Math 446

Statistical analysis of the published paper “Appraising healthcare systems’ efficiency in facing COVID-19 through data envelopment analysis”

By: Jordon Silcox and Lormel Bationo

I. Introduction

The COVID-19 pandemic emergence was characterized by the rapid transmission of the virus resulting in high mortality rates across various countries worldwide. This unprecedented global health crisis has placed the global healthcare system under immense pressure which revealed critical gaps in both their state of readiness as well as their response capabilities. For many countries, it revealed that needed to be investigated and then addressed. The severe consequences of this pandemic proved the need to develop a solid healthcare system which is capital in effectively preventing and managing socio economic crisis. This led to the need to conduct research into healthcare systems efficiency of different countries to improve its structures and improve efficiency. This study seeks to assess the level of readiness of 29 countries and their effectiveness in facing a global health crisis. Seven scenarios were employed using the Data Envelopment Analysis methodology (DEA) incorporating several variables such as the number of medical practitioners, available hospital beds, the number of infections and recovery rates as well as the relative mortality rate...Other methods were employed such as the Tobit regression analysis. Some non-parametric tests such as the tests Wilcoxon signed test, Mann-Whitney U, and Kruskal-Wallis H were used to assess the significance of the chosen DEA models. The result of this study would provide valuable insight for implementing the necessary changes to build a strong and efficient health care system capable of efficiently dealing with the challenges posed by unexpected health crisis.

II. Summary

Covid-19 is an infectious disease that affects the respiratory system and swept across the globe when it emerged in December 2019. Its infectiousness caused a worldwide pandemic, without an effective vaccine or treatment, it left many hospitals to deal with a surplus of patients in critical need of health care. This pandemic lasted until September 2020 and can be a period to gather information on the efficiency of different medical systems in highly populated countries. This is what the article "Appraising healthcare systems' efficiency in facing COVID-19 through data envelopment analysis" (Mourad, N. et. al, 2021) aims to do by using six different variables: medical practitioners, Covid-19 test conducted, affected cases, hospital beds, recovered cases, and death cases. These variables are separated into two groups input and output used for the data envelopment analysis (DEA) which is a nonparametric test implanted in the paper to determine the efficiency of different countries health systems capability to keep up with demand resulted from the unique situation like the coronavirus pandemic.

The four input variables are tested conducted, affected cases, medical practitioners, and hospital beds. For the two output variables are recovered cases and death cases, all numbers are per million. These variables responsible for converting inputs into outputs are called Decision Making Units (DMUs) which the DEA model measures to provide information for assessing and optimizing these DMUs. Twenty-nine countries were chosen for this study, this included seven different scenarios involving the input and output variables:

- Scenario 1: Inputs: tests conducted, affected cases, medical practitioners, and hospital beds; Outputs: recovered cases and death cases.
- Scenario 2: Number of death cases is removed from the variables in scenario 1.
- Scenario 3: Number of conducted tests is removed from the variables in scenario 1.
- Scenario 4: Number of conducted test and death cases is removed from the variables in scenario 1.

- Scenario 5: Number of affected cases and recovered cases are the only variables used in the DEA model.
- Scenario 6: The number of medical practitioners is added to the input variable from scenario 5.
- Scenario 7: The number of hospital beds is added to the input variable from scenario 5.

These scenarios with different types of inputs and outputs are plugged into the DEA model of formula:

$$e_n = \max_{(\mu,v)\in\mathbb{R}_+^{m\times s}} \frac{\sum_{j=1}^{s} v_j\, y_{jn}}{\sum_{i=1}^{m} \mu_i\, x_{in}},\ where\ \frac{\sum_{j=1}^{s} v_j\, y_{jn}}{\sum_{i=1}^{m} \mu_i\, x_{in}} < 1\ for\ n = 1,2,3,\dots,N$$

For the above equation $\{x_{in}\}_{\{1\le i\le m}$ is the n -th input variables while $\{y_{jn}\}_{\{1\le j\le s}$ is the n -th output variables. The sample size N, which represents the number of comparable DMUs, and the suggested appropriate sample size should be twice the number of inputs and output by Golany and Roll (1989) or three times the number inputs and outputs suggested by Banker et al. (1989) and Cooper et al. (2007). The variables v andμ are weights associated with associated with inputs and outputs.

