STAT 222 Lecture 23
Single Replicate Two-Level Factorial Designs and Half Normal Probability Plots

Yibi Huang

Two-Level Factorial Designs (2^k Designs)

- *Two-level factorial designs (2^k designs)* are factorial designs in which each factor is investigated at only two levels.

Why using 2^k designs?

- The early stages of a study usually involve the investigation of a large number of potential factors to discover the "vital few" factors.
- The # of observations required by a full factorial design grows exponentially with the number of factors. E.g., it takes at least $2^k = 2^{12} = 4096$ observations to investigate $k = 12$ factors. If any of the 12 factors has 3 or more levels, it takes at least $3 \times 2^{11} = 6144$ observations for just a single replicate
- Hence, we can only afford *2 levels* for each factor, and just a *single replicate* for each factor combination
- Two level factorial experiments are often used during these stages to quickly filter out unwanted effects and identify the important ones

Notation of Two-Level Factorial Designs

As all the factors have 2 levels only, their levels are usually referred to as (low, high) and coded as

$$0 = \text{low}, \quad 1 = \text{high}.$$

E.g., the observations y_{ijk} of a 2^3 design are hence denoted as

$$y_{000},\ y_{001},\ y_{010},\ y_{011},\ y_{100},\ y_{101},\ y_{110},\ y_{111}.$$

Parameter Estimates of a 2^k Design

Parameter estimates for the full 3-way factorial model

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + \alpha\beta_{ij} + \beta\gamma_{jk} + \alpha\gamma_{ij} + \alpha\beta\gamma_{ijk} + \varepsilon_{ijk},$$

under the zero-sum constraints can be shown to be of the form $\sum_{ijk} c_{ijk} y_{ijk}/2^k$ where the coefficients c_{ijk} are as shown in the table below.

	$\widehat{\mu}$ (1)	$\widehat{\alpha}_1$ A	$\widehat{\beta}_1$ B	$\widehat{\gamma}_1$ C	$\widehat{\alpha\beta}_{11}$ AB	$\widehat{\alpha\gamma}_{11}$ AC	$\widehat{\beta\gamma}_{11}$ BC	$\widehat{\alpha\beta\gamma}_{111}$ ABC
000	1	−1	−1	−1	1	1	1	−1
001	1	−1	−1	1	1	−1	−1	1
010	1	−1	1	−1	−1	1	−1	1
011	1	−1	1	1	−1	−1	1	−1
100	1	1	−1	−1	−1	−1	1	1
101	1	1	−1	1	−1	1	−1	−1
110	1	1	1	−1	1	−1	−1	−1
111	1	1	1	1	1	1	1	1

	$\widehat{\mu}$ (1)	$\widehat{\alpha}_1$ A	$\widehat{\beta}_1$ B	$\widehat{\gamma}_1$ C	$\widehat{\alpha\beta}_{11}$ AB	$\widehat{\alpha\gamma}_{11}$ AC	$\widehat{\beta\gamma}_{11}$ BC	$\widehat{\alpha\beta\gamma}_{111}$ ABC
000	1	−1	−1	−1	1	1	1	−1
001	1	−1	−1	1	1	−1	−1	1
010	1	−1	1	−1	−1	1	−1	1
011	1	−1	1	1	−1	−1	1	−1
100	1	1	−1	−1	−1	−1	1	1
101	1	1	−1	1	−1	1	−1	−1
110	1	1	1	−1	1	−1	−1	−1
111	1	1	1	1	1	1	1	1

For example,

$$\begin{aligned}
\widehat{\mu} &= (y_{000} + y_{001} + y_{010} + y_{011} + y_{100} + y_{101} + y_{110} + y_{111})/8 \\
\widehat{\alpha}_1 &= (-y_{000} - y_{001} - y_{010} - y_{011} + y_{100} + y_{101} + y_{110} + y_{111})/8 \\
\widehat{\alpha\beta}_{11} &= (y_{000} + y_{001} - y_{010} - y_{011} - y_{100} - y_{101} + y_{110} + y_{111})/8
\end{aligned}$$

and so on.

