Week 10: *DOE & Power Analysis*

EMSE 6035: Marketing Analytics for Design Decisions

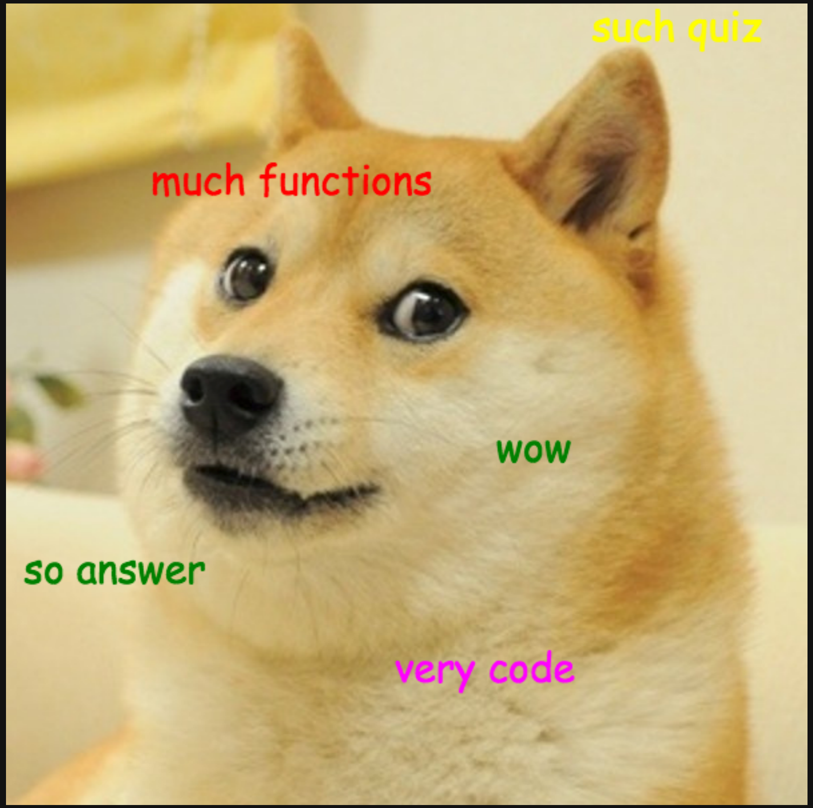
John Paul Helveston

November 03, 2021

Quiz 4

Make sure to download the zip file on the first page!

