Tests on Dispersion by Capon's and Savage Statistic

A Simulation Study

Pritam Dey
Sweata Sen
Sourojyoti Barick

April 14, 2020

Indian Statistical Institute

What is a Scale Parameter?

If a family of probability distributions is such that there is a parameter s (and other parameters θ) for which the cumulative distribution function satisfies

$$F(x; s, \theta) = F(\frac{x}{s}; 1, \theta)$$

then s is called a scale parameter, since its value determines the **"scale"** or statistical dispersion of the probability distribution.

Lets take an example of scale parameter:

$$F(x; \lambda) = \begin{cases} 0 \;\; when\ x \leq 0 \\ 1 - \exp\left(\frac{x}{\lambda}\right) \;\; o.w. \end{cases}$$

Here λ is the scale parameter.

Problem of interest

Let,

$$\textit{independent} \begin{cases} X_1, X_2, ..., X_n & \text{iid } F(x) \\ Y_1, Y_2, ..., Y_m & \text{iid } G(x) \end{cases}$$

Here we assume $G(x)$ to be of the form $F(\theta x)$.

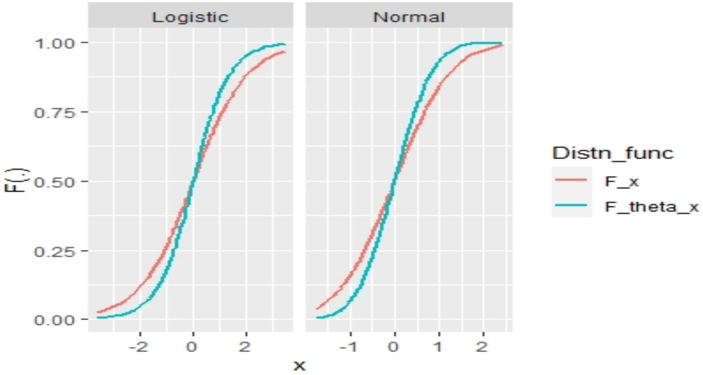
Our concern is the test for dispersion i.e.,

$$\begin{aligned} H_0 : \theta = 1 \; vs. \; & H_1 : \theta > 1 \\ vs. \; & H_1 : \theta < 1 \\ vs. \; & H_1 : \theta \neq 1 \end{aligned}$$

We have drawn all our inferences for testing H_0 *vs.* H_1. Under H_0, $X \stackrel{\mathrm{d}}{=} Y$. Now we shall see how the distribution functions of X and Y look under $\theta > 1$.

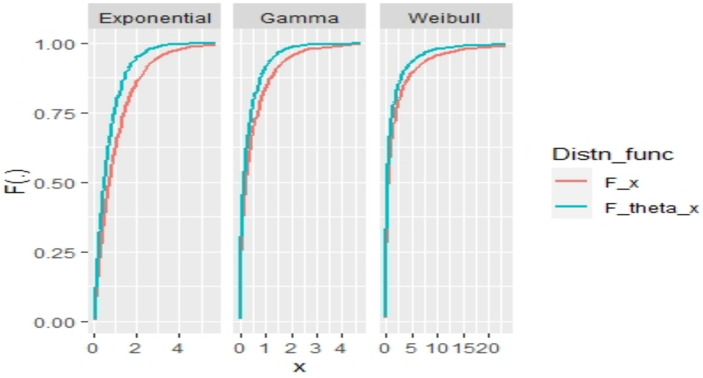
Illustrations

We have random variables from $N(0,1)$ and $L(0.5,1)$ distributions i.e., with scale parameter $= 1$ under H_0.



Illustrations (Contd.)

We have random variables from $Exp(0,1), G(0.5,1)$ and $Weibull(0.5,1)$ distributions i.e., with scale parameter $= 1$ under H_0.



Test Statistics

In our project we mainly focus on the non-Parametric Test of scale parameter using distribution free Test statistic (mainly **Savage** and **Capon's** Statistic).
Let $R_1, R_2, ..., R_n$ denote the ranks of the X observations in the combined ranking of n $X's$ and m Y observations.

$$C = \sum_{i=1}^{n} E(Z_{R_i}^2)$$

$$S = \sum_{i=1}^{n} \sum_{j=1}^{R_i} \frac{1}{n-j+1}$$

where $Z_{(i)}$ denotes the i^{th} order statistic from $N(0,1)$.

Behaviour under H_0

- The mean and variance of Capon's statistic under H_0 are:

$$E(C) = n$$

$$Var(C) = \frac{nm}{n(n-1)} \sum_{i=1}^{n+m} (E(Z_{(i)}^2))^2 - \frac{nm}{(n-1)}$$

- The mean and variance of Savage statistic under H_0 are:

$$E(S) = n$$

$$Var(S) = \frac{S - n}{\sqrt{\frac{nm}{n-1}(1 - \frac{1}{n} \sum_{i=1}^{n+m} \frac{1}{i})}}$$

Under H_0, $\frac{C - E_{H_0}(C)}{\sqrt{Var_{H_0}(C)}}$ and $\frac{S - E_{H_0}(S)}{\sqrt{Var_{H_0}(S)}}$ follow $N(0,1)$ asymptotically.

Distribution Free

To show that the statistics are distribution free under H_0,

- We have simulated both Capon and Savage statistics for $n = 25, m = 25$.
- Then we have replicated the same 1000 times for each of the 5 distributions mentioned earlier.
- We have done qqplot of statistic of one distribution vs the statistic of another distribution.

Since, we have 5 distributions, we have 10 qqplots under each statistic.

Distribution Free

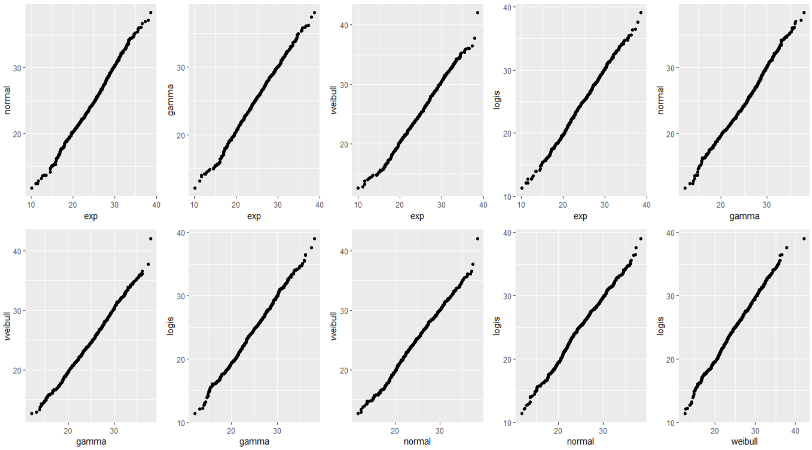


Figure 1: Capon's statistic

Distribution Free

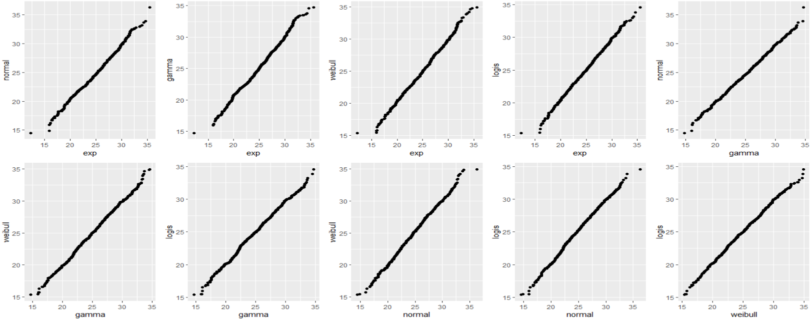


Figure 2: Savage statistic

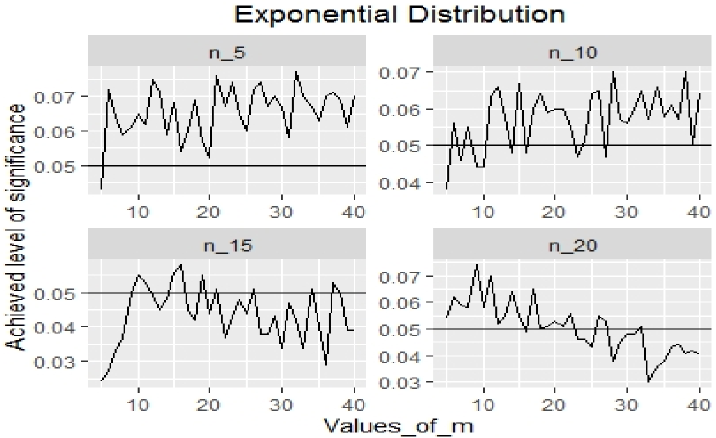
We can observe that the qqplots for statistics from each pair of distributions show the same behaviour. So both the test statistics are distribution free.

