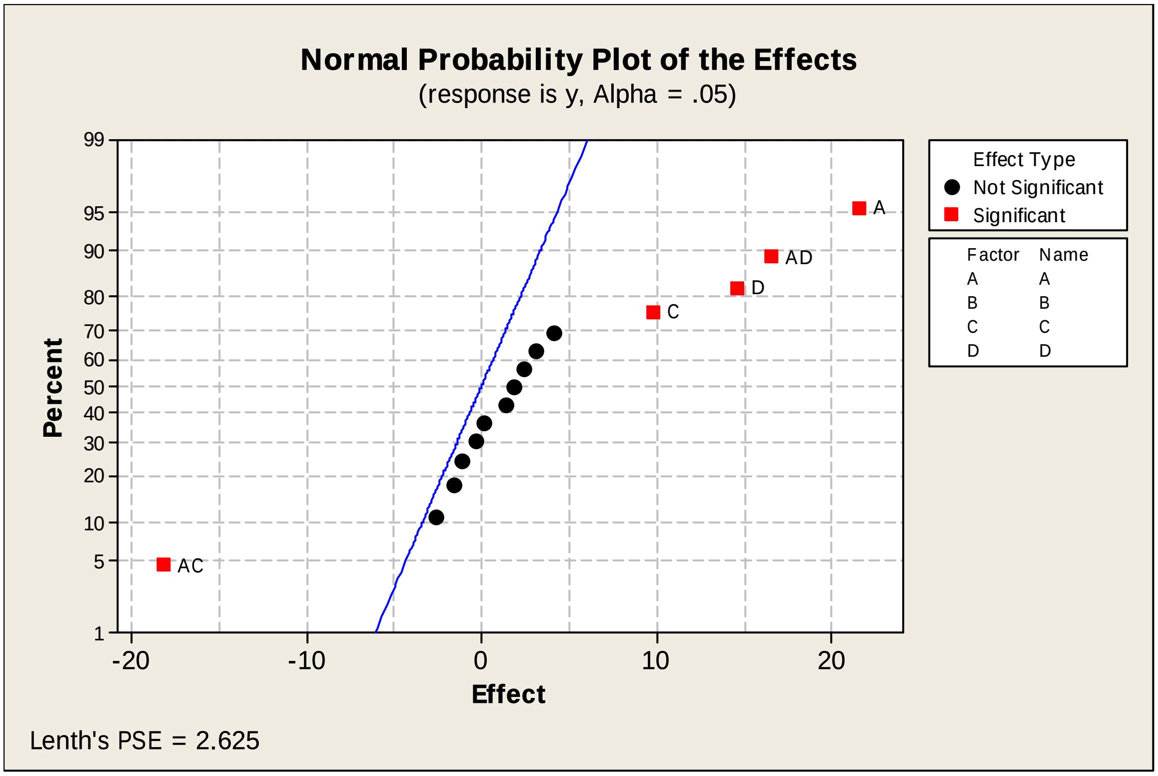
Montgomery and try to determine if a pattern develops for writing any contrast in base-2 designs. Some of you are probably familiar with the pattern that I am referring to, namely Odd and Even Rules.

<u>The Odd and Even Rules.</u> If the effect has an odd number of letters, such as ACD 2^{nd}-order interaction, then we must use the Odd Rule to determine the signs in its contrast, i.e., odd number of letters in common with ACD must be assigned a "+" sign and even number of letters a "–" sign, and vice a versa for the Even Rule.

As an example we use the Even Rule to obtain the contrast for the AC effect (i.e., even number of letters in common with AC will receive a positive sign, where 0 is even).

Contrast (AC) = (1) – *a* + *b* – *ab* – *c* + *ac* – *bc* + *abc* + *d* – *ad* + *bd* – *abd* – *cd* + *acd* – *bcd* + *abcd* = 45 – 71 + 48 – 65 – 68 + 60 – 80 + 65 + 43 – 100 + 45 – 104 – 75 + 86 – 70 + 96 = – 145 $\rightarrow$ $\overline{\text{AC}}$ = –145/8 = –18.125, which is in agreement with Montgomery's Table 6.12 on page 258 of the 8th edition.

