

Repairable Systems (Times)

Summary

The **Repairable Systems (Times)** procedure is designed to analyze data consisting of failure times from systems that can be repaired. It is assumed that when the system fails, it is immediately repaired and placed in service again. Further, it is assumed that the repair time is negligible compared to the time between failures. The goal of the analysis is to develop a model that can be used to estimate failure rates or quantities such as the MTBF (mean time between failures).

This procedure differs from the *Distribution Fitting* and *Weibull Analysis* procedures in that it allows for a failure rate that changes as the system ages.

Sample StatFolio: *repairs2.sgp*

Sample Data:

The file *repairs2.sgd* contains the repair history of three simulated systems, similar to the data reported by Tobias and Trindade (1995) on page 360 of their textbook. The times (in hours) at which the systems failed are shown in the following table:

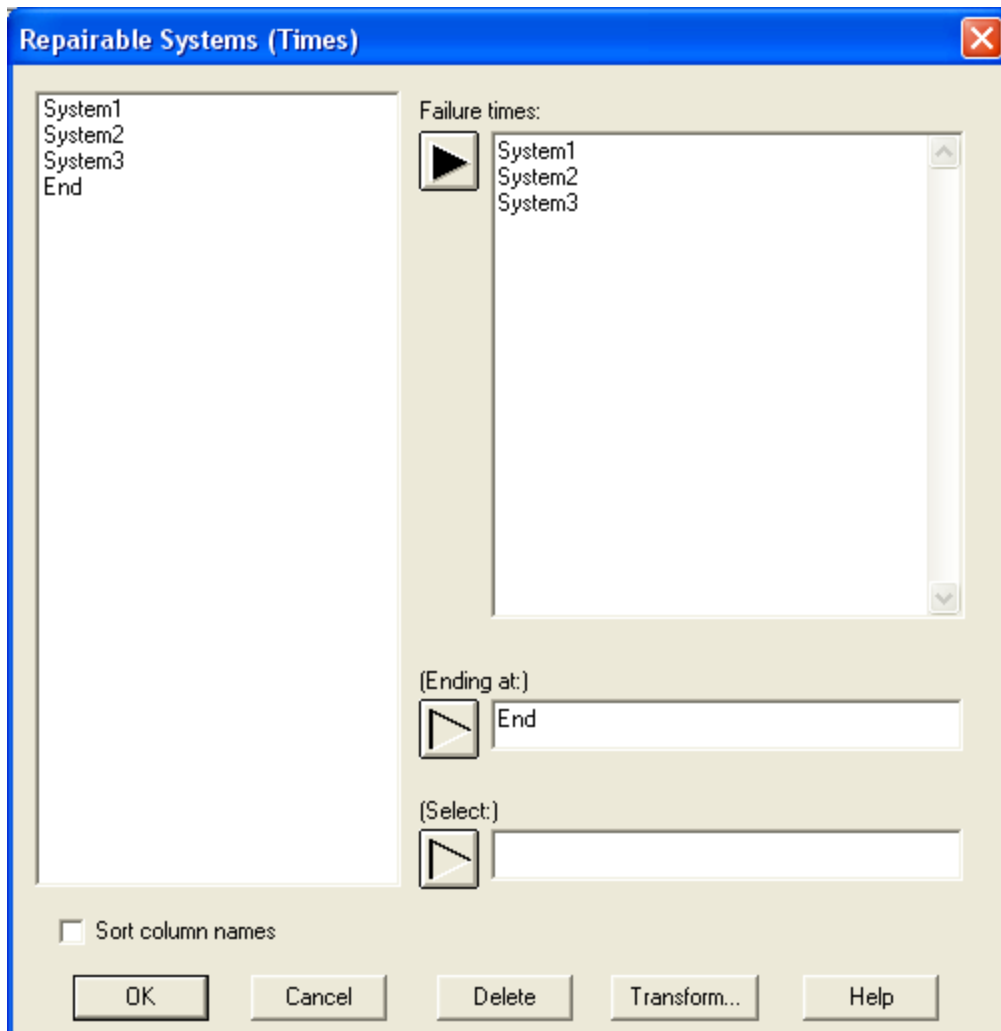
<i>System 1</i>	<i>System 2</i>	<i>System 3</i>
0.6	29.1	12.7
8.5	221.1	60.5
33	366.6	93.3
45.8	392.4	124.5
158.5	732.9	193.7
267.4		240.5
286.8		419.6
497.7		555.6
539		616.5