Achieved Level of Significance

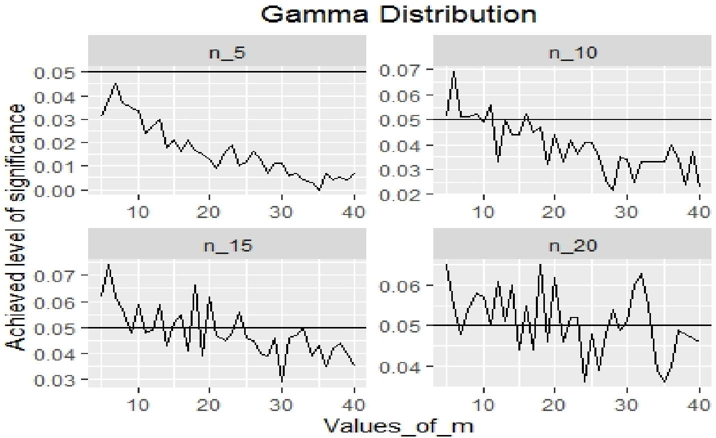
Both Capon and Savage statistics reject H_0 against H_1 for large values of the statistic.

- We simulate the values of Capon and Savage statistics for
  - $n < m$
  - $n > m$
  - $n = m$
- We replicate this simulation 1000 times and get the level of significance achieved using the asymptotic cut-off points.
- We will now observe the behaviour of the level of significance attained for different values of m and n for each of the 5 distributions.

Capon's statistic :

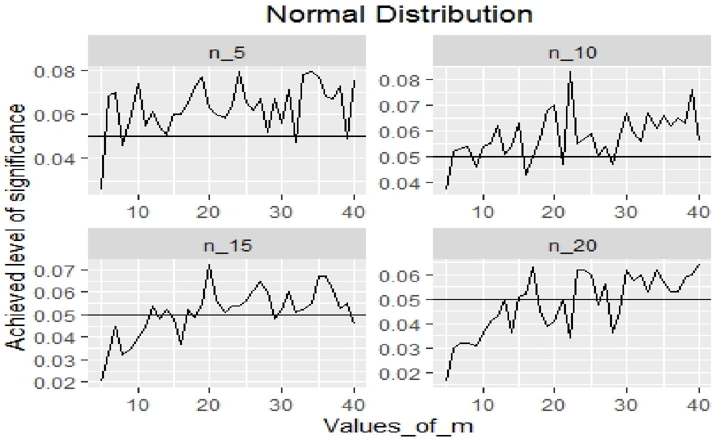


Contd.

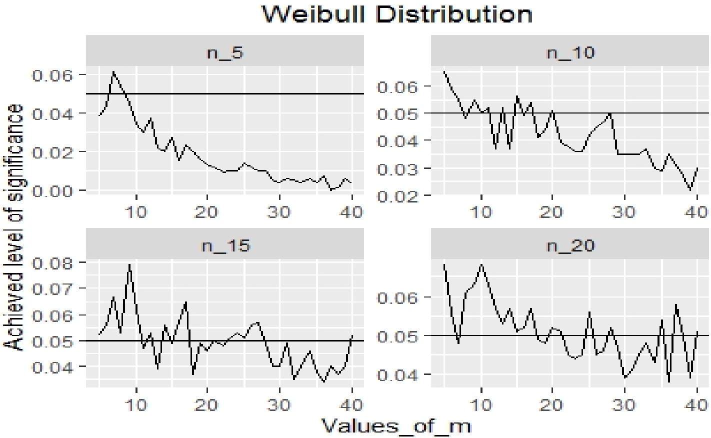


Gamma Distribution
n_5
n_10
n_15
n_20
Achieved level of significance
Values_of_m
0.00
0.01
0.02
0.03
0.04
0.05
0.06
0.07
10
20
30
40

Contd.

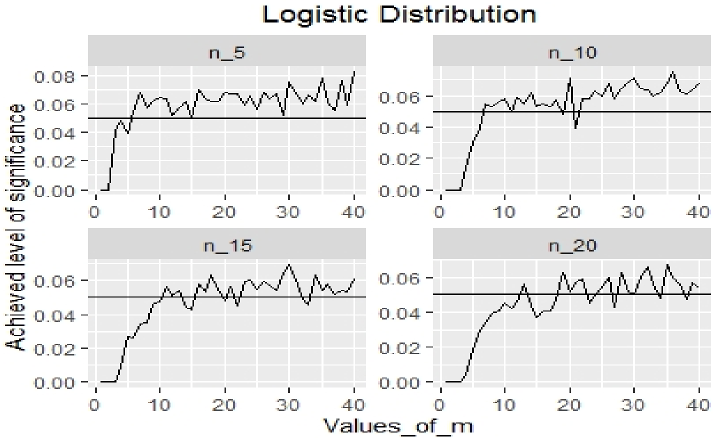

Normal Distribution
n_5
n_10
n_15
n_20
Achieved level of significance
Values_of_m
0.08
0.06
0.04
0.07
0.05
0.03
0.02
10
20
30
40

Contd.



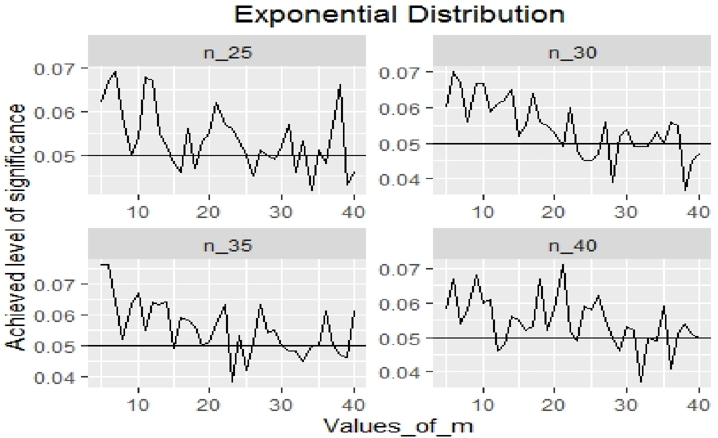
Weibull Distribution
n_5
n_10
n_15
n_20
Achieved level of significance
Values_of_m
0.00
0.02
0.04
0.06
0.03
0.05
0.07
0.08
10
20
30
40

Contd.



