Homework 9 Submission by Haritha Pulletikurti **2020-10-20**

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

Design of Experiments can be used to determine:

The combination of ingredients (factors) in soap making in a cosmetic manufacturing company that are more liked by the customers.

Here the different factors might be different kind essential oils and other soap organic and inorganic ingredients.

Here we can use fractional factorial design and may create about 15 different subsets of the ingredients and we can survey against 100 sample customers and record which combinations are best liked.

We need to have **control** of factors such as Age and gender among the customer samples.

We can also use **blocking factor** such as the purpose of the soap, is it medicinal to treat any illnesses or is it just used for regular use. This will reduce the variability in the estimates as the regular soaps might be more in demand.

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:

Fractional Factorial Design: This method is used to create a subset of combinations of the features.

In a balanced Factorial Design,

1. **Test each choice the same number of times**
2. Test each pair of choices the same number of times.

If there are Independent Factors:

1. **Test each subset of combinations**
2. Use Regression to estimate the effects.

In the question 12.2,
Number of features/factors = 10.

Number of people to survey = 50

Number of groups for the subset of these 10 factors need to be created = 16.

In this example, we do not know the names of the factors nor we have any information whether the factors are independent or not. So, I will use a simple FrF2() function to get the best subset of factors for each of the 16 groups.

```
library(FrF2)

## Warning: package 'FrF2' was built under R version 4.0.3

## Loading required package: DoE.base

## Warning: package 'DoE.base' was built under R version 4.0.3

## Loading required package: grid

## Loading required package: conf.design

## Warning: package 'conf.design' was built under R version 4.0.3

## 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

# We need to create subsets of 10 different factors in 16 Fictitous houses.

FrF2(16,10)
```

```
##     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


Analysis : The above output shows

Fictious House 1 can include {C,D,E,K} features.

Fictious House 2 can include {A,C,F,J} features.

Fictious House 3 can include {A,G} features etc..
```

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).

Answer

1. **Binomial Distribution:**
   Binomail Distribution models the probability of getting n successes before a failure where each individual success or failure is a **Bernoulli trail** with success probability = p and failure probability = 1-p.

   Binomial Distribution Example: How many wins in a series of table tennis games before a losing to the opponent?

2. **Geometric Distribution:**
   Geometric Distribution models the probability of getting n failures before a success where each individual success or failure is a **Bernoulli trail** with success probability = p and failure probability = 1-p.

   Geometric Distribution Example: How many loses in a series of table tennis games before winning over the opponent by the participants can be modelled by Geometric distribution

3. **Poisson Distribution:**
   This is good at modeling random arrivals. Lambda is average number of arrivals per time period.

   Poisson Distribution Example: Average number of people arriving at the movies ticketing queue in a day can be modelled using Poisson Distribution.

4. **Exponential Distribution:** Time between successive arrivals is modeled using Exponential Distribution.
   Exponential Distribution Example: The time between successive arrivals of people arriving at the movies ticketing queue can be modelled using Exponential Distribution.

5. **Weibull Distribution:** Time between failures is modeled using Weibull.
   Weibull Distribution Example: The time between the failures for the manufactured TV units that are bought during a Thanksgiving sale which were reported for replacement before a certain warranty period. Here k > 1 (time between failures is more for the units as they are new) can be modelled using Weibull Distribution.

Question 13.2

In this problem you, can simulate a simplified airport security system at a busy airport. Passengers arrive according to a Poisson distribution with $\lambda_1 = 5$ per minute (i.e., mean interarrival rate $\mu_1 = 0.2$ minutes) to the ID/boarding-pass check queue, where there are several servers who each have exponential service time with mean rate $\mu_2 = 0.75$ minutes. [Hint: model them as one block that has more than one resource.] After that, the passengers are assigned to the shortest of the several personal-check queues, where they go through the personal scanner (time is uniformly distributed between 0.5 minutes and 1 minute).

Use the Arena software (PC users) or Python with SimPy (PC or Mac users) to build a simulation of the system, and then vary the number of ID/boarding-pass checkers and personal-check queues to determine how many are needed to keep average wait times below 15 minutes. [If you're using SimPy, or if you have access to a non-student version of Arena, you can use $\lambda_1 = 50$ to simulate a busier airport.]