	$\widehat{\mu}$ (1)	$\widehat{\alpha}_1$ A	$\widehat{\beta}_1$ B	$\widehat{\gamma}_1$ C	$\widehat{\alpha\beta}_{11}$ AB	$\widehat{\alpha\gamma}_{11}$ AC	$\widehat{\beta\gamma}_{11}$ BC	$\widehat{\alpha\beta\gamma}_{111}$ ABC
000	1	−1	−1	−1	1	1	1	−1
001	1	−1	−1	1	1	−1	−1	1
010	1	−1	1	−1	−1	1	−1	1
011	1	−1	1	1	−1	−1	1	−1
100	1	1	−1	−1	−1	−1	1	1
101	1	1	−1	1	−1	1	−1	−1
110	1	1	1	−1	1	−1	−1	−1
111	1	1	1	1	1	1	1	1

- Note only 1 parameter for each main effect or interaction. Parameter at the levels can be determined using the zero-sum constraints
- Except for the grand mean, all other estimates are *contrasts* as $\sum_{ijk} c_{ijk} = 0$
- Hence, we often just refer to the estimates as contrasts and denote them as

$$A, B, C, AB, AC, BC, ABC.$$

	A	B	C	AB	AC	BC	ABC
000	-1	-1	-1	1	1	1	-1
001	-1	-1	1	1	-1	-1	1
010	-1	1	-1	-1	1	-1	1
011	-1	1	1	-1	-1	1	-1
100	1	-1	-1	-1	-1	1	1
101	1	-1	1	-1	1	-1	-1
110	1	1	-1	1	-1	-1	-1
111	1	1	1	1	1	1	1

Observe the coefficients of the A, B, C, contrasts are

$$c_{ijk}^A = \begin{cases} -1 & \text{if } i = 0 \\ 1 & \text{if } i = 1 \end{cases}, \quad c_{ijk}^B = \begin{cases} -1 & \text{if } j = 0 \\ 1 & \text{if } j = 1 \end{cases}, \quad c_{ijk}^C = \begin{cases} -1 & \text{if } k = 0 \\ 1 & \text{if } k = 1 \end{cases}.$$

In other words, for a main effect contrast of a factor

- $c_{ijk} = 1$ if the factor is at the *high* level
- $c_{ijk} = -1$ if the factor is at the *low* level

	A	B	C	AB	AC	BC	ABC
000	-1	-1	-1	1	1	1	-1
001	-1	-1	1	1	-1	-1	1
010	-1	1	-1	-1	1	-1	1
011	-1	1	1	-1	-1	1	-1
100	1	-1	-1	-1	-1	1	1
101	1	-1	1	-1	1	-1	-1
110	1	1	-1	1	-1	-1	-1
111	1	1	1	1	1	1	1

For the interaction contrasts, their coefficients c_{ijk} are just the products of the coefficients of the main effect contrasts of corresponding factors.

For example,

- $c_{ijk}^{AB} = c_{ijk}^{A} c_{ijk}^{B}$
- $c_{ijk}^{AC} = c_{ijk}^{A} c_{ijk}^{C}$
- $c_{ijk}^{ABC} = c_{ijk}^{A} c_{ijk}^{B} c_{ijk}^{C}$

	A	B	C	AB	AC	BC	ABC
000	-1	-1	-1	1	1	1	-1
001	-1	-1	1	1	-1	-1	1
010	-1	1	-1	-1	1	-1	1
011	-1	1	1	-1	-1	1	-1
100	1	-1	-1	-1	-1	1	1
101	1	-1	1	-1	1	-1	-1
110	1	1	-1	1	-1	-1	-1
111	1	1	1	1	1	1	1

For the interaction contrasts, their coefficients c_{ijk} are just the products of the coefficients of the main effect contrasts of corresponding factors.

For example,

- $c_{ijk}^{AB} = c_{ijk}^{A} c_{ijk}^{B}$
- $c_{ijk}^{AC} = c_{ijk}^{A} c_{ijk}^{C}$
- $c_{ijk}^{ABC} = c_{ijk}^{A} c_{ijk}^{B} c_{ijk}^{C}$

	A	B	C	AB	AC	BC	ABC
000	−1	−1	−1	1	1	1	−1
001	−1	−1	1	1	−1	−1	1
010	−1	1	−1	−1	1	−1	1
011	−1	1	1	−1	−1	1	−1
100	1	−1	−1	−1	−1	1	1
101	1	−1	1	−1	1	−1	−1
110	1	1	−1	1	−1	−1	−1
111	1	1	1	1	1	1	1

For the interaction contrasts, their coefficients c_{ijk} are just the products of the coefficients of the main effect contrasts of corresponding factors.