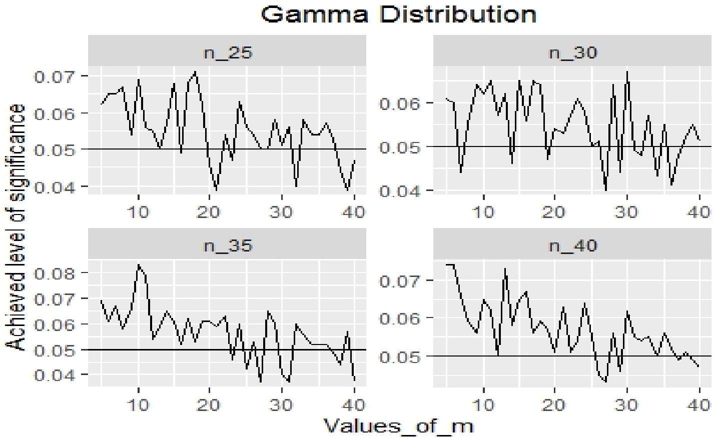
Logistic Distribution
n_5
n_10
n_15
n_20
Achieved level of significance
0.08
0.06
0.04
0.02
0.00
0 10 20 30 40
Values_of_m

Contd.



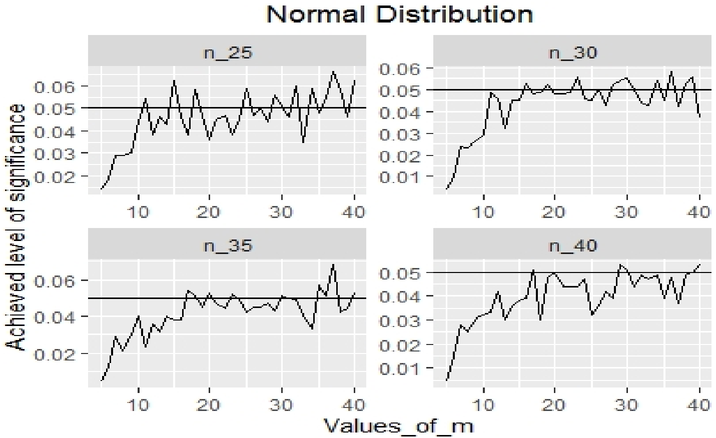
Exponential Distribution
n_25
n_30
n_35
n_40
0.07
0.06
0.05
0.04
10
20
30
40
Achieved level of significance
Values_of_m

Contd.

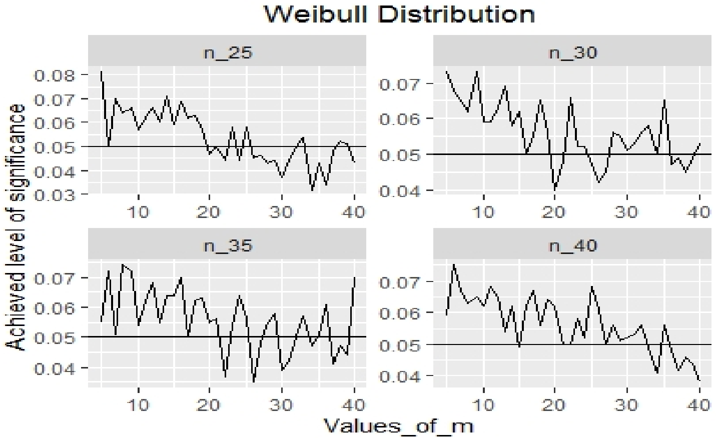


Gamma Distribution
n_25
n_30
n_35
n_40
Achieved level of significance
Values_of_m
0.07
0.06
0.05
0.04
0.08
10
20
30
40

Contd.

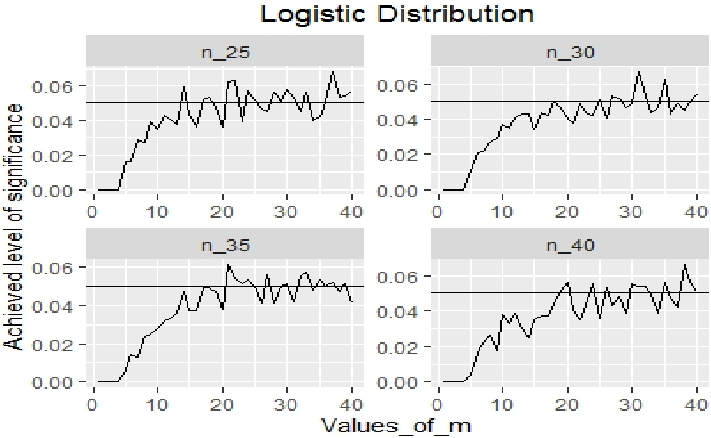

Normal Distribution
n_25
n_30
n_35
n_40
Achieved level of significance
Values_of_m
0.06
0.05
0.04
0.03
0.02
0.01
10
20
30
40

Contd.



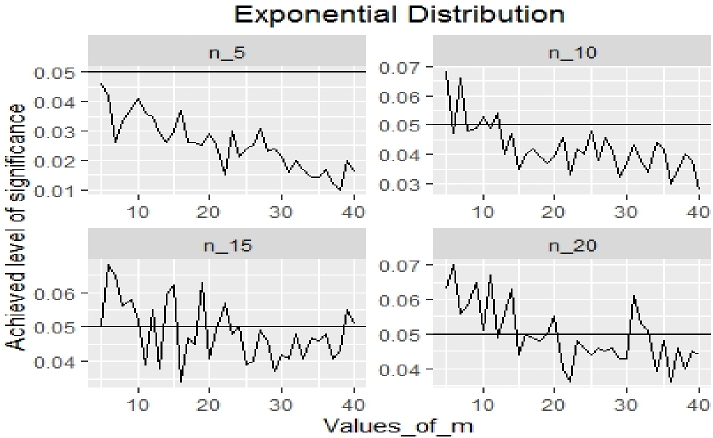
Weibull Distribution
n_25
n_30
n_35
n_40
0.08
0.07
0.06
0.05
0.04
0.03
10
20
30
40
Achieved level of significance
Values_of_m

Contd.

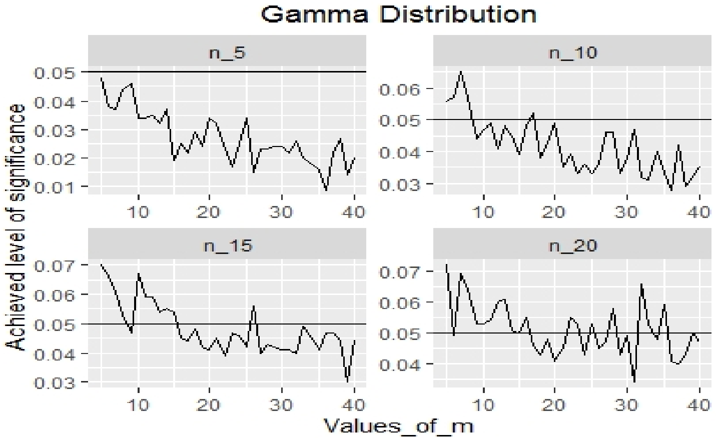


Logistic Distribution
n_25
n_30
n_35
n_40
0.06
0.04
0.02
0.00
0
10
20
30
40
Achieved level of significance
Values_of_m

Savage statistic :

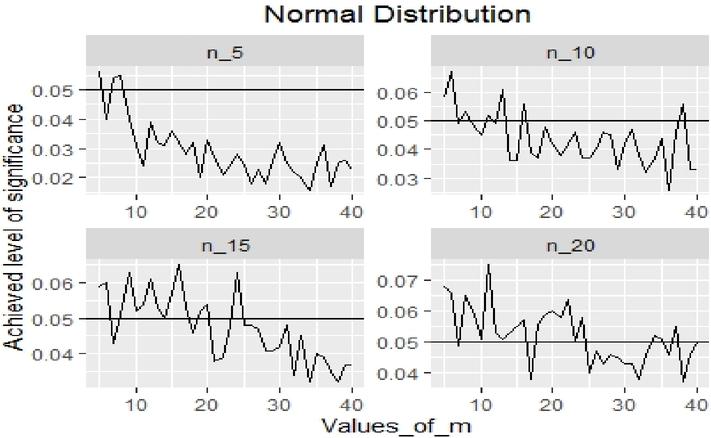


Contd.