Observation started when the systems were first placed into service and was truncated as follows:

- System 1: Observation truncated after 1,000 hours.
- System 2: Observation truncated after 5 failures (at time 732.9).
- System 3: Observation truncated after 1,000 hours.

Data Input

The data input dialog box requests information about the failure times and the sampling intervals:



- **Failure times:** one or more numeric variables containing the failure times.
- **(Ending at):** an optional column indicating the times when observation ended for each variable.
- **Select:** subset selection.

Two kinds of sampling may be indicated:

- (1) *Failure truncated sampling* - If the *End at* field is blank or if the *End at* value is equal to the last failure time, the data will be treated as *failure truncated*. This normally occurs when a system is observed until a predetermined number of failures occur.
- (2) *Time truncated sampling* - If the *End at* value is not equal to the last failure time, the data will be treated as *time truncated*. This normally occurs when a system is observed for a predetermined amount of time.

Analysis Summary

The *Analysis Summary* displays a table showing the number of failures observed in each sample.

Repairable Systems (Times)

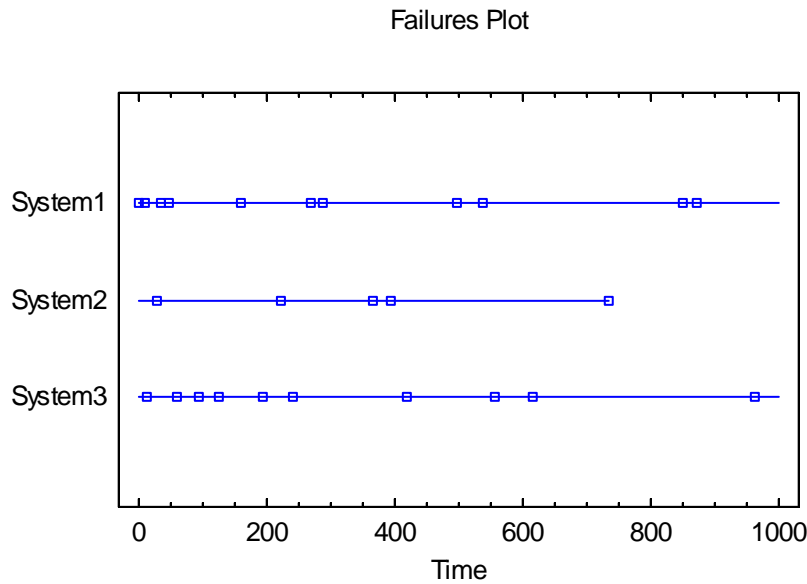
<i>Sample</i>	<i>Start</i>	<i>End</i>	<i>Events</i>
System1	0.0	1000.0	11
System2	0.0	732.9	5
System3	0.0	1000.0	10

The StatAdvisor

This procedure constructs models for the occurrence of failures in repairable systems. For the 3 samples entered, there were a total of 26 observed failures.

Failure Plot

The events plot plots the times of each failure.