For example,

- $c_{ijk}^{AB} = c_{ijk}^{A} c_{ijk}^{B}$
- $c_{ijk}^{AC} = c_{ijk}^{A} c_{ijk}^{C}$
- $c_{ijk}^{ABC} = c_{ijk}^{A} c_{ijk}^{B} c_{ijk}^{C}$

We just showed the parameter estimates for 2^3 designs, the parameter estimates for a general 2^k designs are in the form

$$c_{ijk\ldots} y_{ijk\ldots} / 2^k$$

where

- the coefficients $c_{ijk\ldots}$ for the main effect of a factor is 1 if the factor is at the high level and -1 if the factor is at the low level
- the coefficients c_{ijk} of an interaction are just the products of the coefficients of the main effect contrasts of corresponding factors,

e.g., for a 2^4 design

- $c^{ABD}_{ijk\ell} = c^{A}_{ijk\ell} c^{B}_{ijk\ell} c^{D}_{ijk\ell}$
- $c^{ABCD}_{ijk\ell} = c^{A}_{ijk\ell} c^{B}_{ijk\ell} c^{C}_{ijk\ell} c^{D}_{ijk\ell}$

Contrasts in a 2^k Design Are Orthogonal

We said two contrasts $C_1 = \sum_i c_i^{(1)} \mu_i$ and $C_2 = \sum_i c_i^{(2)} \mu_i$ are *orthogonal* to each other if

$$\sum_i c_i^{(1)} c_i^{(2)} = 0.$$

Observe the 7 contrasts on the right in a 2^3 design are *orthogonal* to each other.

	A	B	C	AB	AC	BC	ABC
000	−1	−1	−1	1	1	1	−1
001	−1	−1	1	1	−1	−1	1
010	−1	1	−1	−1	1	−1	1
011	−1	1	1	−1	−1	1	−1
100	1	−1	−1	−1	−1	1	1
101	1	−1	1	−1	1	−1	−1
110	1	1	−1	1	−1	−1	−1
111	1	1	1	1	1	1	1

In general, the main effect and interaction contrasts below for a 2^k design are orthogonal to each other.

$$A, B, C, \ldots, AB, AC, \ldots, ABC, ABD, \ldots, ABCD, \ldots.$$

Contrasts in a 2^k Design Are Uncorrelated w/ Each Other

As y_{ijk}'s are independent with equal variance σ^2, for any two contrasts U, V of the 7 contrasts on the right, their covariance is

	A	B	C	AB	AC	BC	ABC
000	−1	−1	−1	1	1	1	−1
001	−1	−1	1	1	−1	−1	1
010	−1	1	−1	−1	1	−1	1
011	−1	1	1	−1	−1	1	−1
100	1	−1	−1	−1	−1	1	1
101	1	−1	1	−1	1	−1	−1
110	1	1	−1	1	−1	−1	−1
111	1	1	1	1	1	1	1

$$\begin{aligned}
\mathrm{Cov}(U,V) &= \mathrm{Cov}\left(\sum\nolimits_{ijk} c^U_{ijk} y_{ijk}, \sum\nolimits_{i'j'k'} c^V_{i'j'k'} y_{i'j'k'}\right) \\
&= \sum_{ijk} c^U_{ijk} c^V_{ijk} \underbrace{\mathrm{Var}(y_{ijk})}_{=\sigma^2} + \sum_{ijk} \sum_{i'j'k'} c^U_{ijk} c^V_{i'j'k'} \underbrace{\mathrm{Cov}(y_{ijk}, y_{i'j'k'})}_{=0 \text{ by indep.}} \\
&= \sigma^2 \underbrace{\sum\nolimits_{ijk} c^U_{ijk} c^V_{ijk}}_{=0} = 0 \quad \text{since } U, V \text{ are orthogonal}
\end{aligned}$$

The same argument applies to other 2^k designs in general.

