Homework-9

Question 12.1

Describe a situation or problem from your job, everyday life, current events, etc., for which a design of experiments approach would be appropriate.

In Risk and Compliance, Assurance audits happen across all lines of business. Methodology used by auditing teams often slightly defer based on individual experiences in the field or level of relationship with the stakeholders. In this circumstance, businesses are finding audit reviews troublesome and does not provide enough value to the company. My team was tasked to identify factors that undermine the perceived value or actual value of audit reviews and at the end of preliminary data collection, we suggested a quasi experimentation on new ways of working aiming to address issues identified. In the experiment design, different review teams were given nudge packs on slightly different ways of engagement with the stakeholders designed to address those issues. Data would then be collected through surveys longitudinally from all parties involved for further analysis on effectiveness.

Question 12.2

To determine the value of 10 different yes/no features to the market value of a house (large yard, solar roof, etc.), a real estate agent plans to survey 50 potential buyers, showing a fictitious house with different combinations of features. To reduce the survey size, the agent wants to show just 16 fictitious houses. Use R's FrF2 function (in the FrF2 package) to find a fractional factorial design for this experiment: what set of features should each of the 16 fictitious houses have? Note: the output of FrF2 is "1" (include) or "-1" (don't include) for each feature.

```
rm(list=ls())

library(FrF2)
```

```
## Warning: package 'FrF2' was built under R version 4.0.4
```

```
## Loading required package: DoE.base
```

```
## Warning: package 'DoE.base' was built under R version 4.0.4
```

```
## Loading required package: grid
```

```
## Loading required package: conf.design
```

```
## Registered S3 method overwritten by 'DoE.base':
##   method           from
##   factorize.factor conf.design
```

```
##
## Attaching package: 'DoE.base'
```

```
## The following objects are masked from 'package:stats':
##
##     aov, lm
```

```
## The following object is masked from 'package:graphics':
##
##     plot.design
```

```
## The following object is masked from 'package:base':
##
##     lengths
```

```
set.seed(42069)

design = FrF2(nruns = 16, nfactors = 10)
design
```

```
##     A  B  C  D  E  F  G  H  J  K
## 1   1 -1  1  1 -1  1 -1  1 -1 -1
## 2   1  1  1  1  1  1  1  1  1  1
## 3   1 -1 -1 -1 -1 -1  1 -1 -1 -1
## 4   1 -1 -1  1 -1 -1  1  1  1  1
## 5  -1 -1  1  1  1 -1 -1 -1 -1  1
## 6  -1 -1 -1  1  1  1  1 -1  1 -1
## 7   1  1 -1 -1  1 -1 -1 -1  1  1
## 8  -1 -1 -1 -1  1  1  1  1 -1  1
## 9  -1 -1  1 -1  1 -1 -1  1  1 -1
## 10  1  1 -1  1  1 -1 -1  1 -1 -1
## 11  1 -1  1 -1 -1  1 -1 -1  1  1
## 12 -1  1 -1 -1 -1  1 -1  1  1 -1
## 13 -1  1  1 -1 -1 -1  1  1 -1  1
## 14  1  1  1 -1  1  1  1 -1 -1 -1
## 15 -1  1 -1  1 -1  1 -1 -1 -1  1
## 16 -1  1  1  1 -1 -1  1 -1  1 -1
## class=design, type= FrF2
```

Question 13.1

For each of the following distributions, give an example of data that you would expect to follow this distribution (besides the examples already discussed in class).

a. Binomial

Number of loan defaults across a group of customers.

b. Geometric

Number of successful repaid loans until first default across a group of customers.

c. Poisson

Rate of customers arriving to a bank during peak hours.

d. Exponential

Interval timings between each customer's arrival to a bank during peak hours.

e. Weibull

Modeling new sign-up credit cards cancellation rates, using $k<1$ as customers tend to have a higher likelihood of canceling their cards within the first few days of card activation and this decreases with time.