10:00



Week 10: *DOE & Power Analysis*

1. Design of Experiment

BREAK

2. Design Efficiency

3. Power Analysis

4. Interactions

Week 10: *DOE & Power Analysis*

1. Design of Experiment

BREAK

2. Design Efficiency

3. Power Analysis

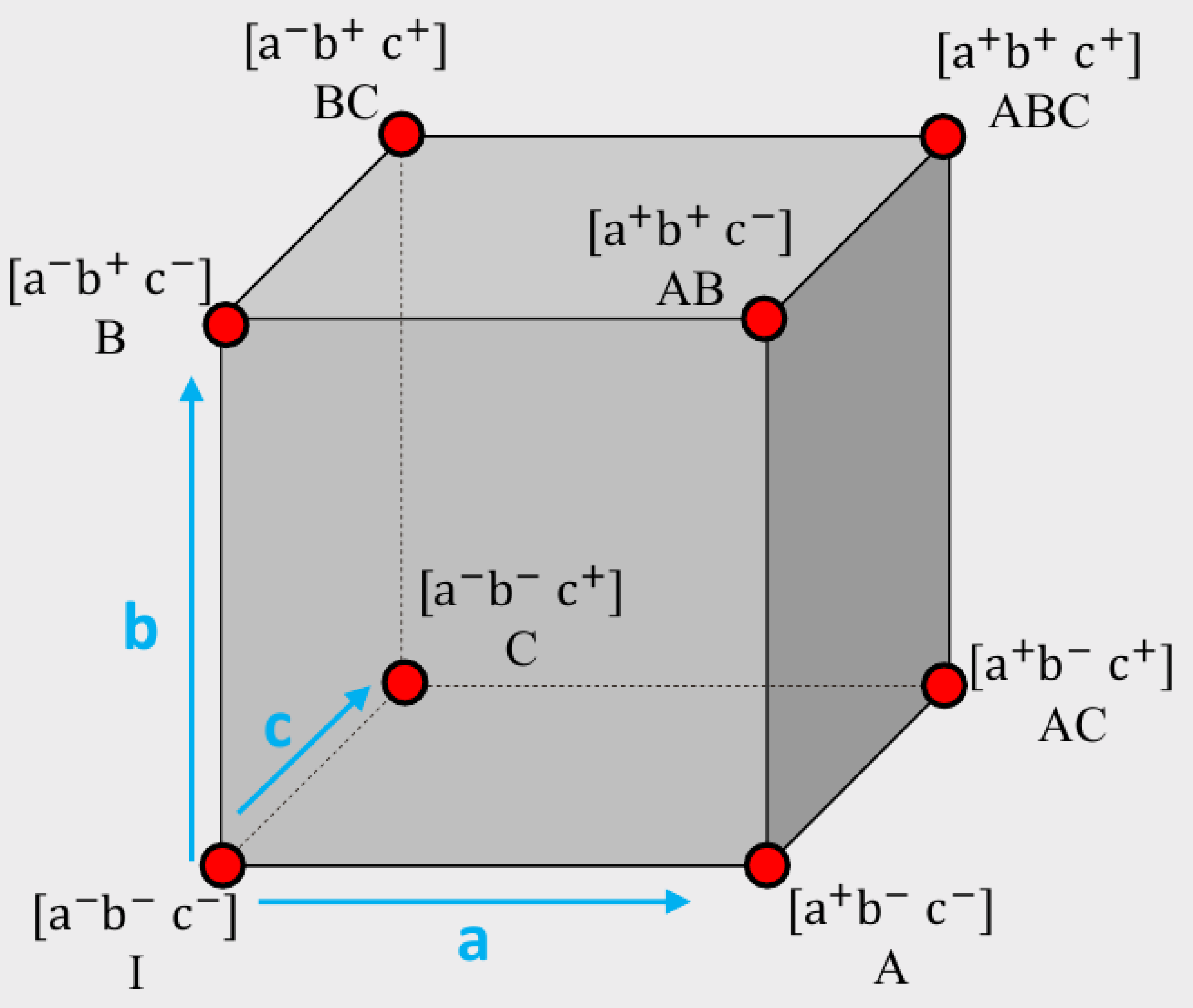
4. Interactions

Before we start, re-install {conjointTools}

```r
remotes::install_github("jhelvy/conjointTools")
```

Main & Interaction Effects

Full design space for 3 effects: A, B, C



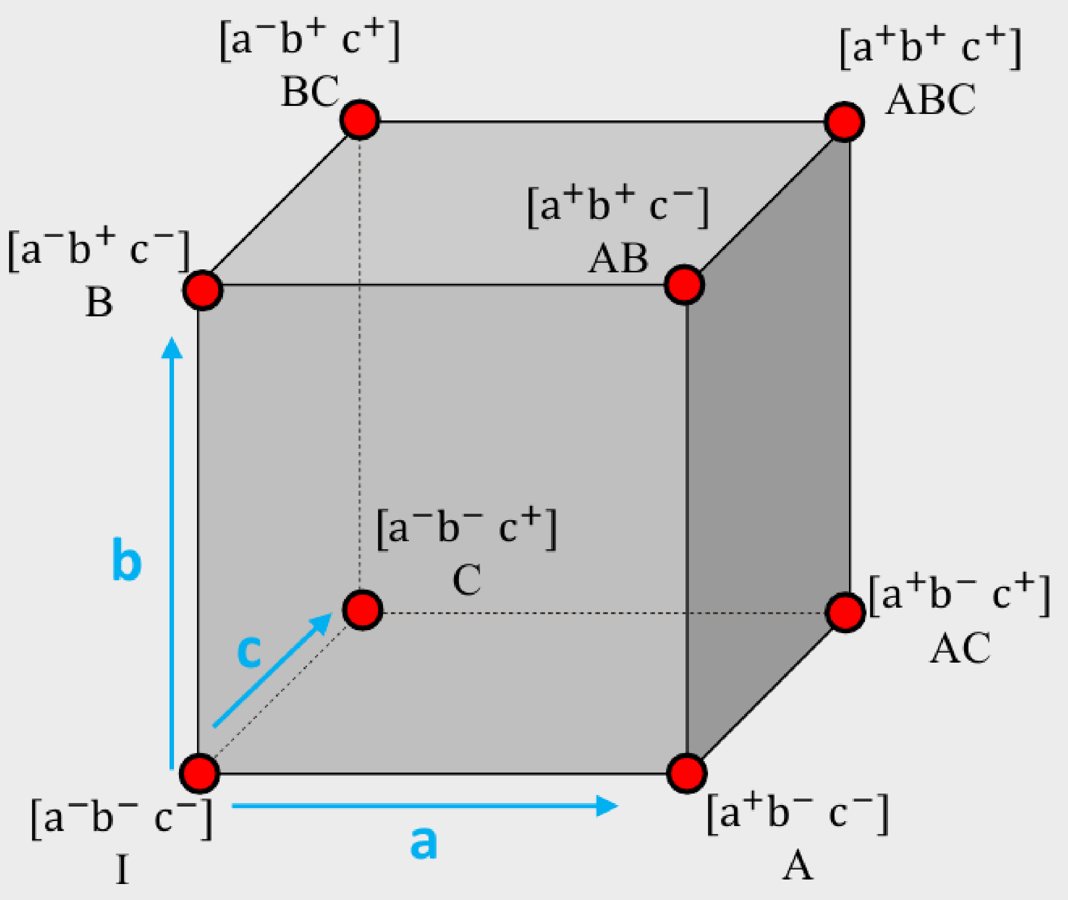
Full design space for 3 effects: A, B, C

