# Choosing Appropriate Distribution by Minitab's 17 Software to Analysis System Reliability

**Asst. Prof. Dr. Ahmed Abdulrasool Ahmed**
Department of Mechanical Engineering
College of Engineering
Baghdad University
Email: aa.alkhafaji@yahoo.com

**B.Sc. Layla khudheir Jadwa**
Department of Mechanical Engineering
College of Engineering
Baghdad University
Email: laylaalhamd@yahoo.com

## ABSTRACT

This research aims to choose the appropriate probability distribution to the reliability analysis for an item through collected data for operating and stoppage time of the case study.

Appropriate choice for .probability distribution is when the data look to be on or close the form fitting line for probability plot and test the data for goodness of fit .

Minitab's 17 software was used for this purpose after arranging collected data and setting it in the theprogram.

The program results gave the best or well-fitting distribution among four of default probability distributions, that will use in order to estimate the distribution parameters values, for reliability determination and analysis. From probability plot can estimate time that designates the percent of the item's failure.

**Key words**: Reliability, Life data analysis, Probability distribution, Exponential distributions, Simple Weibull

# اختيار التوزيع المناسب باستخدام برنامج (Minitab's 17 software) لتحليل معولية منظومة

**ا.م.د. احمد عبد الرسول الخفاجي**
قسم الهندسة الميكانيكية
كلية الهندسة / جامعة بغداد

**ر. م: ليلى خضير جدوع**
قسم الهندسة الميكانيكية
كلية الهندسة / جامعة بغداد

## الخلاصة

هذا البحث يهدف الئ اختيار التوزيع الاحتمالي المناسب لتحليل المعولية لمعدة من خلال بيانات مجمعة لاوقات اشتغال وتوقف معدة قيد الدراسة ,الاختيار الافضل للتوزيع الاحتمالي يكون للبيانات التي على او قريبة من الخط المستقيم الرابط للبيانات في مخطط الاحتمالية و اختبار البيانات لجودة التوفيق.

وظف برنامج (Minitab's 17) لهذا الغرض بعد ترتيب البيانات المجمعة وادخالها الى البرنامج.

نتائج البرنامج اعطت التوزيع الاحتمالي الصحيح من بين اربعة من التوزيعات الاحتمالية المفترضة والذي يستخدم في تخمين قيم متغيرات التوزيع لايجاد وتحليل المعولية, ومن مخطط الاحتمالية يمكن تخمين الوقت الذي يمثل النسبة المئوية لفشل المعدة .

**الكلمات الرئيسيه:** معولية, تحليل بيانات العمر, التوزبع الاحتمالي, التوزيع الاسي, توزيع ويبل البسيط.

## 1. INTRODUCTION

The increasing in complexity of equipment and systems lead to failures and as a result for the aspects of reliability, maintainability and availability have appeared into forefront. The failure of machineries and equipment causes interruption in production as a resulting from a loss of availability **Bose, et al., 2013**. Design for reliability is an important research area **Soleimania and Mohammad, 2014.** In life data analysis, statistical distribution represents the failure behavior of the equipment population through time, and subsequently it is possible to calculate the reliability indices of the equipment **Iqbal Ridwan, et al., 2010**. The term life data refers to the measurements of the lifetime of the equipment, whether in hours, years or cycles **Rausand , and Hoylanc, 2004** The repair of a system will initially sound like replacing a failed part with another one **Manortey, 2006.** the most common metric used to represent the reliability of repairable systems is Mean Time Between Failures" MTBF or "Average Run Life", that is used to characterize "life span" **MichelleI , 2011**, which is calculated by adding all the operating hours of all the systems and dividing by the number of failures. The popularity of the MTBF metric is due to its simplicity and its skill to supply to the one number syndrome, but MTBF hides information by not accounting for any trends in the appearance of failures and treating machines of all ages as coming from the same population

The probability plot represents the sum of the frequencies from the lowest value up to the considered point. The cumulative curve is the integral of the density function **Venkataramana.,et al., 2013**

## 2. THEORY OF RESEARCH

The theoretical population models used to describe unit lifetimes are known as Lifetime Distribution Models **Venkataramana.,et al., 2013**. The model selection tests were applied based on statistical analysis **Bose,2013.**

Minitab's 17 software used for the Distribution Identification and as a good tool can help to find the distribution that best fits the data.

A lifetime distribution model can be any probability density function (PDF) $f(t)$ defined over the range of time from $t = 0$ to $t =$ infinity. The corresponding cumulative distribution function (CDF) $F$(t) is a very useful function, as it gives the probability that a randomly selected

1. $F(t)$ = the area under the PDF $f(t)$ to the left of $t$.