As before, the SS of any contrast is given by SS(contrast) = $(\text{Contrast value})^2 / \sum_{i=1}^{N} n_i c_i^2$, where $N = 2^k$, which is $N = 16$ for the Example 6.2 of Montgomery on pp. 257-262. For Table 6.10 of Montgomery presented above, SS(ACD) = $(-13)^2/16$ = 10.5625 and SS(AC)= $(-145)^2/16$ = 1314.0625, which agree to 2 decimals with the ANOVA Table 6.13 on page 260 of Montgomery(8e). It can be shown that for base-2 designs, the SS of a contrast can also be obtained from SS(Contrast) = $n \times (\text{Average Effect})^2 \times 2^{\boldsymbol{k}-2}$. Since in Montgomery's Example 6.2 there is only 1 observation per cell and in general the residual $e_{ijkLm} = y_{ijkLm} - \bar{\mathbf{y}}_{ijkL.} = y_{ijkL} - y_{ijkL} \equiv 0$ for all 16 observations, then we must develop a regression model for the data in order to obtain the 16 residuals. Further, because there is no estimate of pure error, no exact statistical test of significance can be conducted. However, there are procedures that identify relatively large effects, the most important of which is the Normal Pr Plot (NPrPL) of average effects as shown in Figure 6.11 on page 259 of Montgomery. In order to draw the NPrPL of average effects, go to Minitab's Stat $\rightarrow$ DOE $\rightarrow$ Factorial $\rightarrow$ Define Custom factorial $\rightarrow$ Input factors $\rightarrow$ Click on High/Low $\rightarrow$ Select coded; then go to Analyze Factorial Design $\rightarrow$ click on Yes to info that Minitab requires for factors, providing the factor names (A B C D) $\rightarrow$ click on High/Low $\rightarrow$ click on coded $\rightarrow$ ok $\rightarrow$ Response: y $\rightarrow$ Graphs $\rightarrow$ click on normal for Effects Plots $\rightarrow$ ok $\rightarrow$ ok. Minitab's NPrPL is provided atop the next page. In order to ascertain which average effects are important, first one must put a fat pencil on the Pr plot line and conclude that the covered effects under the fat pencil are relatively unimportant. The above figure from Minitab clearly shows that the only important effects are A, AC, AD, D, and C because they will not be covered by a fat pencil. Another conservative rule-of-thumb observed by me in the past 30 years is to obtain the range of absolute average effects, denoted by R_E, and divide this range by



4 and 3; then declare any effect below $R_E/4$ relatively insignificant. For the Example 6-2 on p. 257 of Montgomery $R_E/4 = (21.625 - 0.125)/4 = 5.375$ and $R_E/3 = 7.1667 \rightarrow$ This implies that any average effect in Table 6.12 on page 257 of Montgomery(8e) whose absolute value is less than 5.375 is relatively unessential and should be ignored and all effects above 7.1667 are relatively influential. Average effects whose value lie between $R_E/4$ and $R_E/3$ are at best moderately influential. In order to understand how Minitab computes Lenth's pseudo standard error PSE = 2.625, the reader must refer to page 262-264 of Montgomery's 8th edition.

We will include only the relatively large effects in the model (see Figure 6.11 on page 259 of Montgomery), namely $A = x_1$, $C = x_3$, $D = x_4$, $A\times C = x_1x_3 = x_{13}$, $A\times D = x_1x_4 = x_{14}$, where x_1, x_3 and $x_4 = -1$ or $+1$ just like given in Table 6.10 on page 257 of Montgomery. Thus our regression model is:

$$y_{ijkL} = \beta_0 + \beta_1x_1 + \beta_3x_3 + \beta_4x_4 + \beta_{13}x_1x_3 + \beta_{14}x_1x_4 + \epsilon_{ijkL}. \qquad (15)$$