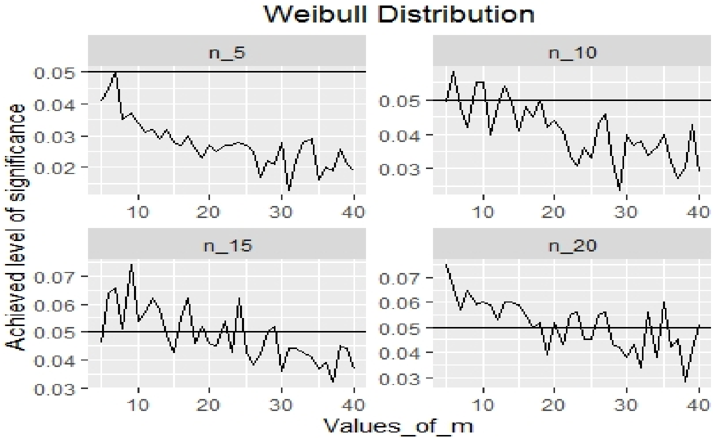
Gamma Distribution
n_5
n_10
n_15
n_20
Achieved level of significance
Values_of_m
0.05
0.04
0.03
0.02
0.01
0.06
0.07
10
20
30
40

Contd.



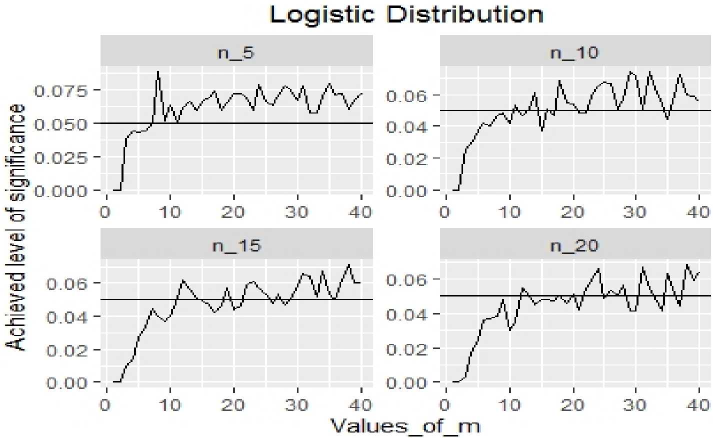
Normal Distribution
n_5
n_10
n_15
n_20
Achieved level of significance
Values_of_m
0.05
0.04
0.03
0.02
0.06
0.07
10
20
30
40

Contd.



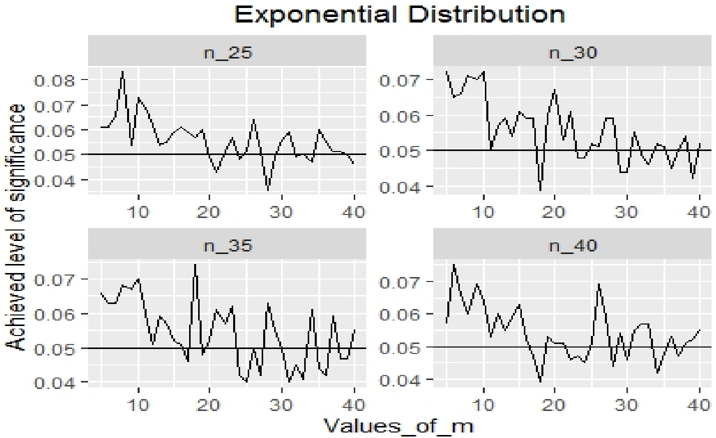
Weibull Distribution
n_5
n_10
n_15
n_20
Achieved level of significance
Values_of_m
0.05
0.04
0.03
0.02
0.07
0.06
10
20
30
40

Contd.



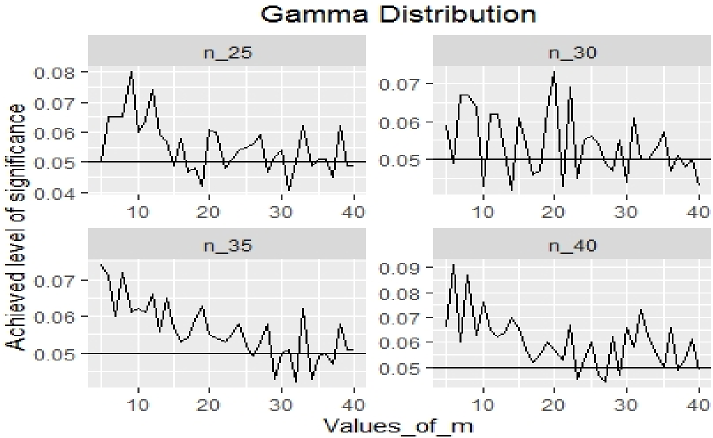
Logistic Distribution
n_5
n_10
n_15
n_20
Achieved level of significance
0.075
0.050
0.025
0.000
0.06
0.04
0.02
0.00
0
10
20
30
40
Values_of_m

Contd.



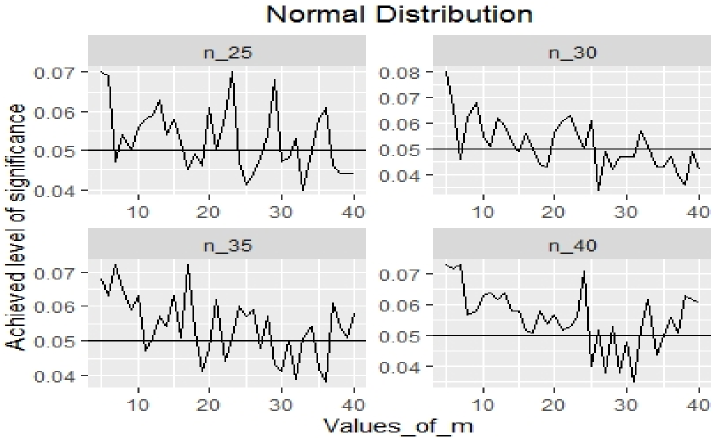
Exponential Distribution
n_25
n_30
n_35
n_40
0.08
0.07
0.06
0.05
0.04
10
20
30
40
Achieved level of significance
Values_of_m

Contd.



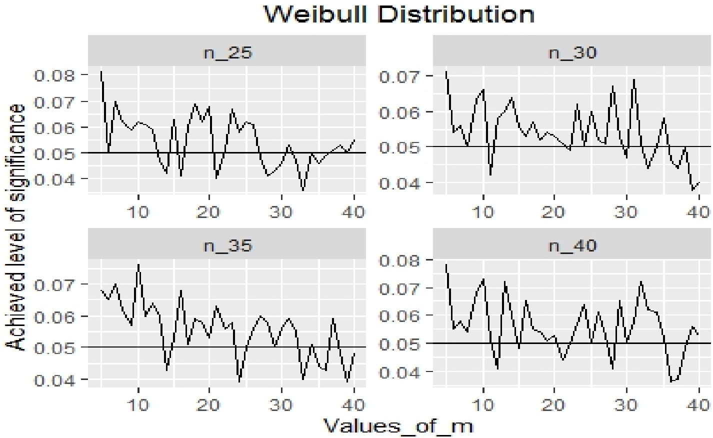
Gamma Distribution
n_25
n_30
n_35
n_40
Achieved level of significance
Values_of_m
0.04
0.05
0.06
0.07
0.08
0.09
10
20
30
40

Contd.