Example: *Cars*

A: Electric? (Yes+ or No-)

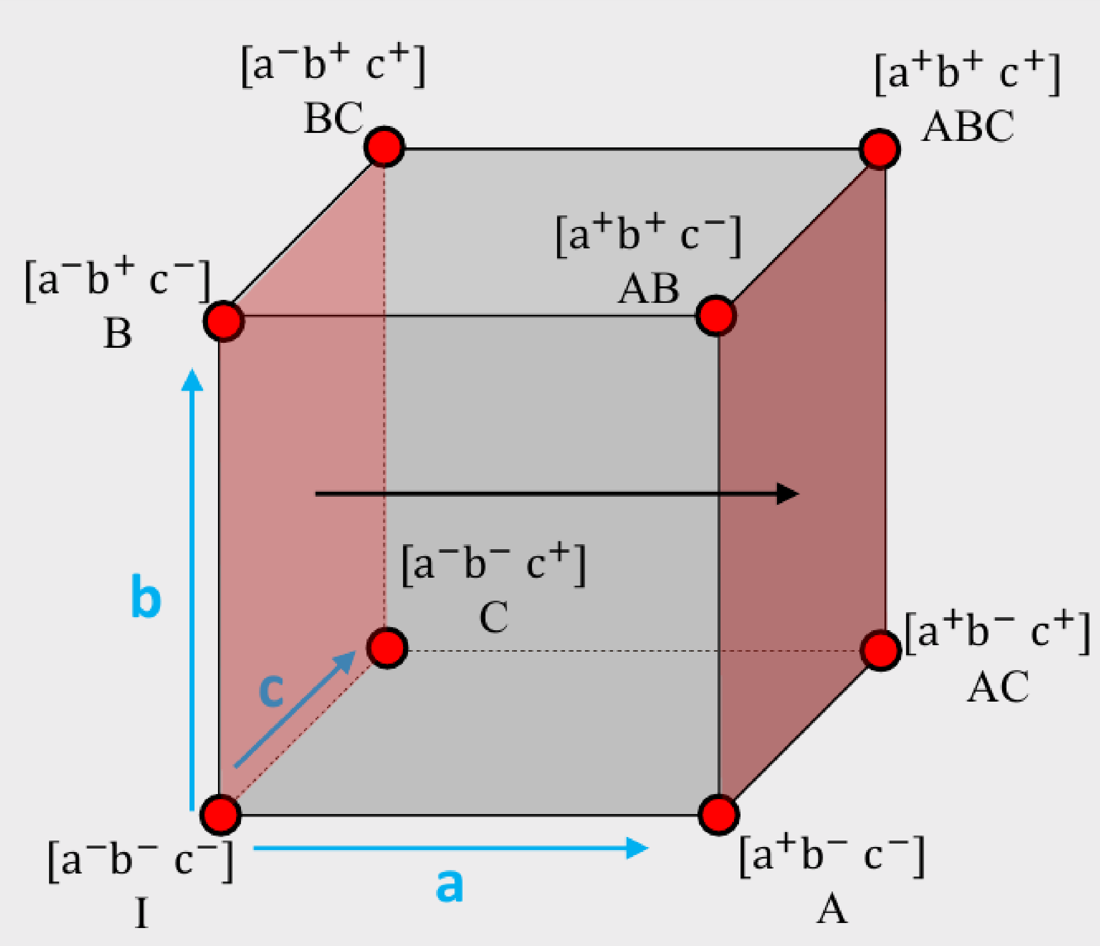
B: Warranty? (Yes+ or No-)