The fractional problem can be turned into a linear one, but doing this transformation comes with a choice of orientation. This is done by having control over decreasing the inputs or when they have control over maximizing the outputs while maintaining the input variables:

Input-oriented	Output-oriented.
$\sum_{i=1}^{m} \mu_i\ x_{in} = 1$	$\sum_{j=1}^{s} v_j\ y_{jn} = 1$
$\sum_{j=1}^{s} v_j\ y_{jn} - \sum_{i=1}^{m} \mu_i\ x_{in} \le 0$	$\sum_{i=1}^{m} \mu_i\ x_{in} - \sum_{j=1}^{s} v_j\ y_{jn} \le 0$

Also, since the number of comparable DMUs (N) is larger than the amount of considered variables (m+s) the dual of these problems need to be used to determine so the fewest number of constraints are obtained:

Input-oriented	Output-oriented
$e_n = \min_{\lambda\in\mathbb{R}_+^N}(\theta_n)$	$\frac{1}{e_n} = \max_{\lambda\in\mathbb{R}_+^N}(\theta_n)$
$\sum_{j=1}^{N} \lambda_j\ x_{ij} \le \theta_n\, x_{in}\,,\ i = 1,\dots,m$	$\sum_{j=1}^{N} \lambda_j\ x_{ij} \le x_{in}\,,\ i = 1,\dots,m$
$\sum_{j=1}^{N} \lambda_r\ y_{rn} \ge y_{rn}\,,\ r = 1,\dots,s$	$\sum_{j=1}^{N} \lambda_r\ y_{rn} \ge \theta_n\, y_{rn}\,,\ r = 1,\dots,s$

The variable λ is the vector of the associated weights with the DMUs. An efficiency score of one will mean that the DMU is efficient, if it is less than one will mean it is inefficient.

The other model used in this paper is the Tobit regression analysis, which is done to verify the major factors influencing the efficiency of different healthcare systems in highly populated countries. Tobit regression was chosen because it is best used when there are restraints on the dependent variable. The Tobit model used is:

$$y_i^* = X_i^T\beta + \epsilon_i\,,\ i = 1,2,\dots n$$

ϵ_i is the Gaussian white noise, which is the random error of the model. $X_i$is the vector containing the data corresponding to the independent variables. Regression is done on the non-observed data and observed dependent variables are given by:

$$y_i = \max(y_i^*, 0)\,,\ i = 1,2,..,n$$

With m independent variables, the linear multivariate regression can be expressed as:

$$y = X^T\beta = \beta_\circ + \beta_1 x_1 + \cdots + \beta_m x_m$$

where $\theta = (\beta_\circ, \beta_1, \ldots, \beta_m, \sigma)$ gets maximized by the log-likelihood function:

$$L_{n(\theta)} = \frac{1}{n}\sum_{i=1}^{n}\left(1 - \delta_{yi=0}\right)\log\left[\frac{1}{\sigma}\phi\left(\frac{y_i - X_i^T\beta}{\sigma}\right)\right] + \delta_{yi=0}\log\left[\phi\left(\frac{y_i - X_i^T\beta}{\sigma}\right)\right]$$

Statistical analysis was used to verify certain aspects of the data between the input and output orientation. The first of these non-parametric tests used was the Sign test to determine if there is a difference between medians of the two data sets. The second test used was the Wilcoxon Sign Rank test, which is like the Sign test, but is a stronger test for paired samples which takes the assumption that the data come from the same population and needs to be symmetric. Another statistical test employed is the Mann-Whitney test is used to determine if the two data sets have identical shaped distributions. The final statistical test used was the Kruskal-Wallis H test to find if there is a significant difference between the efficiency scores obtained from the DEA model. Lastly, the Spearman Rank test was used to indicate correlation between the input and output orientations which did show there is high positive correlation. For another analysis not presented in the paper performed on the different seven scenarios is the Siegel-Tukey test to determine if they come from distributions with equal variances.

There are four hypotheses being tested:

Hypotheses 1: The variance between the healthcare systems' scores obtained from the DEA model based on the various considered scenarios is not statistically significant.