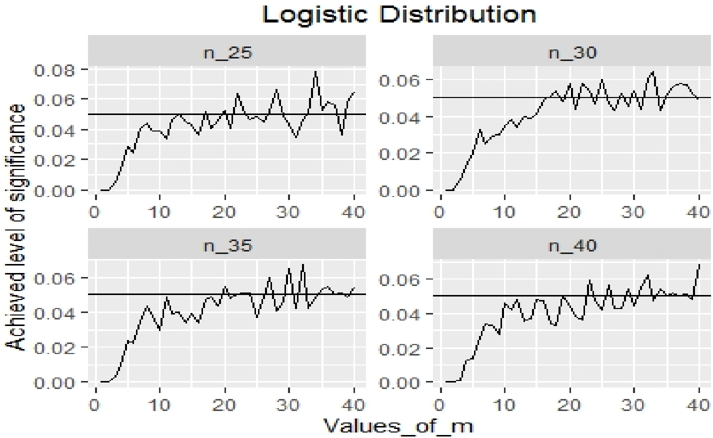
Normal Distribution
n_25
n_30
n_35
n_40
Achieved level of significance
Values_of_m
0.07
0.06
0.05
0.04
0.08
10
20
30
40

Contd.



Weibull Distribution
n_25
n_30
n_35
n_40
0.08
0.07
0.06
0.05
0.04
10
20
30
40
Achieved level of significance
Values_of_m

Contd.

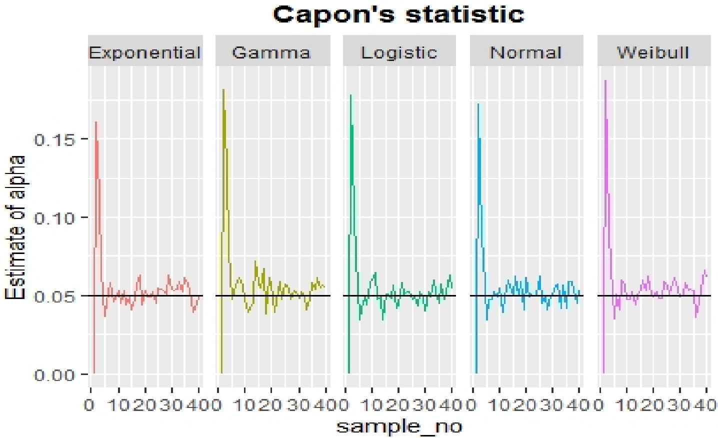


Logistic Distribution
n_25
n_30
n_35
n_40
Achieved level of significance
Values_of_m
0.00
0.02
0.04
0.06
0.08
0
10
20
30
40

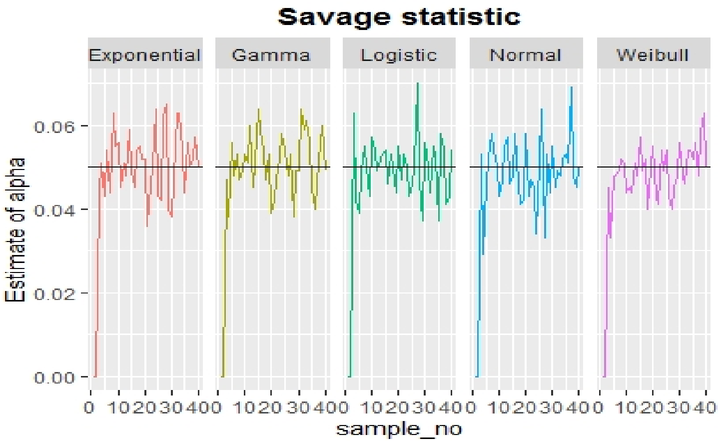
Observations

- We can clearly see that for when $n < m$, the values of achieved level of significance is equal to the desired size when m lies in a very small boundary of n.
- The same happens when $n > m$.
- Let us look at the case $n = m$
- So we can see that the achieved level of significance converges to the pre-fixed size $\alpha = 0.05$

Observations



Observations



Observations

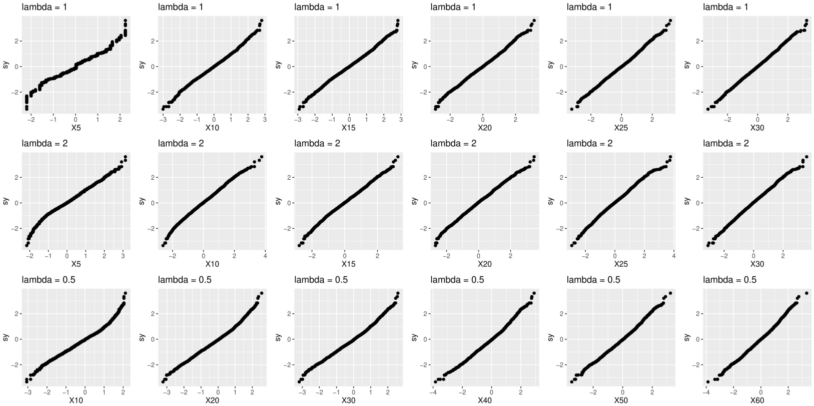
- We can clearly see that for when $n < m$, the values of achieved level of significance is equal to the desired size when m lies in a very small boundary of n.
- The same happens when $n > m$.
- Let us look at the case $n = m$
- So we can see that the achieved level of significance converges to the pre-fixed size $\alpha = 0.05$

Normality Approach

We have already seen the distribution free property of Savage and Capon's statistic. Now we are are interested to see how quickly does the distributions of the statistics converge to normality.
To observe if the statistics are realizing normality, for a given value of m and n, we shall plot the qqplots and note the changes in them,

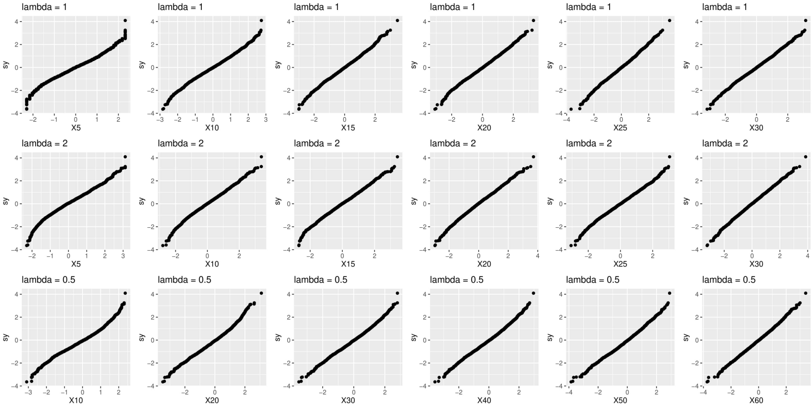
qqplots

Figure 3: Capon Statistic qqplots for different m and n



qqplots

Figure 3: Savage Statistic qqplots for different m and n



qqplots

- Clearly when the total sample size exceeds 50, both the statistics have approximately normal distribution.
- Whether $m > n$ or $n < m$ does not affect the asymptotic distribution of both the statistics

Further to confirm that the distribution of the statistics are normal we have done kolmogorov-smirnov test of normality for $m = 30$ and $n = 30$.