Failure Rate Statistics

This table shows the observed failure rate for each sample and for all of the samples combined, with confidence limits:

	<i>Count</i>	<i>Sample size</i>	<i>Observed rate</i>	<i>95.0% lower confidence limit</i>	<i>95.0% upper confidence limit</i>
System1	11	1000.0	0.011	0.00549116	0.019682
System2	5	732.9	0.00682221	0.00221516	0.0159208
System3	10	1000.0	0.01	0.00479539	0.0183904
COMBINED	26	2732.9	0.0095137	0.00621467	0.0148076

If

$$n_q = \text{number of observed failures for sample } q \tag{1}$$

and

$$T_q = \text{sample size (duration of observation) for sample } q, \tag{2}$$

then the observed rate equals

$$\hat{\lambda}_q = \frac{n_q}{T_q} \tag{3}$$

The 100(1- α)% confidence limits for the failure rate of system q , assuming that the failure count comes from a Poisson distribution, are given by:

$$\left[\frac{\chi^2_{1-\alpha/2, 2n_q}}{2T_q}, \frac{\chi^2_{\alpha/2, 2(n_q+1)}}{2T_q} \right] \tag{4}$$

if the data are time truncated and

$$\left[\frac{\chi^2_{1-\alpha/2, 2n_q}}{2T_q}, \frac{\chi^2_{\alpha/2, 2n_q}}{2T_q} \right] \tag{5}$$

if the data are failure truncated. The confidence limits show the margin of error in estimating the true failure rate in the underlying process that generated the failures.

If more than one sample has been entered, a comparison of the samples is also included:

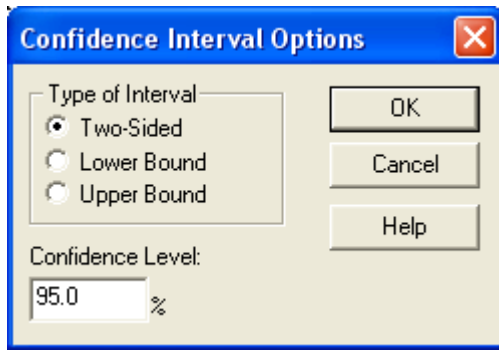
<p>Group Comparisons Likelihood ratio statistic = 0.86511 P-value = 0.648849</p>
--

A likelihood ratio test is constructed to test the null hypothesis that the underlying failure rate that generated each observed failure count is the same. For k samples, the test is constructed by calculating

$$H = 2 \left\{ \sum_{q=1}^k n_q \log(n_q / T_q) - \left(\sum_{q=1}^k n_q \right) \log \left(\sum_{q=1}^k n_q / \sum_{q=1}^k T_q \right) \right\} \tag{6}$$

and comparing it to a chi-squared distribution with $k - 1$ degrees of freedom. If the corresponding P-value is less than a selected value of α (such as 0.05), then there is a statistically significant difference amongst the observed failure rates.

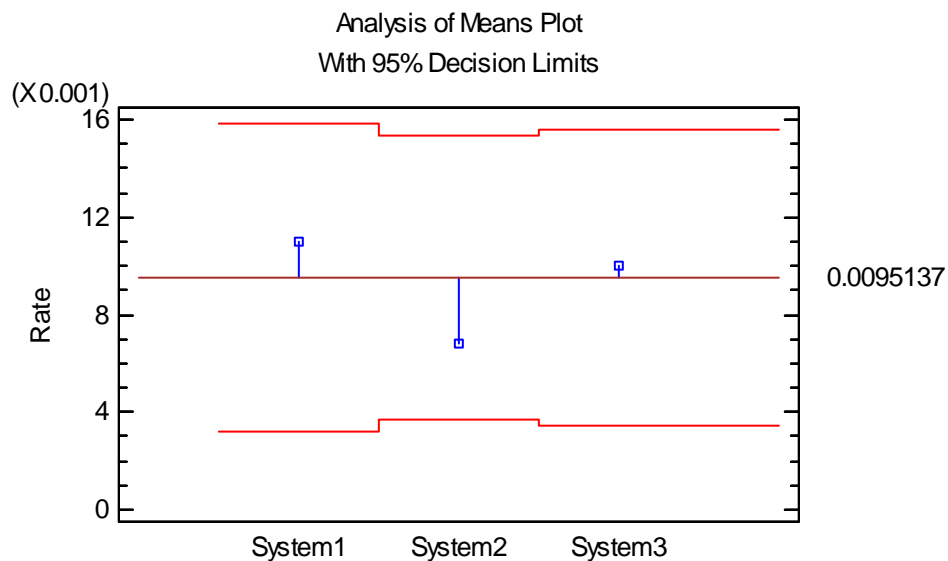
For the sample data, there is no reason to question the hypothesis that the samples come from processes with the same failure rate.