2. $F(t)$ = the probability that a single randomly chosen new unit will fail by time $t$.

3. $F(t)$ = the proportion of the entire population that fails by time $t$**. Iqbal Ridwan,et al., 2010,Venkataramana.,et al., 2013**

PDF $f(t)$ has only non-negative values and eventually either becomes 0 as $t$ increases, or decreases towards 0, The CDF $F(t)$ is increasing and goes from 0 to 1 **Ronniger, C. 2012** as $t$ approaches infinity, in other words, total area under the curve is always 1.**Iqbal Ridwan,2010**

Probability plots are simple visual ways of summarizing reliability data by plotting CDF estimates vs. time on specially constructed probability paper **Venkataramana, P. 2013**.

### 2.1 Anderson Darling

The Anderson-Darling (AD) statistic measures how well the data follow a particular distribution.

For a specified data set and distribution, the better the distribution fits the data, the smaller this statistic will be, Goodness of Fit (GOF) test to determining how well a given curve can be modeled by a distribution.

In statistics, there is GOF measure including AD (used in this paper) which may give powerful test of fit **MichelleI P, 2011**, GOF test was evaluated and also whether the data was come from the same probability distributions was checked. The analysis based on AD criteria, Minitab calculates the AD statistic using either (maximum likelihood estimation method or least squares estimation) **Engmann, S .2011**

## 2.2 The Exponential Distribution

The exponential distribution is a continuous distribution that is used to represent the constant failure rates. It is characterized by the parameter λ which represents the failure rate of the item The failure rate λ of the exponential distribution is constant with respect to time **Rausand , M.2004,MichelleI P., 2011.**

$\theta$ = The distribution mean

$\lambda = 1/\theta$ = The failure rate (1)

$\lambda = 1 / MTBF$

The exponential reliability function is

$$R(x) = e^{-\frac{t}{\theta}} = e^{-\lambda t} \quad , t \geq 0 \tag{2}$$

The cumulative distribution function

$$F(t) = 1- e^{-\lambda t} \tag{3}$$

The exponential probability density function is

$$f(x) = \frac{1}{\text{MTBF}} e^{\frac{-t}{\theta}} = \lambda e^{-\lambda t} \qquad t \geq 0 \tag{4}$$

Where: $\text{MTBF} = \frac{1}{\lambda} = \theta$

The exponential hazard function is constant

$$h(x) = \frac{f(t)}{R(t)} = \frac{1}{\theta} = \lambda \tag{5}$$

## 2.3 The Weibull Distribution

Weibull distribution is one of the most widely used life distributions in reliability analysis, very flexible and can through parameters, models many types of failure rate behaviors so it’s known with two parameters **,mIqbal Ridwan,et al., 2010, Ronniger , 2012.**

- Scale parameter $\eta > 0$
- Shape parameter $\beta > 0$

Survival Function

$$R(t)= pr\ (t > 0\ ) = e^{-\left(\frac{t}{\eta}\right)^{\beta}} \quad (6)$$

Cumulative Distribution Function

$$F(x) = 1 - e^{-\left(\frac{t}{\eta}\right)^{\beta}} \qquad t \geq 0:\ \beta > 0 \quad (7)$$

Probability density function

$$f(t) = \frac{\beta(t)^{\beta-1}}{\eta^{\beta}}\ e^{-\left(\frac{t}{\eta}\right)^{\beta}} \quad (8)$$

Hazard Function

$$h\ (t)= \frac{\beta(t)^{\beta-1}}{\eta^{\beta}} \qquad t \geq 0:\ \beta > 0 \quad (9)$$

$\beta = 1$ Failure rate is constant

$\beta > 1$ Failure rate function is increasing

$\beta < 1$ h (t) is decreasing **(Bose, et al., 2013)**.

$\beta = 2$ The resulting distribution is known as (Rayleigh Distribution) (**Venkataramana, P., 2013).**

## 3. APPLICATION

**Fig. (1)** indicated the methodology for research to reach its aim. The following steps for analysis data indicated the used path to find the suitable probability distribution to fit life data and estimate the parameters of distributions.

1. The reliability was studied through a compressor of Gasoline Reformer unit in Daura Refinery in Baghdad, this has helped to optimize the schedule for preventive maintenance.
2. The exact failure time data was collected for a part under study measured the time by hours that included operating and stoppage time for six years from (2005 to 2010).
3. Table (1) life data was sorting and arranged to be suitable to use in Minitab's 17 software application, included [From , To, TTR] in [hr.] as columns.
4. Data, table (1) was input to Minitab's 17 software for Reliability and Distribution analysis, this would discover which distribution that the recorded time to failure utilizing "Distribution ID Plot" function **Fig. (2)** shows the worksheet of Minitab's 17 software included the menus for software.
5. Output of Minitab's 17 software were, four probability distributions **Fig. (3)** as identical distributions (ID), (notice the software is able to analysis eleven probability distributions) and Session window with table of statistics **Fig. (4),** was summarized in **table (2).**
6. According to resultes the Weibull and Exponential distributions were good, and it is really the Exponential distribution is a special case of the Weibull distribution with shape parameter β = 1 but for constant failure rate .