	Capon's Statistic	Savage Statistic
Kolmogorv-Smirnov	0.6287	0.4829

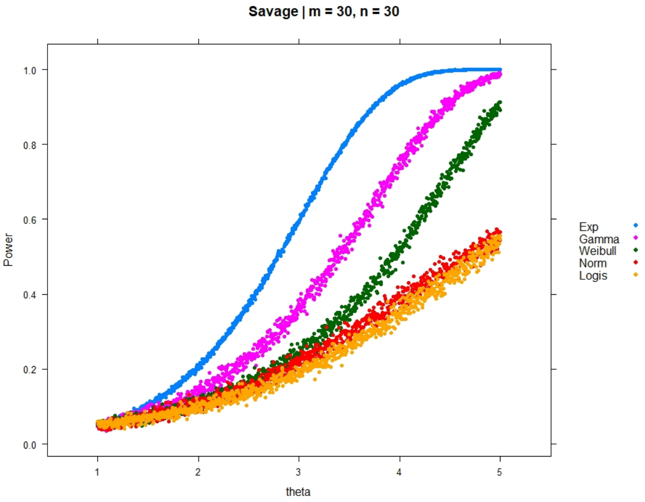
Power function — Savage statistic

We already know, Savage statistic follows a normal distribution, for sufficiently large values of m and n. We can define the critical region using the asymptotic distribution of the Savage Statistic.

The power of the test (using Savage statistic) for different values of θ has been calculated by simulating $X_n's$ and $Y_n's$ 1000 times for each value of θ.

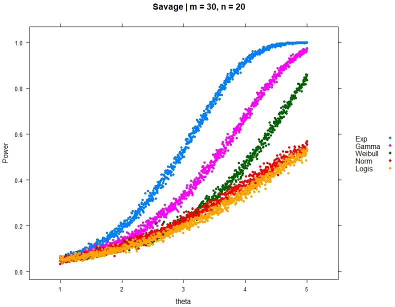
Power plots

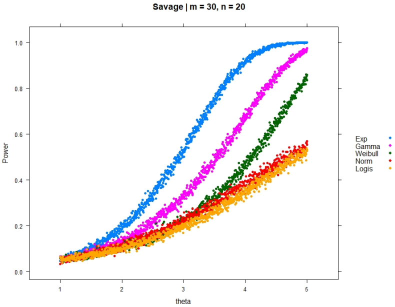
Figure 3: Savage Statistic θ vs Power



Power plots

Figure 3: Savage Statistic θ vs Power





Observation

- The asymptotic behaviour of the distribution is same for $m < n$, $m = n$, $m > n$.
- Savage statistic is most powerful for exponential distribution which is expected since, when the underlying distribution is exponential, the locally most powerful (LMP) test gives rise to Savage statistic.
- Power of savage statistic for the distributions Exponential, Gamma, Weibull, Normal and Logisitic follow the exact order Exponential $>$ Gamma $>$ Weibull $>$ Normal $>$ Logistic for all values of θ.

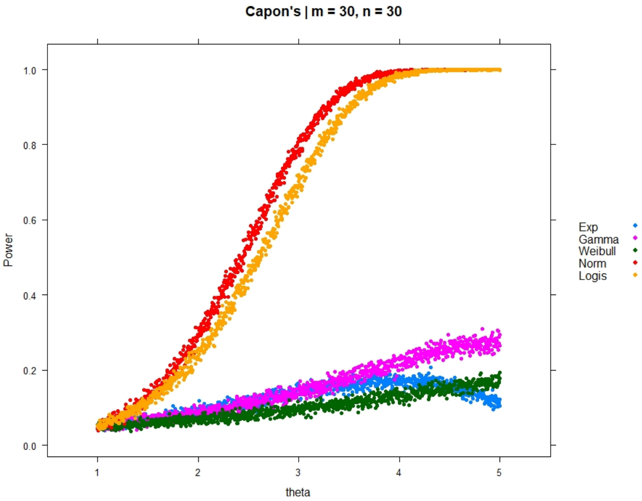
Power function — Capon's statistic

As discussed earlier, Capon's statistic follows a normal distribution, for sufficiently large values of m and n. We can define the critical region using the asymptotic distribution of the Capon's Statistic.

The power of the test (using Capon's statistic) for different values of θ has been calculated by simulating $X_n's$ and $Y_n's$ 1000 times for each value of θ.

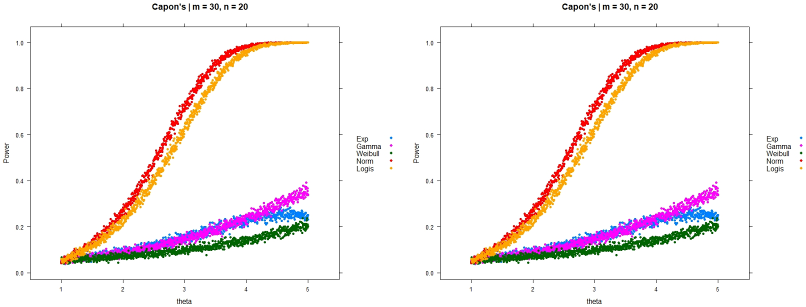
Power plots

Figure 4: Capon Statistic θ vs Power



Power plots

Figure 4: Capon Statistic θ vs Power



Observation

- The asymptotic behaviour of the distribution is same for $m < n$, $m = n$, $m > n$.
- Capon statistic is most powerful for Normal distribution
- Power of Capon's statistic for Normal distribution and logistic distribution is almost same, but the power for normal distribution is always greater than the power for Logistic distribution. This is because Capon's test statistic arises as the locally most powerful test when sample is from normal distribution.
- Capon statistic performs very poorly for Exponential, Gamma and Weibull distribution. This happens because, for positive valued random variables when scale changes, stochastic dominance comes into play, and hence only scale test is not sufficient.

Alternative Approach

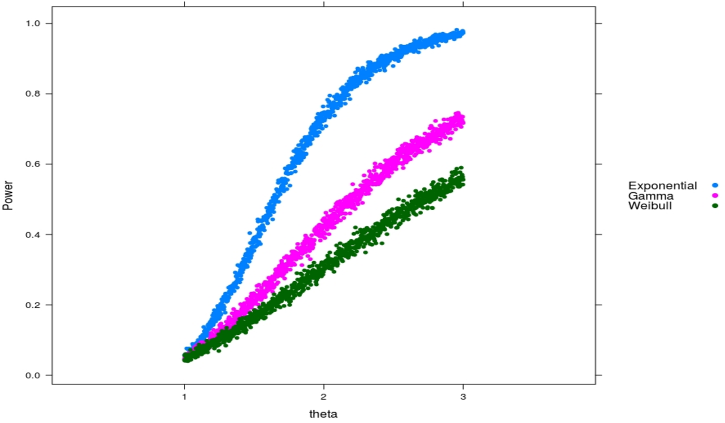
As we have seen that for positive valued random variables, the power graph is not satisfactory. So we have performed Mann-Whitney test for the problem of dispersion.

$$U = \sum_{i=1}^{n} \sum_{j=1}^{m} \phi(X_i, Y_j) = \begin{cases} 1 & \text{if } X_i > Y_j \\ 0 & \text{o.w} \end{cases}$$

We reject for large values of the test statistic.
We have varied values of θ between 1 and 3 and plotted the power and fixed $n = m = 30$ in each case for Exponential, Gamma and Weibull distributions.

Power Graph

Figure 5: Power of Mann-Whitney statistic for positive valued random variables



Parametric Counterpart

- At first we look at the F-statistic for the scale parameter of the normal distribution and focus ourselves how this statistic is affected on changing the basic assumption of the data (i.e. data comes from normal distribution) and taking various underlying distributions like Exponential, Logistic, Gamma and Weibull.
- We do the same thing as stated above for both large sample and small sample size.
- Our next motive is to compare it with our non-parametric test statistics for the given distributions.