Hypothesis 2: The variance between the healthcare systems' scores based on the orientation of the DEA model is not statistically significant.

Hypothesis 3: All the considered inputs and outputs variables under analyses have no statistically significant impact on the healthcare systems' scores over the study period.

Hypotheses 4: The GDP per capita has no statistically significant influence on the healthcare systems' scores over the study period.

Of the scenarios tested, using both Wilcoxon and Sign test, 5 and 7 had significant p-values. With scenario 5 preferring the input-oriented model at a confidence level of 0.01, while scenario 7 favoring the output-oriented model at a confidence level of 0.05. The results of the Mann-Whitney test did indicate they different scenarios have similar shaped orientation. For another analysis not presented in the paper performed on the different seven scenarios is the Siegel-Tukey test to determine if they come from distributions with equal variances.

III. Statistical Analysis

The DEA method gave us relative efficiency scores to evaluate the performance of the healthcare system regarding the output-oriented method and the input-oriented method. The result indicates that there is relative efficiency in the performance of 13 healthcare systems according to the first and second scenario facing COVID-19. It also suggested that there was a relative efficiency in the performance of 11 healthcare systems according to the third, fourth and seventh scenarios.

We used the results compiled in the relative efficiency summary to evaluate the DEA method results and to test whether there is a statistically significant difference in the healthcare system efficiency scores between the DEA output and input oriented models. We perform two non-parametric tests: the Wilcoxon test and the sign test. The result (figure1) suggests that the median of differences between both models is equal to zero according to all considered scenarios except the 5th and 7th scenarios, where there was a significant difference between the efficiency scores according to scenario 5 in favor of the output-oriented model, while there was a significant difference according to scenario 7 in favor of the input-oriented model.

To replicate the tests in R, we considered the 7 different scenarios of the DEA method. For the Wilcoxon test, we created two vectors in each scenario to store the relative efficiency score from the two different outputs method and used the R function Wilcox.test. For the sign test we used the same approach, but we created other variables in each of the scenarios to compute and store the difference between each method. We used the R function binom.test. Our findings (figure2) line up with the result from the analysis. These results suggest that in terms of decision making to improve the relative efficiency, any of the orientation whether input or output can be adopted except for scenario 5 and 7 because the scenario 5 yields better result with the output-oriented model while the scenario 7 is best with input-oriented model.

A Tobit regression was performed in the analysis to determine the impact of the drivers of the healthcare systems' efficiency. We assumed that the variance between the healthcare systems' scores obtained from the DEA model based on the various considered scenarios is not statistically significant. Different variables like the total number of recovered cases, reverse ratio of death cases, total tests, affected cases, medical practitioners, hospital beds, and GDP per capita are considered over the study period. The analysis suggests that the overall scores of these healthcare systems' efficiencies are positively correlated with the total number of recovered cases per million at the 0.01 significance level. This means that the countries will be more efficient if they can increase the total cases recovering from the COVID-19 pandemic. On the other side, the total number of affected cases per million has a significant negative impact on the efficiency scores at the 0.01 significance level. This means that the countries will be more efficient if they can decrease the total affected cases from the COVID-19 pandemic.

We replicated a validity test as in the analysis, to test that the distribution of all models is the same across all the scenarios. We used R to replicate the test Kruskal Wallis test. We created two data frames to regroup the DEA data by input or output method. The result obtained (figure3) confirmed the robustness of the DEA model.

To finish our analysis, we performed a parametric test to check whether there was a difference between the mean of the paired groups regarding the method used (input or output). Since the data are paired, we used. The same vectors were used to store the DEA score values and used the t test in R with the paired function. Only the results of scenario 5 were significant. All the other scenarios showed that there was no significant difference between the means of the groups.

IV. Discussion

The aim of the study was to assess the performance of the different healthcare systems analyzed and gain insight in their response capacity towards global health crisis like COVID-19. The method uses a multiple variables approach coupled with the DEA method to ensure the efficiency of the decision-making units. The DEA provides scores ranging from 0 to 1, where 1 represents perfect efficiency and scores less than 1 represent inefficiency. Multiples scenarios were considered which can help evaluating the different assumptions and provide a more comprehensive assessment of efficiency.