From **Fig . (3)**:

- Distribution: four Distributions to select which plot has a majority of the data points closest to the line (representing the specific distribution trend).
- Anderson-Darling statistic measures for Goodness-of-Fit was shown how well the data follow a particular distribution, the better the distribution fits the data, the smaller this statistic will be, use the Anderson-Darling statistic to compare the fit of several distributions to see which one is best or to test whether a sample of data comes from a population with a specified distribution.
- Mean, MTTFs allow seeing how conclusions may change with different distributions, from session window **Fig. (3).**

## 4. RESULTS

For Weibull distribution the Minitab analysis provides displays of the future prediction of failures data behavior, to obtain equations those represent the Survival (reliability) function Cumulative failure probability function, Probability density function, and Hazard function, equations using the shape value of 1.45234 and scale value of 5681.08 hr. those were estimated by the maximum likelihood in **Fig.(5**) as substitutions for $\beta$ and $\eta$ respectively, the formulas were determined equations 6,7,8,9, and Minitab's 17 software was used to calculate Probability density, Cumulative distribution , and Hazard functions, .

Percentiles split the data set into parts, nth percentile has n% of the observations lower than it, and (100-n) % of observations higher than it. Percentiles divide a data set into 100 parts (data set arrange into an ascending order in Session window included the output Percentiles values of 1 5 10 50.

Confidence Interval for MTBF for 95%, which means acceptable risk of error $\alpha = 0.05$ lower and upper with standard error for means.

The exponential distribution functions **Fig.(6**) were determined by using equations 2, 3,4,5, with the estimated mean of distribution that equal 5110.37 hr. and the formulas values calculated by using Minitab's 17 software

## 5. DISCUSSION

- Life data analysis is able to provide the information on the reliability indices such as reliability, unreliability, mean time between failure and failure rate**.**
- The estimate MTBF of distribution was 5110.37 hr. equal to calculate mean for Exponential distribution but not equal for Weibull
- All calculations are computed for Weibull and Exponential distribution. Comparing the various results show that the Weibull distribution for upper phase .and Exponential distribution is suitable for lower phase of data life.
- **Fig.(6)** Probability density function (PDF): It is exponentially decreasing curve shows that the likelihood for item to have increasing time to failure values for the beginning time to failure in comparison to the probability of having less time to failure values for the last times to failure .
- **Fig. (6)** Survival (Reliability) Plot: This curve as the (PDF) showing that the proportion of surviving time for item decreases exponentially with time.
- **Fig. (6)** Hazard plot: As is the case here, for an Exponential distribution, the Hazard function displaying the instantaneous time failure rate is constant.

- **Fig. (5)**, for (PDF) and Survival plots for the Weibull distribution may be identical to the Exponential distribution except of the Hazard plot.
- **Fig. (5 )** Hazard plot: For a Weibull distribution the Hazard function can take many shapes, and in this case, the trend is .the hazard rate is increasing over time, which means that the item was more likely to fail as it age.
- η is the value in time by which 63.2% of all failures will have occurred. In this sense, η is one point on the time scale, providing some standard measure of the distribution of times to failure.

## REFERENCE


- Bose, D. Ghosh, G. Mandal,K. Sau S.P. and Kunar, S, 2013 *Measurement and Evaluation of Reliability, Availability and Maintainability of a Diesel Locomotive Engine* International Journal of Engineering Research and Technology, ISSN 0974-3154 Vol. 6, No. 4, pp. 515-534.
- Engmann, S and Cousineau, D. 2011, *Comparing Distribution: The Two-Sample. Anderson Darling Test as an Alternative to The Kolmogorov- Smirnoff Test* Journal of Applied Quantitative Methods Vol. 6 No. 3.

- Iqbal Ridwan, M, Kerk, L Y, Musa, A, Yunus, B ,(2010)., *Application of Life Data Analysis for the Reliability Assessment of Numerical Over current Relays* , International
  - Journal of Electrical, Robotics, Electronics and Communications Engineering Vol. 4 No:12, pp.45 -51.

- Manortey, S. O. 2006, *Life Data Analysis of Repairable Systems. A Case study on Brigham Young University Media Rooms*, (Master of Science Brigham Young University).