Answer:

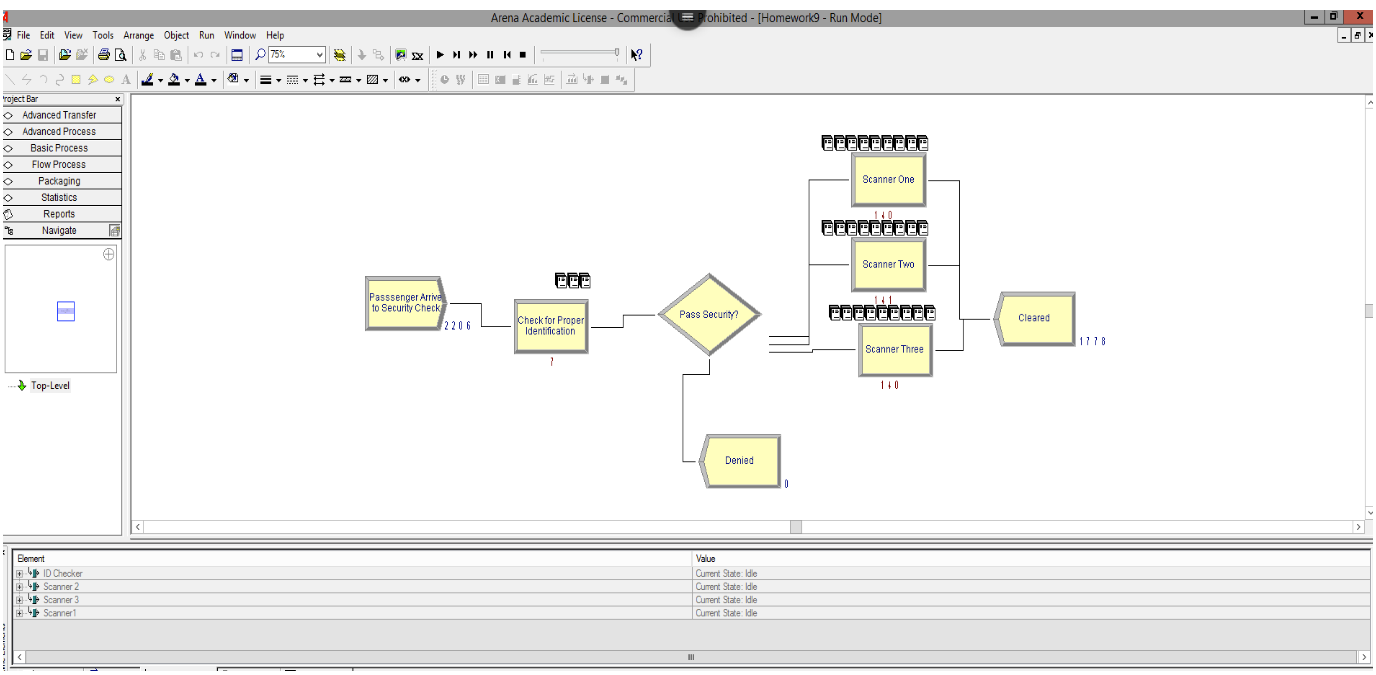
I used Arena software to model the above scenario.

Passenger Arrivals are modelled with mean1 = 0.2 mins (mean arrival rate)

ID Checker is modelled using Exponential service time with mean2 = 0.75

I created 2 models :

Model 1: The Id Checker sends the passengers to a decision block where all the passengers who passed the ID Check go through three scanners . These personal scanners are modelled with a uniform distribution between 0.5 to 1 minute.



Wait Time	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Entity 1	91.7422	3.83	84.6787	101.60	0.00	193.63

Observe that the average wait time with 3 scanners is 91.7422 minutes where the maximum wait time value is 193.63 minutes.