F-Statistic

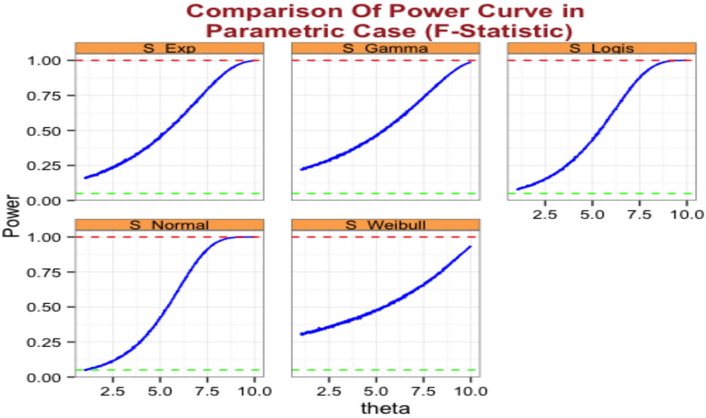
At first we will look at the F statistic for normal distribution. We will perform an F test to compare the variances of two samples from normal populations.Suppose we sample randomly from two independent normal populations. Let σ_x^2 and σ_y^2 be the unknown population variances and s_x^2 and s_y^2 be the sample variances of two independent sample $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_m$ of sizes m and n respectively. Since we are interested in comparing the two sample variances, the F-statistic is:

$$F = \frac{s_x^2}{s_y^2}$$

Small Sample

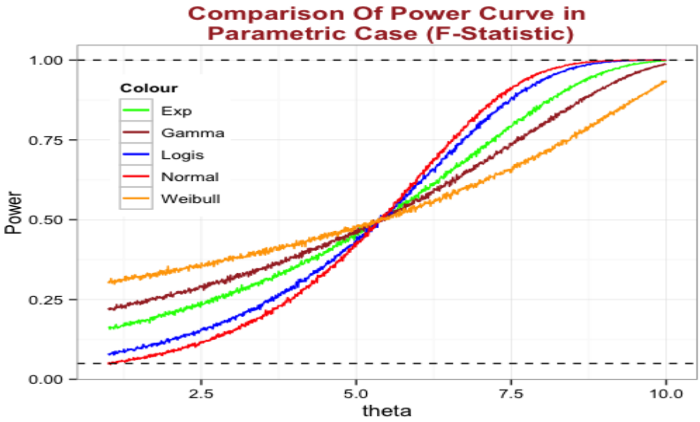
- At first we take a random sample of size 10 from normal distribution with mean 0 and standard deviation 1 and take another random sample of size 10 from another normal distribution mean 0 and standard deviation σ (≤ 1). This is equivalent to our non-parametric case where the scale parameter $\theta > 1$.
- We vary σ from 0.1 to 1 taking 1000 values in between it and perform F-Test and the power function of the test statistic is plotted.
- After completing Normal Distribution we shift to Exponential distribution and so on without changing the test statistic.

Graph



Comparison Of Power Curve in
Parametric Case (F-Statistic)
S_Exp
S_Gamma
S_Logis
S_Normal
S_Weibull
Power
1.00
0.75
0.50
0.25
0.00
2.5
5.0
7.5
10.0
theta

Graph



Comparison Of Power Curve in
Parametric Case (F-Statistic)
1.00
0.75
0.50
0.25
0.00
Power
Colour
Exp
Gamma
Logis
Normal
Weibull
2.5
5.0
7.5
10.0
theta

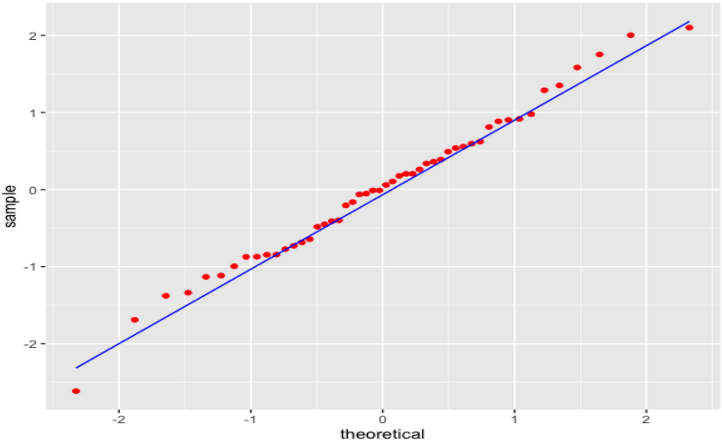
Explanation

- When σ deviates largely from the null hypothesis then for any distribution F-statistic ultimately converges to 1 and unbiased also we can claim the robustness of the F-statistic.
- For the normal distribution F-statistic perform very well power reaches to 1.
- For the exponential distribution F-statistic is not performing well but ultimately it reaches 1. Similar results for Gamma and Weibull distribution with scale parameter 0.5 each.
- No other distributions except normal attempt the size (0.05) of the test.

Large Sample

- At first we take random sample of size 50 from normal distribution with mean 0 and standard deviation 1 and take another random sample of size 50 from another normal distribution mean 0 and standard deviation σ (≤ 1).
- We vary σ from 0.2 to 1 taking 1000 values in between it and perform F-Test and the power function of the test statistic is plotted.
- After completing Normal Distribution we shift to Exponential distribution and so on without changing the test statistic.
- Before proceeding, we at first will check for normality of the data.

Normality



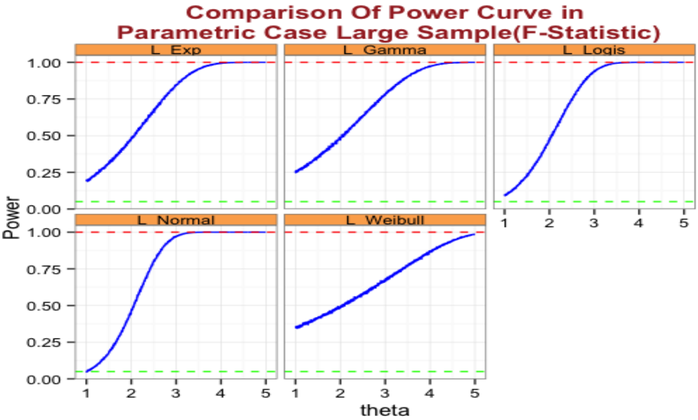
Normality

Hence from the qqplot we can say that this sample is more or less normal and for further check we will test Shapiro-Wilk test and Chisquare Goodness of fit to check it.

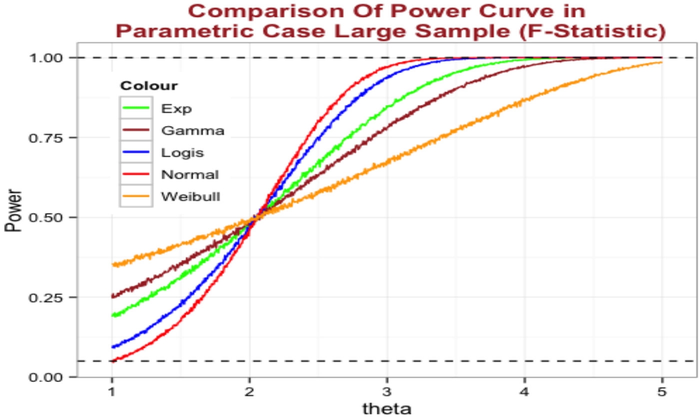
- Shapiro-Wilk test p value: 0.9477
- Chisquare goodness of fit p value: 0.2914

Both are accepted (at the level of significance 0.05) for the sample size m=50 and n=50, so we can proceed our analysis.