- MichelleI P, 2011, "*Electrical Submersible Pump Survival Analysis* Master's Degree Candidate.

- Rausand , M. and Hoylanc, A., (2004) *System Reliability Theory*, Models and Statistical Methods Second Edition, Wiley Series in Probability and Statistics, A John Wiley & Sons, INC., Publication,

- Soleimania, M and -Mohammad, M, P, 2014, *Design for Reliability of Complex System with Limited Failure Data; Case Study of a Horizontal Drilling Equipment* Probabilistic Safety Assessment and Management PSAM 12

- Venkataramana, P., Gurumahesh G..,and Ajay V. 2013, *Reliability Evaluation of Repairable Complex Systems an Analyzing Failure Data* , Int. J. Mech. Eng Vol. 2, No. 1.

- Ronniger, C. 2012 *Reliability Analyses with Weibull* www.crgraph.com.

**Notation and Acronyms**

AD Anderson-Darling
CDF Cumulative distribution function
From Start time, hr.
GOF Goodness-of-fit
ID Identical distribution
MTBF Mean time between failures, hr
PDF Probability density function
TTR Time to repair, hr.
$\alpha$ Acceptable risk of error= 1- Confidence Level
$\mathcal{B}$ Shape parameter
R Reliability
$\theta$ The distribution mean
$\eta$ Scale parameter, hr.
$\lambda$ The failure rate 1/hr.

**Table 1.** Operation time for compressor

| From [hr.] | To [hr.] | TBF[[hr.] | TTR [hr.] |
|---|---|---|---|
| 0.0000 | 1175 | 1175 | 0.5 |
| 1175.5 | 4887.5 | 3712 | 0.5 |
| 4888 | 9931 | 5043 | 2 |
| 9933 | 22953.5 | 13020.5 | 26.5 |
| 22980 | 25544.25 | 2564.25 | 24 |
| 25568.25 | 30715.75 | 5147.5 | 1.5 |

**Table 2.** Summarized outputs from Minitab ̀s 17 software.

| Distribution | AD | Mean | Fitness | Notes |
|---|---|---|---|---|
| Smallest Extreme Value | 2.567 | 4647.35 | Bad | Estimated Mean≠ Calculated Mean |
| Normal | 2.479 | 5110.38 | Bad | Estimated Mean = Calculated Mean |
| Weibull | 2.220 | 5149.56 | Best fitness Smallest AD | Estimated Mean≠ Calculated Mean |
| Exponential | 2.384 | 5110.37 | Good | Estimated Mean = Calculated Mean |

**Table 3.** Functions estimation values for Weibull distribution.

| Time [hr] | R(t) w | F(t). | PDF. | h(t) |
|---|---|---|---|---|
| 0 | 1.000 | 0.000 | 0.0000E+00 | 0.0000E+00 |
| 500 | 0.971 | 0.029 | 8.2571E-05 | 8.5022E-05 |
| 1000 | 0.923 | 0.077 | 1.0744E-04 | 1.1640E-04 |
| 2000 | 0.803 | 0.197 | 1.2798E-04 | 1.5937E-04 |
| 2500 | 0.738 | 0.262 | 1.3019E-04 | 1.7633E-04 |
| 3000 | 0.673 | 0.327 | 1.2898E-04 | 1.9152E-04 |
| 3500 | 0.610 | 0.390 | 1.2524E-04 | 2.0538E-04 |
| 4000 | 0.548 | 0.452 | 1.1968E-04 | 2.1820E-04 |
| 4500 | 0.490 | 0.510 | 1.1285E-04 | 2.3016E-04 |
| 5000 | 0.436 | 0.564 | 1.0520E-04 | 2.4142E-04 |
| 7000 | 0.258 | 0.742 | 7.2573E-05 | 2.8119E-04 |
| 8000 | 0.193 | 0.807 | 5.7688E-05 | 2.9873E-04 |
| 10000 | 0.103 | 0.897 | 3.3997E-05 | 3.3053E-04 |
| 13000 | 0.036 | 0.964 | 1.3323E-05 | 3.7226E-04 |