This approach is efficient because it measures not only the efficiency of the test but can also help to highlight key identify the best practices and provide areas for improvement. The analysis of the results shows that the DEA model showed consistent results in the average efficiency scores across the different scenarios. This shows that either orientation can achieve best practices. Therefore, decision making to improve performance can rely on either of the scenarios. This research was very well conducted in terms of balance because it combined a nonparametric test with a regression model used to feature efficient drivers. This would make decision making easier because it provides the cause of the inefficiency.

Some of the limits of non-parametric tests for statistical analysis are that they are sometimes less powerful than parametric tests. They provide less information about the correlation between variables. Knowing whether the healthcare system of a particular country was efficient or not is not sufficient if we want to address the issue. This is the reason why the use of the Tobit regression was important because it determines the key factors that influence the healthcare system. It is more powerful in providing valuable information in areas that require improvement to increase efficiency. However, even though the approach can yield efficient results and provide tools for improvements, it relies on variables that are constantly changing. Therefore, future research should be conducted to make comparisons with the findings.

Appendix

Mourada, N., Habibb, A. M., & Tharwata, A. (2021), Appraising healthcare systems' efficiency in facing COVID-19 through data envelopment analysis. *Decision Science Letters. 10*, 301-310. https://growingscience.com/dsl/Vol10/dsl_2021_12.pdf

Steering Committee for the Review of Commonwealth/State Service Provision 1997, Data Envelopment Analysis: A technique for measuring the efficiency of government service delivery, AGPS, Canberra. https://www.pc.gov.au/research/supporting/data-envelopment-analysis/dea.pdf

Tables

Test results of the differences among the DEA models orientation

Test		Test Statistics						
		Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5	Scenario 6	Scenario 7
Wilcoxon test	**Sig.**	0.574	0.574	0.412	0.412	0.003**	0.628	0.038*
Sign test	**Sig.**	0.267	0.267	0.180	0.180	0.003**	0.118	0.022*

Note: * and ** are significant at the 5 and 1% levels, respectively.

Figure 1 Wilcoxon and signed test

Scenario	Wilcoxon Signed-Rank Test (p-value)	Sign Test (p-value)
1	0.5738	0.2668
2	0.5738	0.2668
3	0.412	0.1796
4	0.412	0.1796
5	0.003135	0.003418
6	0.6285	0.1185
7	0.04137	0.02246

Table 1: Results from Wilcoxon signed-rank tests and sign tests for 7 different scenarios

Figure 2 Wilcoxon and sign test R result

Model	**Data**	**Kruskal-Wallis**	**p-value**
Output	*a, c, e, g, i, k, m*	3.2096	0.7821
Input	*b, d, f, h, j, l, n*	3.7494	0.7105