Contrasts in a 2^k Design Have an Identical Variance

As y_{ijk}'s are independent with a constant variance σ^2, all of the 7 contrasts in a 2^3 design on the right have an identical variance $2^3\sigma^2$ since

	A	B	C	AB	AC	BC	ABC
000	−1	−1	−1	1	1	1	−1
001	−1	−1	1	1	−1	−1	1
010	−1	1	−1	−1	1	−1	1
011	−1	1	1	−1	−1	1	−1
100	1	−1	−1	−1	−1	1	1
101	1	−1	1	−1	1	−1	−1
110	1	1	−1	1	−1	−1	−1
111	1	1	1	1	1	1	1

$$\mathrm{Var}\left(\sum_{ijk} c_{ijk} y_{ijk}\right)$$

$$= \sum_{ijk} \underbrace{c_{ijk}^2}_{=1} \underbrace{\mathrm{Var}(y_{ijk})}_{=\sigma^2} = \sum_{ijk} \sigma^2 = 2^3\sigma^2,$$

where $c_{ijk}^2 = 1$ since all c_{ijk}'s are 1 or -1.

Contrasts in a 2^k Design Have an Identical Variance

As y_{ijk}'s are independent with a constant variance σ^2, all of the 7 contrasts in a 2^3 design on the right have an identical variance $2^3\sigma^2$ since

	A	B	C	AB	AC	BC	ABC
000	−1	−1	−1	1	1	1	−1
001	−1	−1	1	1	−1	−1	1
010	−1	1	−1	−1	1	−1	1
011	−1	1	1	−1	−1	1	−1
100	1	−1	−1	−1	−1	1	1
101	1	−1	1	−1	1	−1	−1
110	1	1	−1	1	−1	−1	−1
111	1	1	1	1	1	1	1

$$\mathrm{Var}\left(\sum_{ijk} c_{ijk} y_{ijk}\right)$$

$$= \sum_{ijk} \underbrace{c_{ijk}^2}_{=1} \underbrace{\mathrm{Var}(y_{ijk})}_{=\sigma^2} = \sum_{ijk} \sigma^2 = 2^3\sigma^2,$$

where $c_{ijk}^2 = 1$ since all c_{ijk}'s are 1 or -1.

- Parameter estimates for the full model (under the zero-sum constraints) of a 2^3 design also have an identical variance $\sigma^2/2^3$ since they are just the contrasts above divided by 2^3

Contrasts in a 2^k Design Have an Identical Variance

As y_{ijk}'s are independent with a constant variance σ^2, all of the 7 contrasts in a 2^3 design on the right have an identical variance $2^3\sigma^2$ since

	A	B	C	AB	AC	BC	ABC
000	−1	−1	−1	1	1	1	−1
001	−1	−1	1	1	−1	−1	1
010	−1	1	−1	−1	1	−1	1
011	−1	1	1	−1	−1	1	−1
100	1	−1	−1	−1	−1	1	1
101	1	−1	1	−1	1	−1	−1
110	1	1	−1	1	−1	−1	−1
111	1	1	1	1	1	1	1

$$\mathrm{Var}\left(\sum_{ijk} c_{ijk} y_{ijk}\right)$$

$$= \sum_{ijk} \underbrace{c_{ijk}^2}_{=1} \underbrace{\mathrm{Var}(y_{ijk})}_{=\sigma^2} = \sum_{ijk} \sigma^2 = 2^3\sigma^2,$$

where $c_{ijk}^2 = 1$ since all c_{ijk}'s are 1 or -1.

- Parameter estimates for the full model (under the zero-sum constraints) of a 2^3 design also have an identical variance $\sigma^2/2^3$ since they are just the contrasts above divided by 2^3
- Parameter estimates for the full model of a 2^k design also have an identical variance $\sigma^2/2^k$.

Properties of Parameter Estimates of a 2^k Design

Under the zero-sum constraints, parameter estimates of the full main-effect-interaction model of a 2^k design

1. are unbiased estimates of their corresponding parameters;
2. have an *identical variance* $\sigma^2/2^k$;
3. are **uncorrelated** and hence *independent* of each other
   - Recall if normally distributed, zero correlation implies independence
4. are *normally distributed* since they are linear combinations of y's, which are independent normal

Single-Replicate Data (Review)

Recall in L17, we said the MSE of a full k-way model is 0 if there is only a *single replicate*.