**Table 4.** Functions estimation values for exponential distribution

| Time[hr.] | R (t) Exp. | F(t) | PDF | h(t). |
|---|---|---|---|---|
| 0 | 1.0000 | 0.0000 | 1.9568E-04 | 1.9568E-04 |
| 500 | 0.9068 | 0.0932 | 1.7744E-04 | 1.9570E-04 |
| 1000 | 0.8223 | 0.1777 | 1.6090E-04 | 1.9570E-04 |
| 2000 | 0.6761 | 0.3239 | 1.3231E-04 | 1.9570E-04 |
| 2500 | 0.6131 | 0.3869 | 1.1997E-04 | 1.9570E-04 |
| 3000 | 0.5560 | 0.4440 | 1.0879E-04 | 1.9570E-04 |
| 3500 | 0.5041 | 0.4959 | 9.8652E-05 | 1.9570E-04 |
| 4000 | 0.4572 | \0.5428 | 8.9457E-05 | 1.9570E-04 |
| 4500 | 0.4145 | 0.5855 | 8.1119E-05 | 1.9570E-04 |
| 5000 | 0.3759 | 0.6241 | 7.3558E-05 | 1.9570E-04 |
| 7000 | 0.2542 | 0.7458 | 4.9736E-05 | 1.9570E-04 |
| 8000 | 0.2090 | 0.7910 | 4.0896E-05 | 1.9570E-04 |
| 10000 | 0.1413 | 0.8587 | 2.7651E-05 | 1.9570E-04 |
| 13000 | 0.0786 | 0.9214 | 1.5373E-05 | 1.9570E-04 |