Table 2: Results of Kruskal-Wallis tests

Figure 3Kruskal-Wallis R output summary

R script and output

```
#M446 Project R scipt

#Packages
install.packages("car")
library(car)
###Scenario 1#####

a = c(1.000, 1.000, 0.726, 0.787, 1.000,1.000,1.000, 0.916, 0.901,1.000,
0.964, 1.000, 0.881,1.000, 1.000, 1.000, 0.959, 0.916, 0.953, 0.997,
0.329, 0.190, 0.795, 1.000, 1.000,0.297, 0.681, 0.954, 1.000)

b= c(1.000, 1.000, 0.704, 0.788, 1.000, 1.000, 1.000,0.966, 0.900,1.000,
0.964, 1.000, 0.878,1.000,1.000,1.000, 0.958, 0.915, 0.952, 0.997,
0.320, 0.182, 0.791, 1.000, 1.000,0.656,0.690, 0.954, 1.000)

ab = c(a-b)

ab1 = ab[!ab == '0']

length(ab1) #13

wilcox.test(a,b, alternative= "two.sided", paired = TRUE)
#
# Wilcoxon signed rank test with continuity correction
#
# data:  a and b
# V = 54, p-value = 0.5738
# alternative hypothesis: true location shift is not equal to 0

binom.test(sum(ab > 0), length(ab1) ,p = 0.5, alternative = "two.sided")

# Exact binomial test
#
# data:  sum(ab > 0) and length(ab1)
# number of successes = 9, number of trials = 13, p-value = 0.2668
# alternative hypothesis: true probability of success is not equal to
0.5
# 95 percent confidence interval:
#   0.3857383 0.9090796
# sample estimates:
#   probability of success
# 0.6923077


result1 <- t.test(a, b, paired = TRUE)
```

```
result1

# Paired t-test
#
# data:  a and b
# t = -1.0157, df = 28, p-value = 0.3185
# alternative hypothesis: true mean difference is not equal to 0
# 95 percent confidence interval:
#   -0.03838475  0.01293647
# sample estimates:
#   mean difference
# -0.01272414



######Scenario 2######

c = c( 1.000, 1.000, 0.726, 0.787, 1.000, 1.000, 1.000,0.916, 0.901,
1.000, 0.964, 1.000, 0.881,1.000, 1.000, 1.000, 0.959, 0.916, 0.953,
0.997, 0.329, 0.190, 0.795, 1.000, 1.000,0.296, 0.681, 0.954, 1.000)

d = c(1.000, 1.000, 0.704, 0.788, 1.000, 1.000, 1.000,0.966, 0.900,1.000,
0.964, 1.000, 0.878,1.000, 1.000, 1.000, 0.958, 0.915, 0.952, 0.997,
0.320, 0.182, 0.791, 1.000, 1.000,0.656,0.690, 0.954, 1.000)

cd = c(c-d)

wilcox.test(c,d, alternative= "two.sided", paired = TRUE)

# Wilcoxon signed rank test with continuity correction
#
# data:  c and d
# V = 54, p-value = 0.5738
# alternative hypothesis: true location shift is not equal to 0

binom.test(sum(cd > 0), 13, p=0.5, alternative = "two.sided")

# Exact binomial test
#
# data:  sum(cd > 0) and 13
# number of successes = 9, number of trials = 13, p-value = 0.2668
# alternative hypothesis: true probability of success is not equal to
0.5
# 95 percent confidence interval:
#   0.3857383 0.9090796
# sample estimates:
```

```
#   probability of success
# 0.6923077


result2 <- t.test(c, d, paired = TRUE)
result2

# Paired t-test
#
# data:  c and d
# t = -1.0157, df = 28, p-value = 0.3185
# alternative hypothesis: true mean difference is not equal to 0
# 95 percent confidence interval:
#   -0.03848896  0.01297172
# sample estimates:
#   mean difference
# -0.01275862

#######Scenario 3#######

e = c(1.000, 1.000, 0.726, 0.787, 1.000, 1.000, 1.000,0.860, 0.901,
0.792, 0.964, 1.000, 0.881, 0.984, 1.000, 1.000, 0.959, 0.916, 0.953,
0.997, 0.329, 0.190, 0.795, 1.000, 1.000,0.297, 0.681, 0.954, 1.000)

f = c(1.000, 1.000, 0.704, 0.788, 1.000, 1.000, 1.000,0.941, 0.900,
0.789, 0.964, 1.000, 0.878, 0.984, 1.000, 1.000, 0.958, 0.915, 0.952,
0.997, 0.320, 0.182, 0.791, 1.000, 1.000,0.656, 0.689, 0.954, 1.000)

ef = c(e-f)

wilcox.test(e,f,alternative = "two.sided", paired = TRUE)