Replications: 10 Time Units: Minutes

Queue

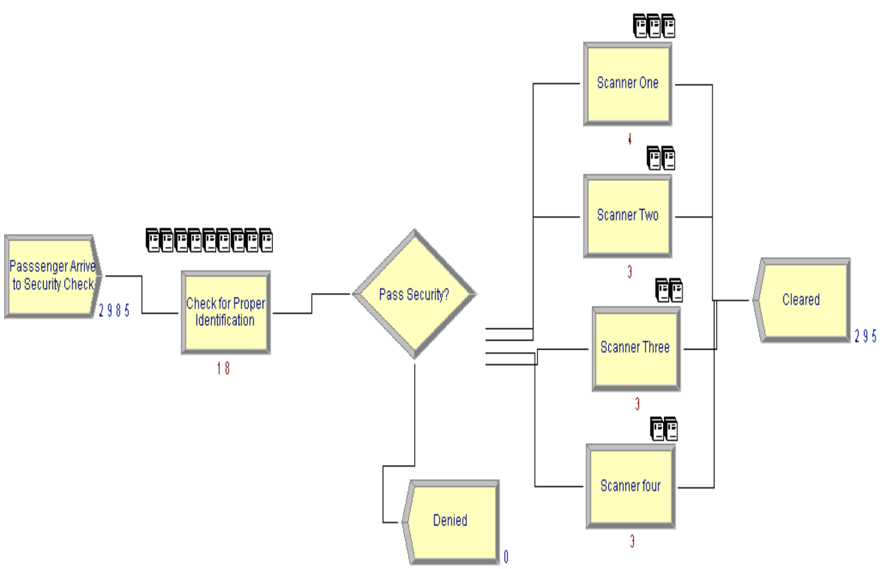
Time

Waiting Time	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Check for Proper Identification.Queue	1.9787	0.68	1.2519	4.5346	0.00	13.1964
Scanner One.Queue	89.9203	3.53	83.7844	99.94	0.00	192.91
Scanner Three.Queue	89.4095	3.44	83.4838	98.8634	0.00	186.93
Scanner Two.Queue	89.8481	3.48	83.1163	99.03	0.00	191.70

Other

Number Waiting	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Check for Proper Identification.Queue	9.8791	3.48	6.2948	22.8733	0.00	69.0000
Scanner One.Queue	149.03	6.00	138.74	164.41	0.00	311.00
Scanner Three.Queue	148.36	6.00	138.10	163.73	0.00	310.00
Scanner Two.Queue	148.70	6.00	138.44	164.10	0.00	311.00

Model 2: The Id Checker sends the passengers to a decision block where all the passengers who passed the ID Check go through four scanners . These personal scanners are modelled with a uniform distribution between 0.5 to 1 minute.



Wait Time	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Entity 1	3.7695	0.58	2.5939	4.6183	0.00	18.6752

Observe that the average wait time for Model 2 with 4 scanners is 3.7695 minutes with the maximum wait value 18.6752.

Replications: 10 Time Units: Minutes

Queue

Time

Waiting Time	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Check for Proper Identification.Queue	2.2093	0.48	1.2982	3.1307	0.00	14.9693
Scanner four.Queue	1.4747	0.22	1.1234	2.1804	0.00	8.6549
Scanner One.Queue	1.7381	0.22	1.3778	2.4474	0.00	9.4057
Scanner Three.Queue	1.4534	0.23	1.0645	2.2082	0.00	8.2291
Scanner Two.Queue	1.5806	0.22	1.2155	2.2984	0.00	9.1385

Other

Number Waiting	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Check for Proper Identification.Queue	11.0793	2.46	6.4186	15.5551	0.00	90.0000
Scanner four.Queue	1.8371	0.29	1.3566	2.7740	0.00	12.0000
Scanner One.Queue	2.3169	0.30	1.8270	3.2557	0.00	12.0000
Scanner Three.Queue	1.6024	0.29	1.1356	2.5418	0.00	11.0000
Scanner Two.Queue	2.0818	0.30	1.6042	3.0243	0.00	12.0000

As we can clearly observe, the average wait times of 4 Scanners -Model 2 (3.7695 minutes) is much better than 3 Scanners - Model 1(91.7422 minutes). Also the Queue wait times are much better for Model 2 than Model 1.

Hence, we can conclude that model with 4 scanners is the best Model.

Please see the attached summary reports of the above experiments for more details.