Note that x_1 is simply the 16×1 vector of -1 and $+1$ listed in column A of Table 6.10 on page 257 of Montgomery, x_3 represents the column C in the same table, and so on; further x_1, x_3 x_4,

x_1x_3, and x_1x_4 are orthogonal (i.e., their pair-wise dot product is zero). To obtain the point LSQ estimate of β_0, we sum both sides of (15) over i, j, k, L = 1, 2 and assume that $\sum \in_{ijkL} \approx 0. \rightarrow y_{....}$ = $16\hat{\beta}_0 \rightarrow \hat{\beta}_0 = \overline{y}_{....}$ = 1121/16 = 70.0625. To obtain a point estimate of β_1, multiply both sides of (15) by x_1 and then sum both sides over i, j, k, L = 1, 2. $\rightarrow$ –45 + 71–48 + 65 +… – 70 + 96 = $\hat{\beta}_1[(-1)^2 + 1^2 +(-1)^2 +\ldots +1^2] \rightarrow \hat{\beta}_1$ = 173/16 = 10.8125. Note that $\hat{\beta}_1$ is simply (1/2) of the average effect of A given in Table 6.12 of page 258. Similarly, – 521 + 600 = $\hat{\beta}_3(16) \rightarrow \hat{\beta}_3$ = 4.9375 which is again equal to $\overline{c}/2$ = 9.875/2, and $\hat{\beta}_4$ = (619 –502)/16 = 7.3125. To obtain a point estimate of β_{13}, multiply both sides of Eq. (15) by x_1 and x_3 (or by the AC column in Table 6.11 at the bottom of page 258 Montgomery) and sum over i, j, k, L = 1, 2. $\rightarrow$ $\hat{\beta}_{13}$ = (488 – 633)/16 = – 9.0625 = $\overline{A \times C}\ /2 = \overline{AC}/2 =$ –18.125/2. Similarly, $\hat{\beta}_{14} =$ (627 – 494)/16 = 8.3125. The corresponding regression model is given on page 261 of Montgomery(8e) and you may verify the residuals given atop page 261 of Montgomery(8e) by Minitab.

OBTAINING AN INDEPENDENT ESTIMATE OF σ_ϵ^2 BY ADDING CENTER POINTS TO A 2^k DESIGN

For the sake of illustration, consider Example 6-6 on pages 271-275 of Montgomery's 5th edition (ISBN: 0-471-31649-0), where we have one replicate of a 2^2 design with n_C = 5 observations at the design center (35 minutes, 155 ^{0}C). Because you do not have the 5th edition of this text, I will exactly quote the details below as given in Montgomery's 5th edition on his page 273.

EXAMPLE 6-6 on Page 273 of Montgomery's 5th Edition Quoted Below:

"A chemical engineer is studying the yield of a process. There are two variables of interest, reaction time and reaction temperature. Because she is uncertain about the assumption of linearity over the region of exploration, the engineer decides to conduct a 2^2 design (with a single replicate of each factorial run) augmented with five center points." The design and yield data are shown in Figure 6-35 of Montgomery's 5th edition, and also shown on the next page.