# Wilcoxon signed rank test with continuity correction
#
# data:  e and f
# V = 66, p-value = 0.412
# alternative hypothesis: true location shift is not equal to 0

binom.test(sum(ef > 0),14, p = 0.5, alternative = "two.sided" )

# Exact binomial test
#
# data:  sum(ef > 0) and 14
# number of successes = 10, number of trials = 14, p-value = 0.1796
# alternative hypothesis: true probability of success is not equal to
0.5
```

```
# 95 percent confidence interval:
#   0.4189647 0.9161107
# sample estimates:
#   probability of success
# 0.7142857

result3 <- t.test(e, f, paired = TRUE)
result3

# Paired t-test
#
# data:  e and f
# t = -1.076, df = 28, p-value = 0.2911
# alternative hypothesis: true mean difference is not equal to 0
# 95 percent confidence interval:
#   -0.03964998  0.01233963
# sample estimates:
#   mean difference
# -0.01365517
#######Scenario 4#########

g = c(1.000, 1.000, 0.726, 0.787, 1.000, 1.000, 1.000,0.860, 0.901,
0.792, 0.964, 1.000, 0.881, 0.984, 1.000, 1.000, 0.959, 0.916, 0.953,
0.997, 0.329, 0.190, 0.795, 1.000, 1.000, 0.296, 0.681, 0.954, 1.000)

h = c(1.000, 1.000, 0.704, 0.788, 1.000, 1.000, 1.000,0.941,0.900, 0.789,
0.964, 1.000, 0.878, 0.984, 1.000, 1.000, 0.958, 0.915, 0.952, 0.997,
0.320, 0.182, 0.791, 1.000, 1.000,0.656, 0.689, 0.954, 1.000)

gh = c(g-h)

gh1 = gh[!gh == '0']

length(gh1)

wilcox.test(g,h, alternative= "two.sided", paired = TRUE)
#
# Wilcoxon signed rank test with continuity correction
#
# data:  g and h
# V = 66, p-value = 0.412
# alternative hypothesis: true location shift is not equal to 0

binom.test(sum(gh > 0), length(gh1) ,p = 0.5, alternative = "two.sided")

# Exact binomial test
```

```
#
# data:  sum(gh > 0) and length(gh1)
# number of successes = 10, number of trials = 14, p-value = 0.1796
# alternative hypothesis: true probability of success is not equal to
0.5
# 95 percent confidence interval:
#   0.4189647 0.9161107
# sample estimates:
#   probability of success
# 0.7142857

result4 <- t.test(g, h, paired = TRUE)
result4

# Paired t-test
#
# data:  g and h
# t = -1.0759, df = 28, p-value = 0.2912
# alternative hypothesis: true mean difference is not equal to 0
# 95 percent confidence interval:
#   -0.03975312  0.01237381
# sample estimates:
#   mean difference
# -0.01368966

######Scenario 5############

i = c(1.000,0.916, 0.726, 0.787, 1.000, 1.000,0.900,0.814, 0.901, 0.792,
0.964, 0.436, 0.881, 0.984, 1.000, 0.999, 0.959, 0.916, 0.953,0.993,
0.329, 0.190, 0.795, 1.000, 1.000, 0.296,0.680, 0.954, 1.000)

j = c(1.000,0.914, 0.704, 0.787, 1.000, 1.000, 0.900,0.807, 0.900, 0.788,
0.964,0.436, 0.878, 0.984, 1.000, 0.999, 0.958, 0.915, 0.952,0.993,
0.320, 0.182, 0.791, 1.000, 1.000,0.297,0.680, 0.954, 1.000)

ij = c(i-j)

ij1 = ij[!ij == '0']

length(ij1)

wilcox.test(i,j, alternative= "two.sided", paired = TRUE)

# Wilcoxon signed rank test with continuity correction
#
# data:  i and j
```

```
# V = 88, p-value = 0.003135
# alternative hypothesis: true location shift is not equal to 0

binom.test(sum(ij > 0), length(ij1) ,p = 0.5, alternative = "two.sided")

# Exact binomial test
#
# data:  sum(ij > 0) and length(ij1)
# number of successes = 12, number of trials = 13, p-value = 0.003418
# alternative hypothesis: true probability of success is not equal to
0.5
# 95 percent confidence interval:
#   0.6397026 0.9980544
# sample estimates:
#   probability of success
# 0.9230769

result5 <- t.test(i, j, paired = TRUE)
result5