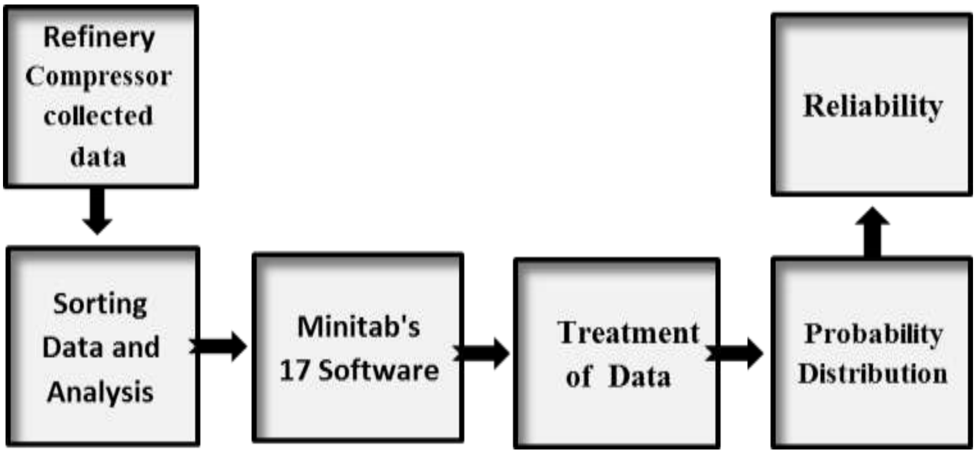


**Figure 1.** Steps for analysis.

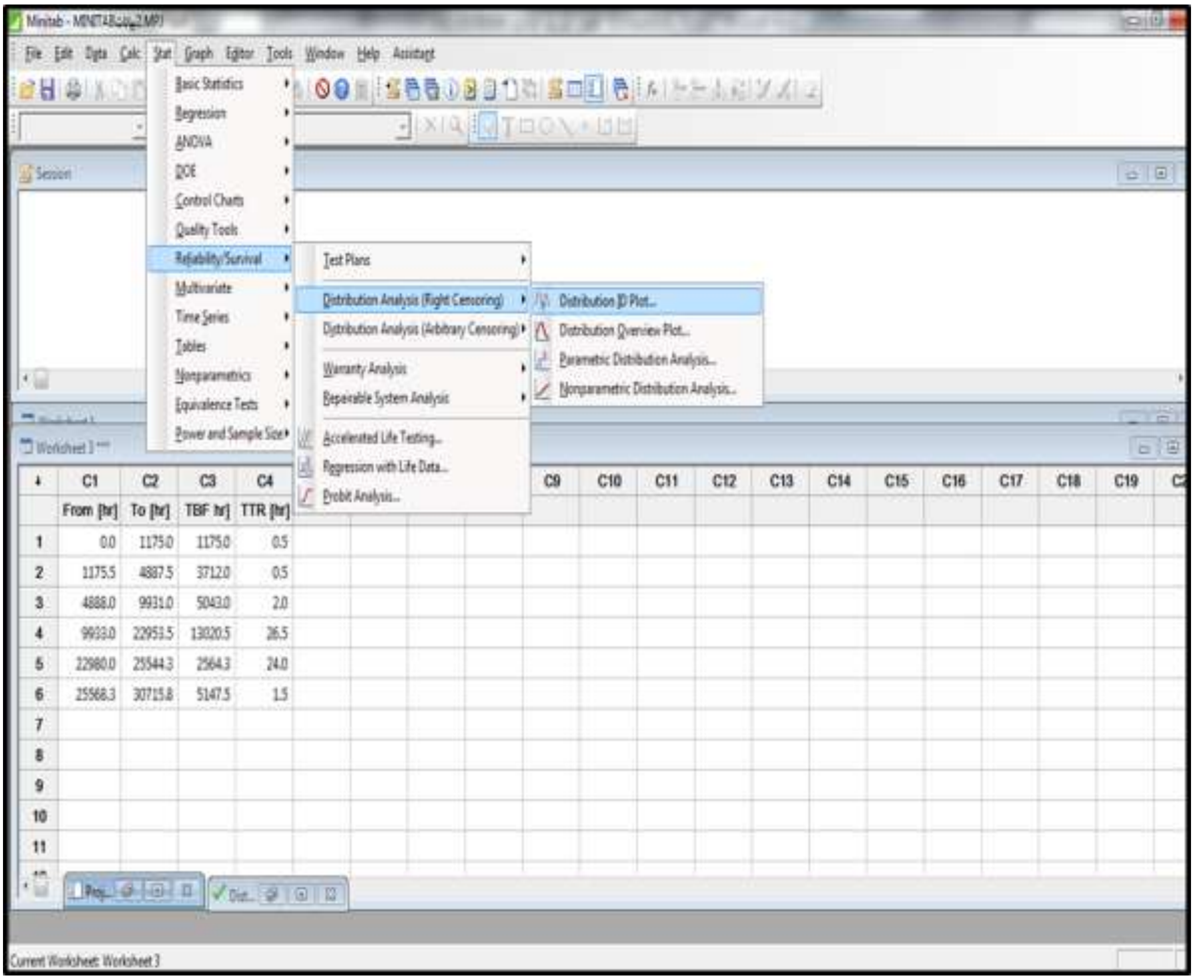


**Figure 2.** Worksheet for reliability analysis for system.

## Results for: Worksheet 3

## Distribution ID Plot: TBF hr]

Goodness-of-Fit

| Distribution | Anderson-Darling (adj) |
|---|---|
| Smallest Extreme Value | 2.567 |
| Normal | 2.479 |
| Weibull | 2.220 |
| Exponential | 2.384 |

Table of Percentiles

| Distribution | Percent | Percentile | Standard Error | 95% Normal CI Lower | 95% Normal CI Upper |
|---|---|---|---|---|---|
| Smallest Extreme Value | 1 | -13231.1 | 6805.64 | -26569.9 | 107.753 |
| Normal | 1 | -3722.67 | 2984.08 | -9571.35 | 2126.01 |
| Weibull | 1 | 239.716 | 264.196 | 27.6421 | 2078.85 |
| Exponential | 1 | 51.3610 | 20.9680 | 23.0745 | 114.323 |
| Smallest Extreme Value | 5 | -5987.35 | 4827.95 | -15450.0 | 3475.26 |
| Normal | 5 | -1135.07 | 2377.66 | -5795.20 | 3525.07 |
| Weibull | 5 | 735.870 | 570.892 | 160.855 | 3366.41 |
| Exponential | 5 | 262.128 | 107.013 | 117.764 | 583.465 |
| Smallest Extreme Value | 10 | -2788.34 | 3989.58 | -10607.8 | 5031.09 |
| Normal | 10 | 244.377 | 2091.88 | -3855.64 | 4344.39 |
| Weibull | 10 | 1207.59 | 769.412 | 346.400 | 4209.78 |
| Exponential | 10 | 538.432 | 219.814 | 241.896 | 1198.48 |
| Smallest Extreme Value | 50 | 5583.73 | 2144.37 | 1380.85 | 9786.62 |
| Normal | 50 | 5110.38 | 1550.10 | 2072.23 | 8148.52 |
| Weibull | 50 | 4414.69 | 1463.37 | 2305.40 | 8453.84 |
| Exponential | 50 | 3542.24 | 1446.11 | 1591.39 | 7884.60 |

Table of MTTF

| Distribution | Mean | Standard Error | 95% Normal CI Lower | 95% Normal CI Upper |
|---|---|---|---|---|
| Smallest Extreme Value | 4647.35 | 2299.40 | 140.60 | 9154.1 |
| Normal | 5110.38 | 1550.10 | 2072.23 | 8148.5 |
| Weibull | 5149.56 | 1476.19 | 2936.03 | 9031.9 |
| Exponential | 5110.37 | 2086.30 | 2295.89 | 11375.1 |