C: Ford? (Yes+ or No-)



Main Effects

$$ME(a) = \left(\frac{A + AB + AC + ABC}{4}\right) - \left(\frac{I + B + C + BC}{4}\right)$$

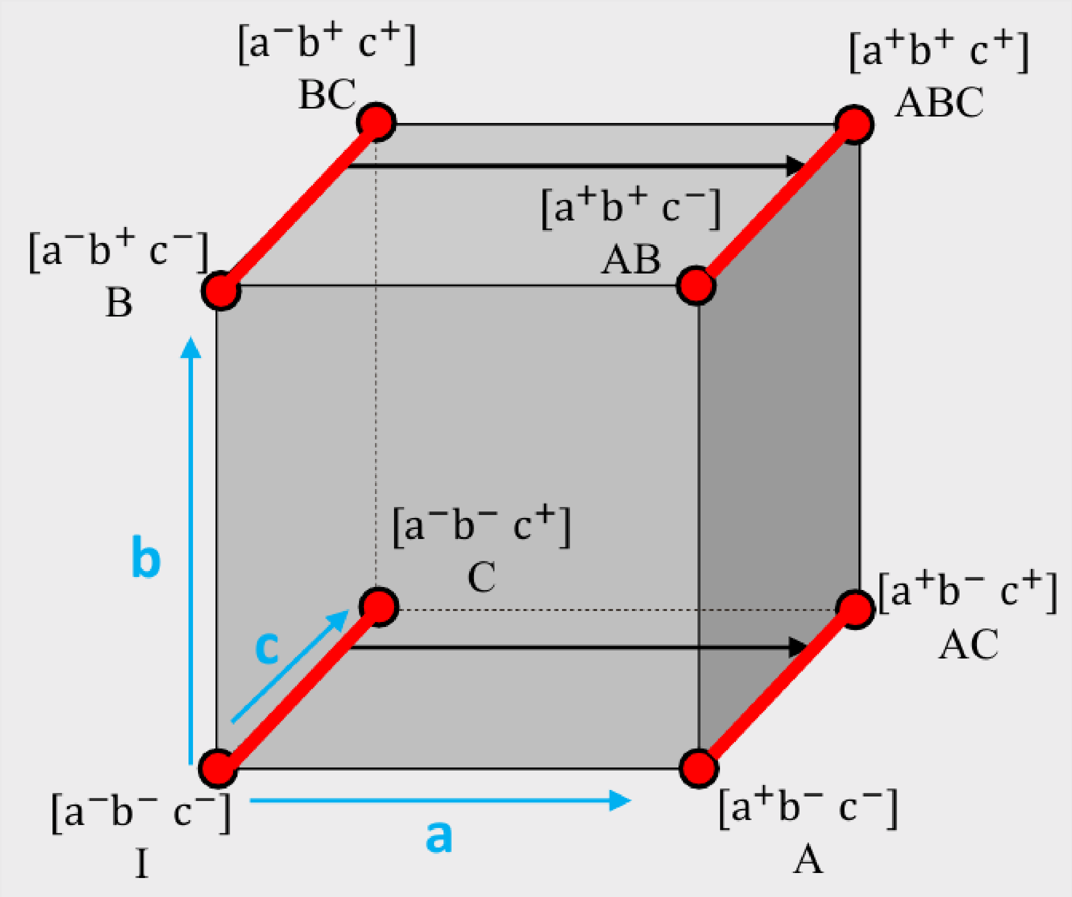


Interaction Effects

$$INT(ab) =$$

$$\frac{1}{2}\left[\left(\frac{AB+ABC}{2}\right) - \left(\frac{B+BC}{2}\right)\right] -$$

$$\frac{1}{2}\left[\left(\frac{A+AC}{2}\right) - \left(\frac{I+C}{2}\right)\right]$$



Example: Wine Pairings

meat	wine
fish	white
fish	red
steak	white
steak	red

Main Effects

1. **Fish** or **Steak**?
2. **Red** or **White** wine?