Graph


Comparison Of Power Curve in
Parametric Case Large Sample(F-Statistic)
L_Exp
L_Gamma
L_Logis
L_Normal
L_Weibull
Power
1.00
0.75
0.50
0.25
0.00
1
2
3
4
5
theta

Graph

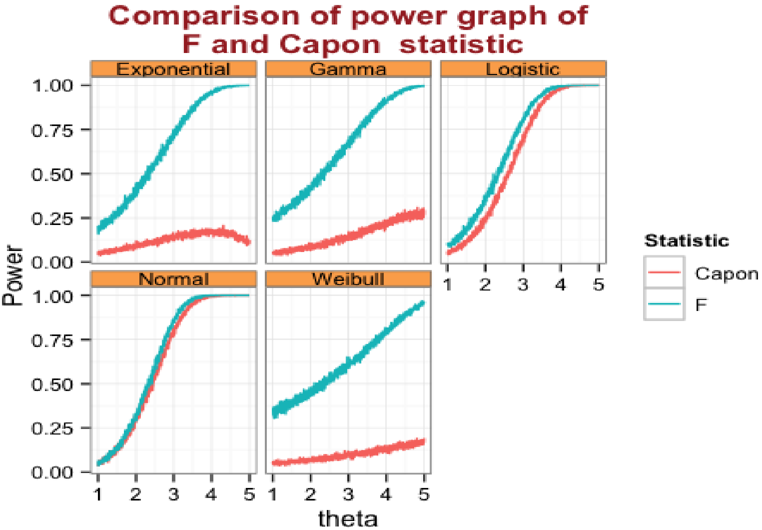


Comparison Of Power Curve in
Parametric Case Large Sample (F-Statistic)
Power
1.00
0.75
0.50
0.25
0.00
1
2
3
4
5
theta
Colour
Exp
Gamma
Logis
Normal
Weibull

Explanation

- The power of the F-statistic for all the distributions approaches 1 rapidly for large samples than small samples.
- When σ decreases significantly from 1 (value under H_0) then for any distribution F-statistic ultimately converges to 1,so we can claim the robustness of the F-statistic.
- The F-statistics for Normal and Logistic distributions perform almost similarly. This can arise because both the distributions have support over the entire real line.
- For the standard exponential and Gamma distribution with scale parameter 0.5, the performance of F-statistic has been satisfactory. The rate of convergence for Weibull distribution with scale parameter 0.5 is much lower than the other distributions.
- No other distributions except normal attain the exact size (0.05) of the test.

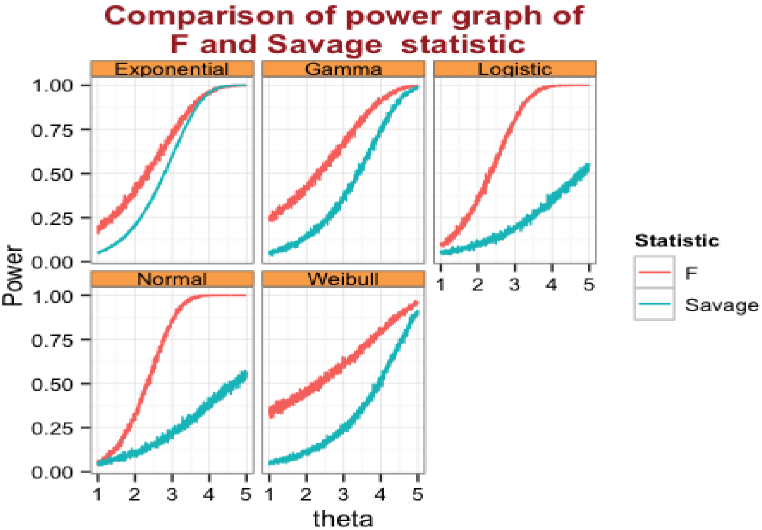
Comparison And Graph



Comparison And Graph

- We can see that Capon's statistic works well when samples are from normal and logistic distribution, the best being for normal distribution. It does not work well when the samples are from other distributions.
- Similarly, we see Savage statistic works well when samples are from exponential, gamma and Weibull distributions but is worse for normal and logistic.
- Finally, we can conclude that F-statistic performs better than the non parametric test statistics for all distributions.

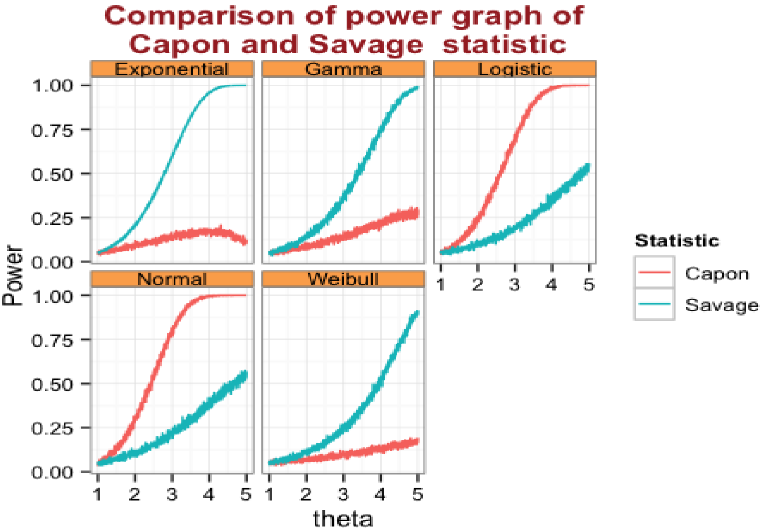
Comparison And Graph



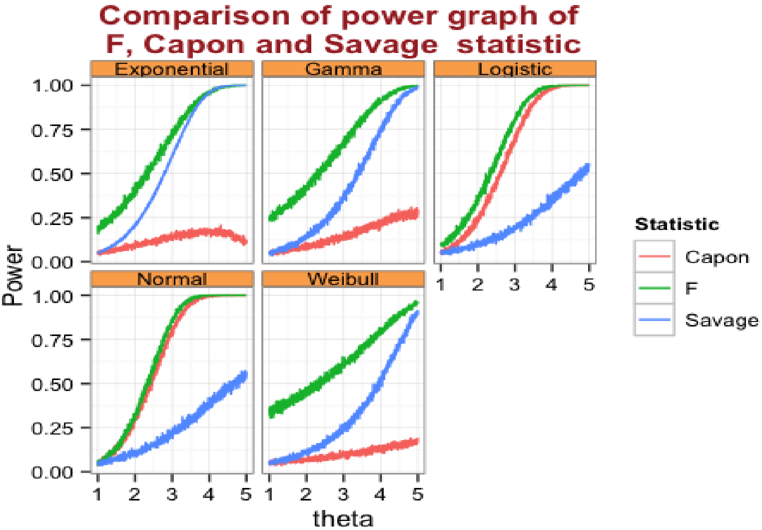
Comparison And Graph

- We can see that Capon's statistic works well when samples are from normal and logistic distribution, the best being for normal distribution. It does not work well when the samples are from other distributions.
- Similarly, we see Savage statistic works well when samples are from exponential, gamma and Weibull distributions but is worse for normal and logistic.
- Finally, we can conclude that F-statistic performs better than the non parametric test statistics for all distributions.

Comparison And Graph



Comparison And Graph



Comparison And Graph

- We can see that Capon's statistic works well when samples are from normal and logistic distribution, the best being for normal distribution. It does not work well when the samples are from other distributions.
- Similarly, we see Savage statistic works well when samples are from exponential, gamma and Weibull distributions but is worse for normal and logistic.
- Finally, we can conclude that F-statistic performs better than the non parametric test statistics for all distributions.

An Observation

Capon's statistic arises as the LMP test when underlying distribution is normal i.e., a distribution with support over the entire real line and Savage statistic came out as the LMP when samples were taken from exponential distribution. So we assume Capon's statistic shows better results for random variables taking both positive and negative values and Savage statistic for positive valued random variables.

Thanking You