# Paired t-test
#
# data:  i and j
# t = 2.4966, df = 28, p-value = 0.0187
# alternative hypothesis: true mean difference is not equal to 0
# 95 percent confidence interval:
#   0.000383807 0.003892055
# sample estimates:
#   mean difference
# 0.002137931
##########Scenario 6#############

k= c(1.000,0.916, 0.726, 0.787, 1.000, 1.000,0.900,0.839, 0.901, 0.792,
0.964,   0.455,   0.881,   0.984,   1.000,   1.000,   0.959,   0.916,
0.953,0.994,0.329,  0.190,  0.795,  1.000,  1.000,  0.296,  0.681,  0.954,
1.000)

l = c(1.000,0.914, 0.704, 0.787, 1.000, 1.000,0.900,0.864, 0.900, 0.788,
0.964,0.801,  0.878,  0.984,  1.000,  1.000,  0.958,  0.915,  0.952,0.994,
0.320, 0.182, 0.791, 1.000, 1.000,0.357, 0.689, 0.954, 1.000)

kl = c(k-l)

kl1 = kl[!kl == '0']

length(kl1)
```

```
wilcox.test(k,l, alternative= "two.sided", paired = TRUE)

# Wilcoxon signed rank test with continuity correction
#
# data:  k and l
# V = 69, p-value = 0.6285
# alternative hypothesis: true location shift is not equal to 0

binom.test(sum(kl > 0), length(kl1) ,p = 0.5, alternative = "two.sided")
# Exact binomial test
#
# data:  sum(kl > 0) and length(kl1)
# number of successes = 11, number of trials = 15, p-value = 0.1185
# alternative hypothesis: true probability of success is not equal to
0.5
# 95 percent confidence interval:
#   0.4489968 0.9221285
# sample estimates:
#   probability of success
# 0.7333333

result6 <- t.test(k, l, paired = TRUE)
result6

# Paired t-test
#
# data:  k and l
# t = -1.0905, df = 28, p-value = 0.2848
# alternative hypothesis: true mean difference is not equal to 0
# 95 percent confidence interval:
#   -0.03811393  0.01163117
# sample estimates:
#   mean difference
# -0.01324138

######Scenario 7#########

m = c(1.000,1.000, 0.726, 0.787, 1.000, 1.000,1.000,0.814, 0.901, 0.792,
0.964, 1.000, 0.881, 0.984, 1.000, 1.000, 0.959, 0.916, 0.953, 0.997,
0.329, 0.190, 0.795, 1.000, 1.000, 0.296,0.680, 0.954, 1.000)

n = c( 1.000, 1.000, 0.704, 0.788, 1.000, 1.000, 1.000,0.807, 0.900,
0.789, 0.964, 1.000, 0.878, 0.984, 1.000, 1.000, 0.958, 0.915, 0.952,
0.997, 0.320, 0.182, 0.791, 1.000, 1.000,0.655,0.680, 0.954,1.000)
```

```
mn = c(m-n)

mn1 = mn[!mn == '0']

length(mn1)

wilcox.test(m,n, alternative= "two.sided", paired = TRUE)

# Wilcoxon signed rank test with continuity correction
#
# data:  m and n
# V = 75, p-value = 0.04137
# alternative hypothesis: true location shift is not equal to 0
#  binom.test(sum(mn  >  0),  length(mn1)  ,p  =  0.5,  alternative  =
"two.sided")



binom.test(sum(mn > 0), length(mn1) ,p = 0.5, alternative = "two.sided")

# Exact binomial test
#
# data:  sum(mn > 0) and length(mn1)
# number of successes = 11, number of trials = 13, p-value = 0.02246
# alternative hypothesis: true probability of success is not equal to
0.5
# 95 percent confidence interval:
#   0.5455289 0.9807933
# sample estimates:
#   probability of success
# 0.8461538

result7 <- t.test(k, l, paired = TRUE)
result7

# Paired t-test
#
# data:  k and l
# t = -1.0905, df = 28, p-value = 0.2848
# alternative hypothesis: true mean difference is not equal to 0
# 95 percent confidence interval:
#   -0.03811393  0.01163117
# sample estimates:
#   mean difference
# -0.01324138
```

```
#Kruskal wallis test(output model)

data <- list(a, c, e, g, i, k, m)

# Perform Kruskal-Wallis test
kruskal.test(data)
# Kruskal-Wallis rank sum test
#
# data:  data
# Kruskal-Wallis chi-squared = 3.2096, df = 6, p-value = 0.7821

#Kruskal wallis test(input model)
data2 <-list(b,d,f,h,j,l,n)

# kruskal.test(data2)
#
# Kruskal-Wallis rank sum test
#
# data:  data2
# Kruskal-Wallis chi-squared = 3.7494, df = 6, p-value = 0.7105
```