Interaction Effects

1. **Red** or **White** wine *with **Steak***?
2. **Red** or **White** wine *with **Fish***?

Open `winePairings.Rmd`

Fractional vs Full Factorial Designs

Full Factorial Design

Example: *Cars*

A: Electric? (Yes+ or No-)

B: Warranty? (Yes+ or No-)

C: Ford? (Yes+ or No-)

```
library(conjointTools)

levels <- list(
    electric = c(1, 0),
    warranty = c(1, 0),
    ford     = c(1, 0)
)

doe <- makeDoe(levels)
recodeDoe(doe, levels)
```

```
#>   electric warranty ford
#> 1        1        1    1
#> 2        0        1    1
#> 3        1        0    1
#> 4        0        0    1
#> 5        1        1    0
#> 6        0        1    0
#> 7        1        0    0
#> 8        0        0    0
```

Full Factorial Design

Balanced?

All levels appear an equal number of times.

Orthogonal?

All pairs of levels appear together an equal number of times.

```r
library(conjointTools)

levels <- list(
    electric = c(1, 0),
    warranty = c(1, 0),
    ford     = c(1, 0)
)

doe <- makeDoe(levels)
doe <- recodeDoe(doe, levels)
doe
```

```
#>   electric warranty ford
#> 1        1        1    1
#> 2        0        1    1
#> 3        1        0    1
#> 4        0        0    1
#> 5        1        1    0
#> 6        0        1    0
#> 7        1        0    0
#> 8        0        0    0
```

Fractional Factorial Design

Balanced?

All levels appear an equal number of times.

Orthogonal?

All pairs of levels appear together an equal number of times.

```r
doe[c(1, 3, 5, 6),]
```

```
#>   electric warranty ford
#> 1        1        1    1
#> 3        1        0    1
#> 5        1        1    0
#> 6        0        1    0
```

Comparing Full and Fractional Factorial Designs

Open `cars.Rmd`

Practice Question 1

Consider the following experiment design

a	b	c	Effect
+	-	-	A
-	+	-	B
+	-	+	AC
-	+	+	BC

a) Is the design balanced? Is is orthogonal?

b) Write out the equation to compute the main effect for a, b, and c.

c) Are any main effects confounded? If so, what are they confounded with?

Break

05:00

Week 10: *DOE & Power Analysis*

1. Design of Experiment

BREAK

2. Design Efficiency

3. Power Analysis

4. Interactions

We want to find β by maximizing the log-likelihood

$$\tilde{u}_j = \boldsymbol{\beta}'\mathbf{x}_j + \tilde{\varepsilon}_j$$
$$= \beta_1 x_{j1} + \beta_2 x_{j2} + \ldots + \tilde{\varepsilon}_j$$

Weights that denote the *relative* value of attributes $x_{j1}, x_{j2}, \ldots$

Estimate β_1 , β_2 , ... , by minimizing the negative log-likelihood function:

$$\text{minimize} - \ln(\mathcal{L}) = -\sum_{j=1}^{J} y_j \ln[P_j(\boldsymbol{\beta}|\mathbf{x})]$$

with respect to $\boldsymbol{\beta}$

$y_j = 1$ if alternative j was chosen
$y_j = 0$ if alternative j was not chosen

For logit model:

$$P_j = \frac{e^{v_j}}{\sum_{k=1}^{J} e^{v_k}} = \frac{e^{\boldsymbol{\beta}'\mathbf{x}_j}}{\sum_{k=1}^{J} e^{\boldsymbol{\beta}'\mathbf{x}_k}}$$

