

# Week 9 Homework - Part 1

## 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.

## Answer 12.1

A Mobile Applications development company might develop a new productivity application or health application with multiple features. If the company would like to know which features to be included for free and which ones should be provided as a part of a paid upgrade, a design of experiments approach could provide the required insights.

The experiment would have trial versions provided to potential users with different sets of features as paid and others as free. The objective is to see which features will encourage users to upgrade and which would not. Also the company would know which features are more important to invest in its development.

## 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.

## Answer 12.2

```
In [4]: # install.packages("FrF2")
library(FrF2)
```

```
In [2]: # Fix Seed Number
set.seed(0)
# Create Experiment using 16 runs (16 fictitious houses) and 10 factors (10 different y
experiment <- FrF2(nruns = 16, nfactors = 10)
# Convert to Data Frame
experiment <- as.data.frame(experiment)
# Display Results of Each experiment (1 = included, -1 = Not included)
experiment
```

A data.frame: 16 × 10

	A	B	C	D	E	F	G	H	J	K
	<fct>	<fct>	<fct>	<fct>	<fct>	<fct>	<fct>	<fct>	<fct>	<fct>
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

Note that Features names are labeled A through K

In [3]:

```
# Store factors in each fictitious house
for (i in seq(1,16)){
  features <- colnames(experiment)[experiment[i,]==1]
  cat("Experiment", i, "features included:", features, "\n")
}
```

```
Experiment 1 features included: A C D F H
Experiment 2 features included: D E F G J
Experiment 3 features included: A B E J K
Experiment 4 features included: B C G H K
Experiment 5 features included: E F G H K
Experiment 6 features included: A G
Experiment 7 features included: C D E K
Experiment 8 features included: B D F K
Experiment 9 features included: B F H J
Experiment 10 features included: A B C E F G
Experiment 11 features included: A B D E H
Experiment 12 features included: C E H J
Experiment 13 features included: A C F J K
Experiment 14 features included: B C D G J
Experiment 15 features included: A D G H J K
Experiment 16 features included: A B C D E F G H J K
```

## 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 b. Geometric c. Poisson d. Exponential e. Weibull

## Answer 13.1

a. Binomial Distribution:

Asking 100 of our Non-American classmates if they have ever been to Georgia. The response might have a binomial distribution.

b. Geometric Distribution:

The number of Petroleum Exploration wells that have to be drilled before a successful hydrocarbon discovery is made could follow a geometric distribution.

c. Poisson Distribution:

The number of expected customers that will arrive at the restaurant per day might have a poisson distribution.

d. Exponential Distribution:

The amount of time between traffic accidents on a certain road might have an exponential distribution

e. Weibull Distribution:

The amount of time before a sensor fails might follow a Weibull distribution. With  $k > 1$  the more time passes the more likely it is the sensor would fail assuming no manufacturing defects.