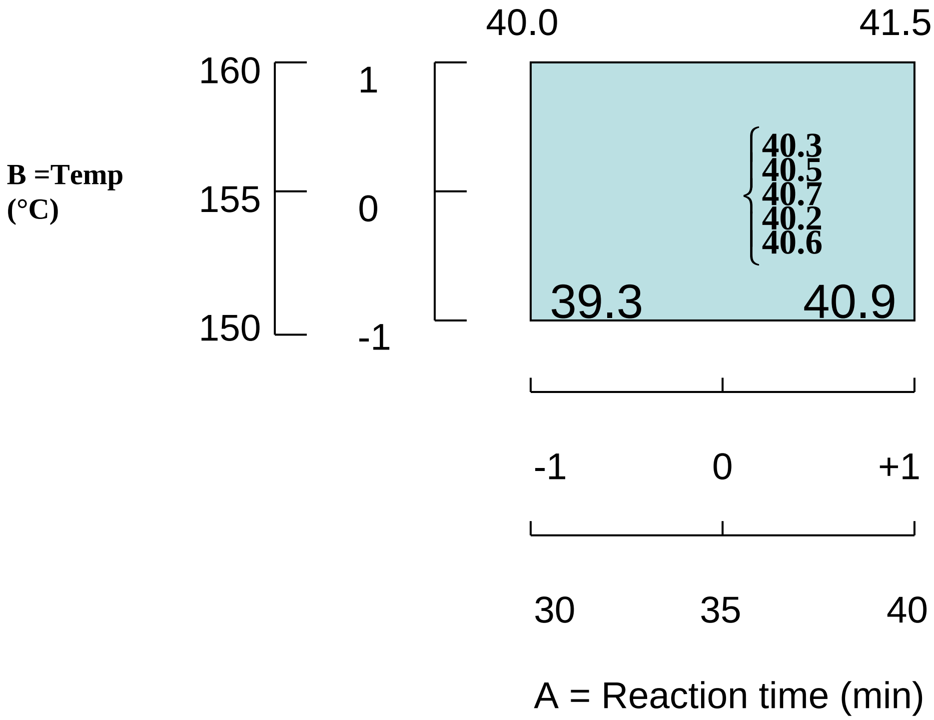


Figure 6-35 on page 273 of Montgomery's 5th Edition

The general 2^2 design with n_C center points is depicted in Figure 6.38 on page 286 of Montgomery's 8th edition. My explanation now follows.

The Total SS is computed in the usual manner as SS_T = USS – CF = 14724.78 – $364^2/9$ = 3.002222 (8 *df*). The pure Error SS is computed from the 5 replicates at the center, i.e., SS(Pure Error) = SS_{PE} = 8185.23 – $202.3^2/5$ = 0.1720 (with 4 *df*), which yields SS(Model) = SS_T – SS(Error) = 2.830222 (4 *df*); this SS(Model) can also be computed from the 5 cell subtotals ($40^2 + 41.5^2 + 39.3^2 + 40.9^2 + 202.3^2/5$ – CF). Clearly, effects A, B and A×B each carry 1 *df* but the model has 4 *df* and hence there remains 1 *df* for the joint quadratic effects of A and B as illustrated by the OA in Table 8 below. Table 8 clearly indicates that the columns A^2 and B^2 are identical, and therefore, the quadratic effects of A and B are hopelessly confounded (i.e., aliased, inseparable, or indistinguishable from each other). Hence the 2 columns A^2 and B^2 together carry 1 *df*. To compute the SS that explains the average quadratic effects of A and B, we note that there are n_F = 4 factorial points but n_C = 5 center points so that in defining the contrast for $\overline{A^2 + B^2}$ we must bear in mind this imbalance. Further, the quadratic effect of factors A or B is simply

Table 8. Note that the 4/9 is simply the average of $A^2 = B^2$ columns.

$A=x_1$	$B=x_2$	$A\times B= x_1x_2$	$A^2 = x_1^2$	$B^2= x_2^2$	$x_1^2 - 4/9$	y_{ijk}
–1	–1	+1	+1	+1	5/9	39.3 = (1)
–1	+1	–1	+1	+1	5/9	40.0 = b
+1	–1	–1	+1	+1	5/9	40.9 = a
+1	+1	+1	+1	+1	5/9	41.5 = ab
0	0	0	0	0	–4/9	40.3 = (0,0)
0	0	0	0	0	–4/9	40.5 = (0,0)
0	0	0	0	0	–4/9	40.7
0	0	0	0	0	–4/9	40.2
0	0	0	0	0	–4/9	40.6 = (0, 0)

the contrast between the factorial and center points. To this end, we define the SS of this contrast as SS($\overline{A^2 + B^2}$) = $\dfrac{(5\times 161.7 - 4\times 202.3)^2}{4(5^2)+5(-4)^2}$ = $0.0027222\overline{2}$, which is identical to the value provided in the ANOVA Table 6-20 on page 274 of Montgomery's 5th edition to 4 decimals as his ANOVA Table provides only 4 decimal accuracy.