- cannot test the significance of all main effects and interactions of all order under the full model
- can test the significance of the main effects and some lower-order interactions by *pooling higher order interactions into error* and get a non-zero MSE.
- However, we are not able to test the significance of terms that are pooled into errors

Half normal probability plot is a tool that one can examine all main effects and interactions altogether and identify non-negligible ones.

How Do Half Normal Probability Plots Work?

- Under the zero-sum constraint, recall that parameter estimates of a full model of a 2^k are *independent* and *normally* distributed with *constant variance*.
- The expected value of any of these contrasts is 0 if the corresponding parameter (main effect or interaction) is 0.
- So, estimates corresponding to zero effects would look like a sample from $N(0, \sigma^2/2^k)$, and estimates corresponding to significant effects looks outliers
- **Sparsity Assumption**: most parameters are 0, only a few are non-zero
- A half-normal probability plot plots the sorted absolute values of the estimates on the vertical axis against the sorted expected scores from a half-normal distribution.

Example 7.5.1 Drill Advance Experiment (p.220 Dean & Voss)

A $2 \times 2 \times 2 \times 2$ experiment to study the effects of 4 factors on the rate of advance of a small stone drill.

- A: load on the drill
- B: flow rate through the drill
- C: speed of rotation
- D: type of mud used in drilling

Each factor was observed at two levels, coded 1 and 2.

Response $= \log_{10}(\text{Advance})$

```
drill = read.table(
          "http://www.stat.uchicago.edu/~yibi/s222/drill.txt", h=T)
drill$A = as.factor(drill$A)
drill$B = as.factor(drill$B)
drill$C = as.factor(drill$C)
drill$D = as.factor(drill$D)
contrasts(drill$A) = contr.sum(2)
contrasts(drill$B) = contr.sum(2)
contrasts(drill$C) = contr.sum(2)
contrasts(drill$D) = contr.sum(2)
```

Fitting a full 4-way model, the SE's of all coefficients are NaN (Not a Number) since SSE $= 0$

```
summary(lm(log10(Advance) ~ A*B*C*D, data=drill))$coef
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.693885        NaN     NaN      NaN
A1           -0.028227        NaN     NaN      NaN
B1           -0.125963        NaN     NaN      NaN
C1           -0.250686        NaN     NaN      NaN
D1           -0.070908        NaN     NaN      NaN
A1:B1        -0.007462        NaN     NaN      NaN
A1:C1         0.002248        NaN     NaN      NaN
B1:C1        -0.010902        NaN     NaN      NaN
A1:D1         0.014527        NaN     NaN      NaN
B1:D1        -0.003244        NaN     NaN      NaN
C1:D1         0.021311        NaN     NaN      NaN
A1:B1:C1     -0.002253        NaN     NaN      NaN
A1:B1:D1     -0.011339        NaN     NaN      NaN
A1:C1:D1     -0.011558        NaN     NaN      NaN
B1:C1:D1      0.007494        NaN     NaN      NaN
A1:B1:C1:D1   0.008385        NaN     NaN      NaN
```

Pooling 4-way interaction into error, we can get a non-zero SSE and calculate the SE for each remaining parameter.

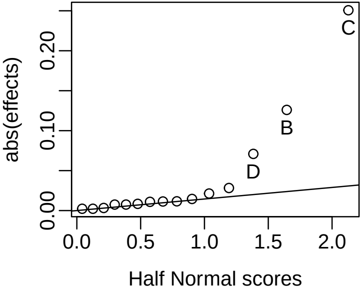
Observe parameters under the zero-sum constraints all have the same SE since they all have an identical variance.

```
summary(lm(log10(Advance) ~ (A+B+C+D)^3, data=drill))$coef
             Estimate Std. Error  t value Pr(>|t|)
(Intercept)  0.693885   0.008385  82.7495 0.007693
A1          -0.028227   0.008385  -3.3662 0.183833
B1          -0.125963   0.008385 -15.0218 0.042317
C1          -0.250686   0.008385 -29.8957 0.021287
D1          -0.070908   0.008385  -8.4562 0.074937
A1:B1       -0.007462   0.008385  -0.8899 0.537056
A1:C1        0.002248   0.008385   0.2681 0.833273
A1:D1        0.014527   0.008385   1.7325 0.333268
B1:C1       -0.010902   0.008385  -1.3001 0.417406
B1:D1       -0.003244   0.008385  -0.3868 0.765012
C1:D1        0.021311   0.008385   2.5415 0.238649
A1:B1:C1    -0.002253   0.008385  -0.2686 0.832926
A1:B1:D1    -0.011339   0.008385  -1.3522 0.405380
A1:C1:D1    -0.011558   0.008385  -1.3784 0.399565
B1:C1:D1     0.007494   0.008385   0.8937 0.535685
```