Covariance of β inversely related to matrix of 2nd derivatives

$$\sum\nolimits_{\boldsymbol{\beta}} = -\left[\overbrace{\nabla^2_{\boldsymbol{\beta}} \ln(\mathcal{L})}^{\text{Hessian}}\right]^{-1}$$

Covariance of $\widehat{\boldsymbol{\beta}}$

Negative of the hessian evaluated at the MLE solution is the **"Observed Information Matrix"**

$$\boldsymbol{I}(\boldsymbol{\beta}) = -\nabla^2_{\boldsymbol{\beta}} \ln(\mathcal{L})$$

"D-optimal" designs attempt to maximize the "D-efficiency" of a design

$$D = \left(\frac{|\boldsymbol{I}(\boldsymbol{\beta})|}{n^p} \right)^{1/p}$$

where p is the number of coefficients in the model and n is the total sample size

D ranges from 0 to 1

Designs are *more* orthogonal as D --> 1

Finding Efficient Designs

Open `efficiency.Rmd`

Your Turn

20:00

1. Individually, create a fractional factorial design of experiment for your team project. Are you able to identify a high D-efficient design with fewer trials than a full factorial design. Can you find a *balanced* design that is also efficient?

2. Compare your results with your teammates.

3. As a team, consider whether there are any restrictions you should make on your design and examine the impact (if any) those restrictions have on your design efficiency.

Week 10: *DOE & Power Analysis*

1. Design of Experiment

BREAK

2. Design Efficiency

3. Power Analysis

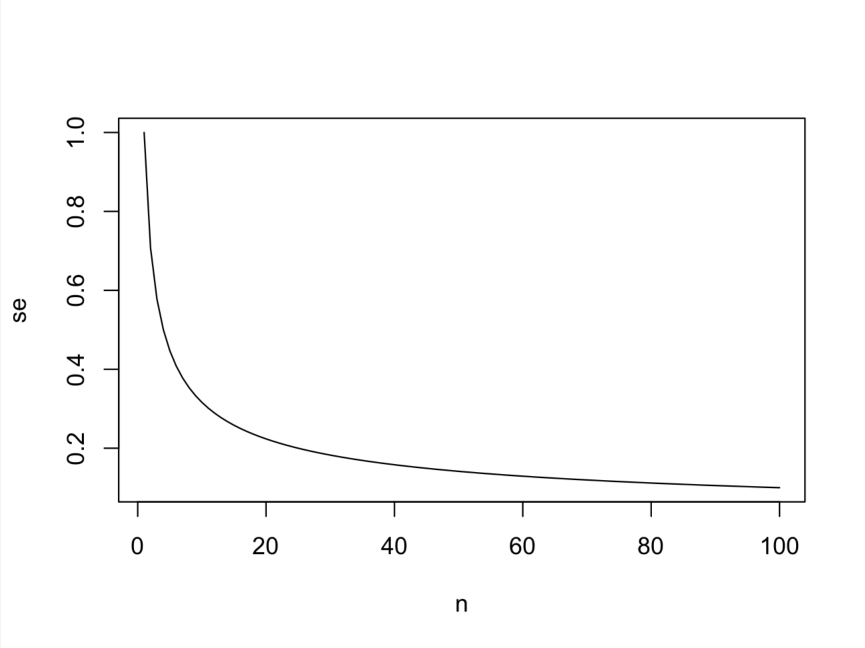
4. Interactions

How many respondents do I need?

How many respondents do I need
to get X level of precision on β?

Standard errors are inversely related to $\sqrt{N}$

```r
n <- seq(100)
se <- 1/sqrt(n)
plot(n, se, type = "l")
```



Standard errors also decrease with:

- Fewer attributes
- Fewer levels in each categorical attribute
- More questions per respondent

Using {conjointTools}, we can run simulations to determine the necessary sample size for a specific model

Open `powerAnalysis.Rmd`

Your Turn

20:00

Individually:

1. Using your design of experiment you just created in the last practice, conduct a power analysis to determine the necessary sample size to achieve a 0.05 significance level on your parameter estimates.

2. Compare your results with your teammates.

Week 10: *DOE & Power Analysis*

1. Design of Experiment

BREAK

2. Design Efficiency
3. Power Analysis
4. Interactions

Open
`powerAnalysis_interactions.Rmd`