## Distribution ID Plot for TBF [hr.]

**Figure 3.** Session window included the output of program.

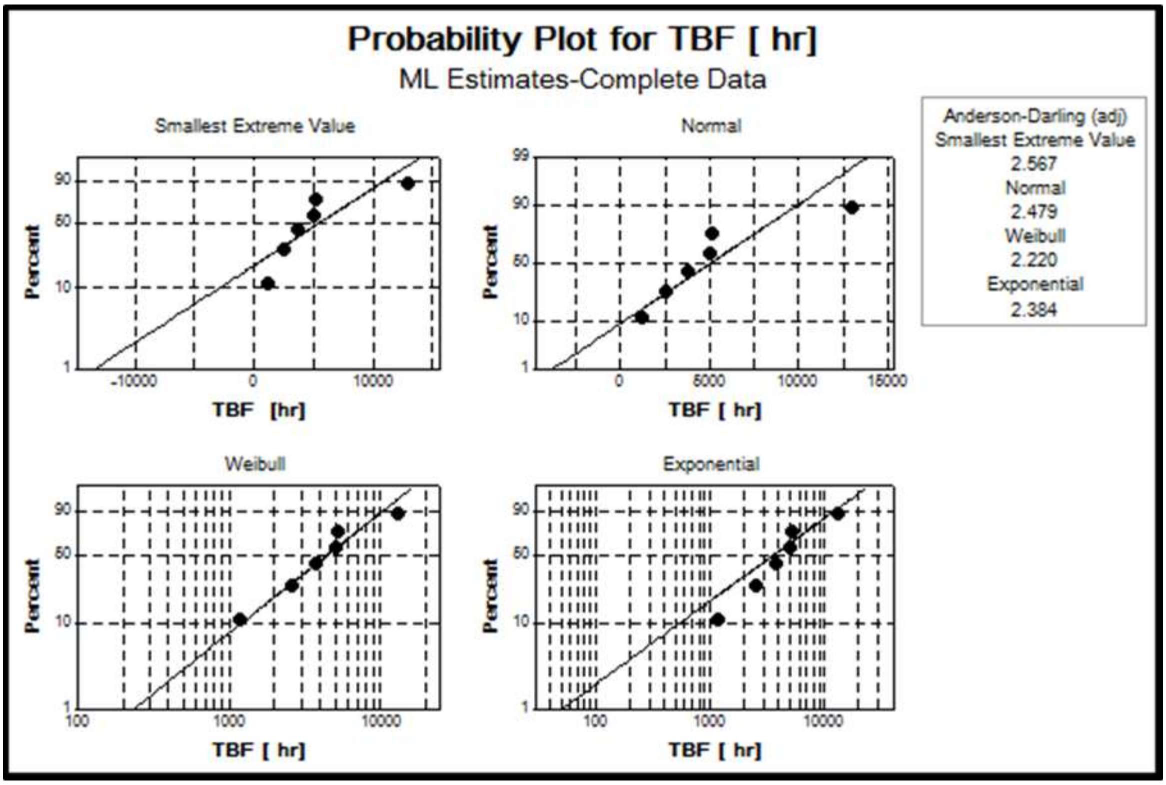


**Figure 4.** Probability Plot of the Failures and the values of Anderson Darling.

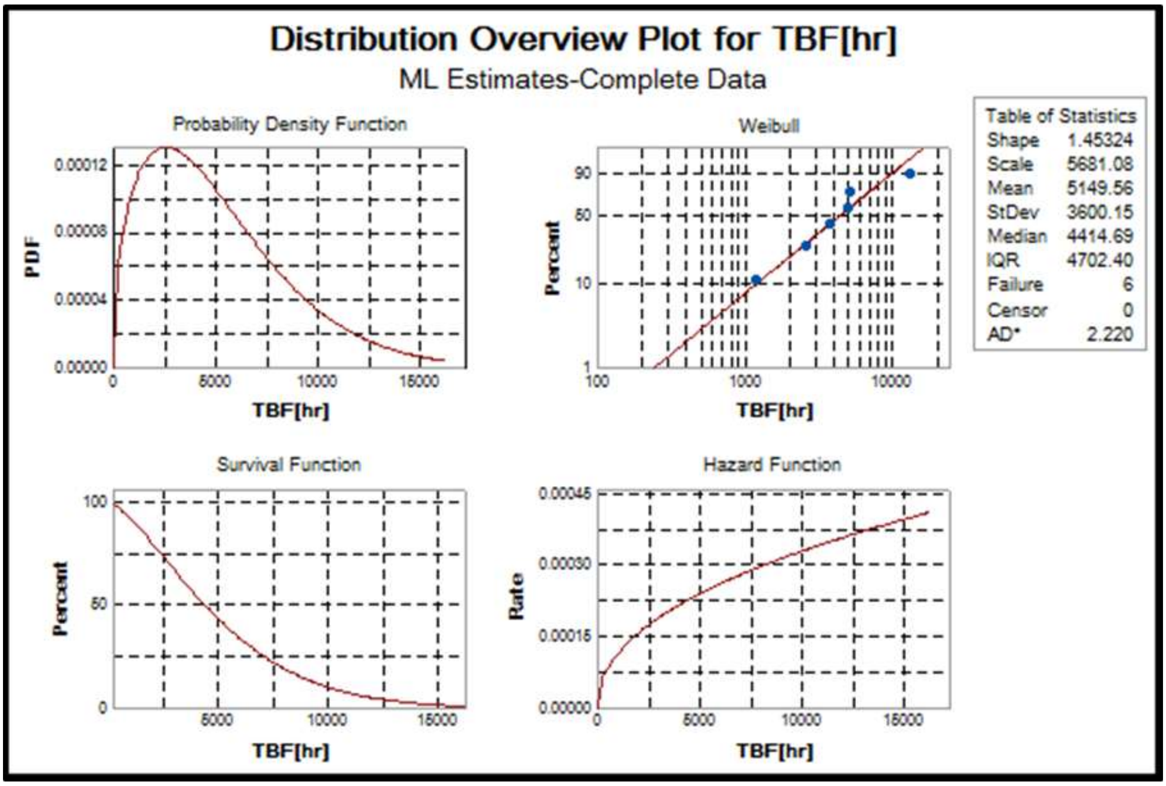