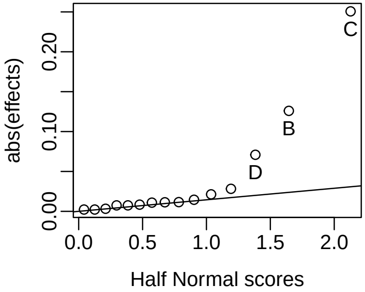
Half Normal Probability Plot in R (using `daewr`)

1. Fit the model and get the parameter estimates under the zero-sum constraint.
2. Exclude the intercept since we don't expect it to be zero
3. Make half-normal probability plot based on the remaining estimates using the `halfnom()` function in the `daewr` library ("daewr" = the book *Design and Analysis of Experiments with R*).

```
library(daewr)
model1 = lm(log10(Advance) ~ A*B*C*D, data=drill)
halfnorm(model1$coef[-1])
```



From the half-normal probability plot, we see estimates B, C, D main effects are outliers compared to other negligible coefficients, consistent with the ANOVA table below obtained by pooling all interactions into errors.




```
zscore= 0.04179 0.1257 0.2104 0.2967 0.3853 0.477 0.573 0.6745 0.7835 0
```

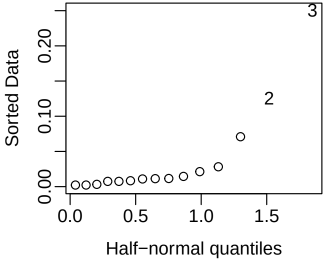
```
anova(lm(log10(Advance) ~ A+B+C+D, data=drill))
Analysis of Variance Table

Response: log10(Advance)
          Df Sum Sq Mean Sq F value  Pr(>F)
A          1  0.013   0.013    7.02   0.023
B          1  0.254   0.254  139.74 1.4e-07
C          1  1.005   1.005  553.46 9.3e-11
D          1  0.080   0.080   44.28 3.6e-05
Residuals 11  0.020   0.002
```

Half Normal Plot Using the faraway Library

If having trouble installing the daewr library, one can use the halfnorm plot in the faraway library[1] instead, which, by default, labels the effects by 1, 2, 3,. . . , rather than A, B, AB,. . . , and the straight line is NOT included.

```
library(faraway)
par(mai=c(.6,.6,.05,.05),mgp=c(2,.5,0))
halfnorm(model1$coef[-1])
```

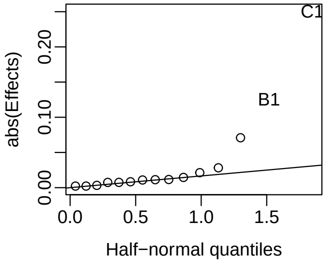


[1]"faraway" = author of the book *Linear Models with R*. This solution is suggested by Antonio Fernandes. Thanks, Antonio!

Half Normal Plot Using the faraway Library (2)

Nonetheless, we can change the effect labels by adding `labs= names(model1$coef[-1])` within `halfnorm()` and add the straight line using `qqline()` ourselves.

```
library(faraway)
par(mai=c(.6,.6,.05,.05),mgp=c(2,.5,0))
halfnorm(model1$coef[-1], labs= names(model1$coef[-1]),
         ylab= "abs(Effects)")
qqline(c(-abs(model1$coef[-1]),abs(model1$coef[-1])))
```



Pros and Cons of Half-Normal Probability Plot

Pros:

- Can check all main effects and interactions of all orders all at once

Cons:

- no p-values are provided. Identification of “significant” effects can be subjective
- doesn’t work well if the **sparsity** assumption is not met(most effects are zero, only a few are non-zero) as we need a sufficient number of null effects to estimate the unknown variance and identify the outliers