The design matrix in Table 8 above is said to be orthogonal because the dot product of any 2 columns is identically zero. Using statistical regression theory, it can be shown that adding center points to an orthogonal design, not only does not alter orthogonality, but also does not alter the SS of different effects already in the design matrix, which in this case are A, B and the AB interaction effect. Hence, SS(A)= $(-79.3 + 82.4)^2/4 = 2.4025$, SS(B) = $(81.5 - 80.2)^2/4 = 0.4225$, SS(A×B) = $(80.8 - 80.9)^2/4 = 0.0025$, which are in agreement with those in Table 6-20 on page 274 of Montgomery's 5th edition. In order to obtain the coefficients of the 2nd-order model

$$y_{ijk} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_{11} x_1^2 \text{ (or } \beta_{22} x_2^2) + \epsilon_{ijk} \quad (16)$$

we observe that neither x_1^2 or x_2^2 columns of Table 8 are contrasts, i.e., they are not orthogonal to a 9×1 vector [1 1 1 … 1]′, and hence we make a slight change in model (16) so that we have complete orthogonality in our regression model. The 4/9 is simply the average of x_1^2 and x_2^2 columns so that the $x_1^2 - 4/9$ column in Table 8 is now a contrast. Our modified regression model now becomes

$$y_{ijk} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_{11}(x_1^2 - 4/9) + \in_{ijk} \quad (17)$$

As always, in order to solve for any of the coefficients in an orthogonal model, simply multiply both sides of Eq. (17) by that variable and sum over all i, j, and k.

For example, to solve for $\hat{\beta}_{11}$, multiply both sides of (17) by the column $x_1^2 - 4/9$ and sum over all the 9 rows of Table 8. This yields (5/9)161.7 – (4/9)202.3 $= \hat{\beta}_{11}\,[4(5/9)^2 + 5(-4/9)^2] \rightarrow \hat{\beta}_{11} = -0.035$, which would be the same as the regression output from Minitab. You should try to similarly obtain the other 4 coefficients in the model (17) and check against $\hat{\mathbf{y}}_{\mathbf{ijk}} = 40.4600 + 0.775x_1 + 0.325x_2 - 0.025x_1x_2 - 0.035x_1^2$. Note that in order to use Minitab to obtain the ANOVA for a design with center points, you must first set up Factorial by going to Stat → DOE → Factorial → Define custom factorial design; insert factor names, click on Low/High and then set up center points by clicking on Designs to tell Minitab that the design has center points, and then go to Analyze Factorial Design. Further, a column equal-length to the design matrix has to be created to represent the center points, where 1 is used for factorial and 0 for center.

OPTIMIZING PROCESS CONDITIONS USING STRONG EFFECTS AND REDUCING PROCESS VARIABILITY USING WEAK EFFECTS IN UNREPLICATED FACTORIALS (when there are at least two weak factors impacting the mean response and n = 1 per FLC)

Montgomery provides an excellent Example 6.4 on pages 271-274 (of 8th edition) of a 2^4 un-replicated factorial, where only the effects of factors A and C have a vital impact in reducing the mean of response y, where y is the number of defects per panel. I have put the data in the OA format in Table 9 below for your better understanding of the analysis that follows. On my website I have provided the Response Table (RT) and the corresponding Residual analysis from Minitab. The RT on my website will clearly verify the author's conclusion that only factors A and C have a vital impact on the mean response and their

Table 9. The OA Format of Montgomery's Example 6.4 (p. 245)

D	C	B	A	FLC	y_{ijkl}
-	-	-	-	(1)	5
-	-	-	+	a	11
-	-	+	-	b	3.5
-	-	+	+	ab	9
-	+	-	-	c	0.5
-	+	-	+	ac	8
-	+	+	-	bc	1.5
-	+	+	+	abc	9.5
+	-	-	-	d	6
+	-	-	+	ad	12.5
+	-	+	-	bd	8
+	-	+	+	abd	15.5
+	+	-	-	cd	1
+	+	-	+	acd	6
+	+	+	-	bcd	5
+	+	+	+	abcd	5%

optimal levels are $A^{-} = A_0$ and $C^{+} = C_1$. Factors B and D are weak in impacting the mean response, and therefore, their levels should be set in order to minimize the variability of the response. Both the Montgomery's and my analyses confirm that the low levels of B and D induce less variability in y. Hence, the final optimal condition is $X_0 = A_0B_0C_1D_0 = c$.

Exercise 20. Work problem 6.28 on page 298 of Montgomery's 8th edition W/O any plots but provide a RT. Further, identify the optimum process condition X_0.

Errata For Chapter 6 of Montgomery's 8th Edition

1. Page 260, near the bottom, change 46.22 to 46.25.
2. Page 263, line 9, change contrast variance to contrast standard error (SE).