**Figure 5.** Weibull Distribution overview plot of the failures via Minitab of compressor.

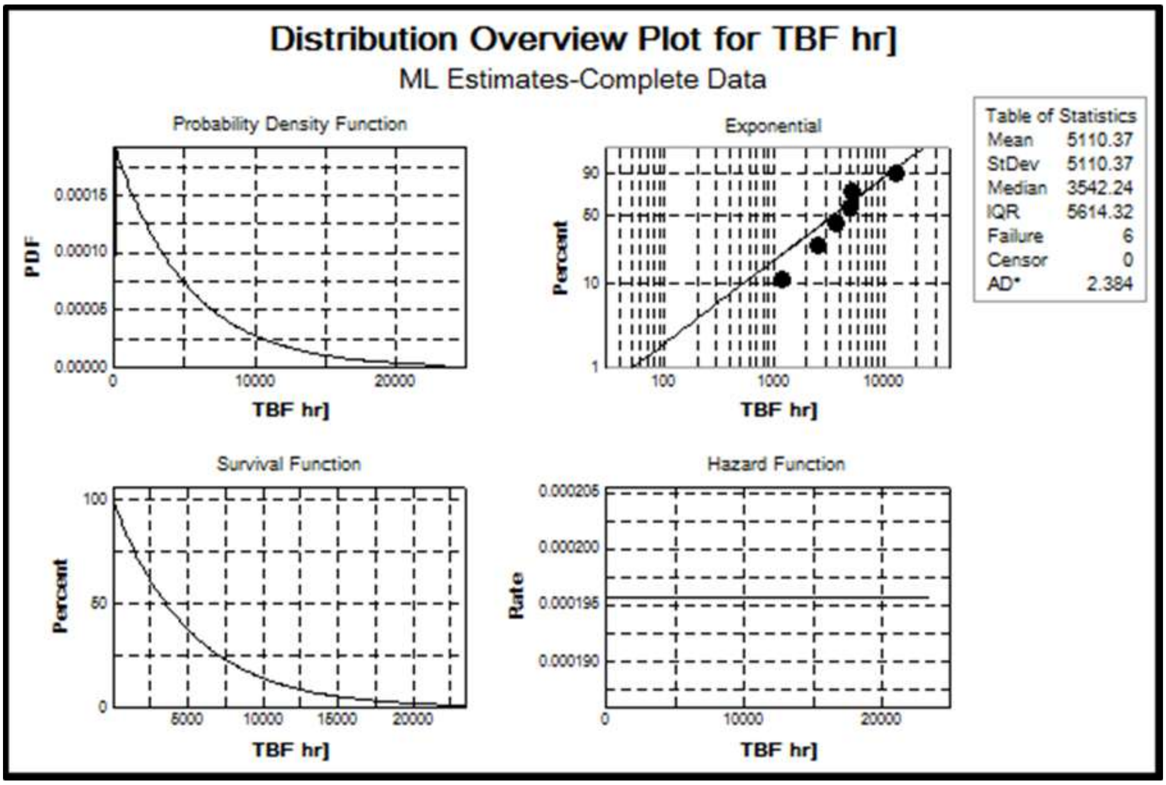


**Figure 6.** Exponential Distribution overview plot of the failures via Minitab of compressor.

.