Pane Options

- *Type of Interval* – select either a two-sided confidence interval or a one-sided confidence bound.
- *Confidence level* – level of confidence for the interval or bound, usually either 90%, 95%, or 99%.

ANOM Plot

This plot compares the observed failure rates when more than one sample is available:



Any samples which fall beyond the decision limits are significantly different than the average failure rate at the specified confidence level, assuming that failures occur according to a homogeneous Poisson process. In this case, there are no significant differences amongst the failure rates in the 3 systems.

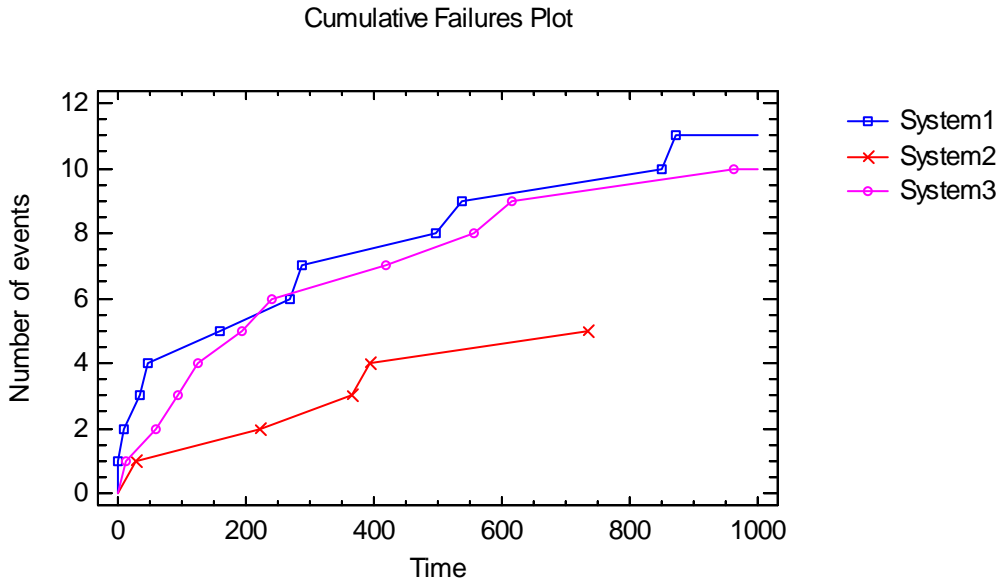
Cumulative Failures Plot

An important function in characterizing the performance of repairable systems is $M(t)$, the expected number of failures per system that will have occurred by time t . If $\lambda(t)$ equals the rate of failures as a function at time t , then

$$M(t) = \int_0^t \lambda(u) du \tag{7}$$

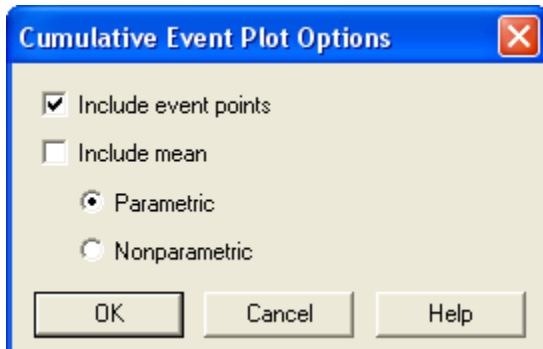
If the failure rate is a constant value λ , then $M(t) = \lambda t$, a straight line. If the failure rate increases or decreases over time, then $M(t)$ will curve.

The *Cumulative Failures Plot* displays $M(t)$ in various ways. First, it can plot the observed cumulative events curve for each sample as shown below:



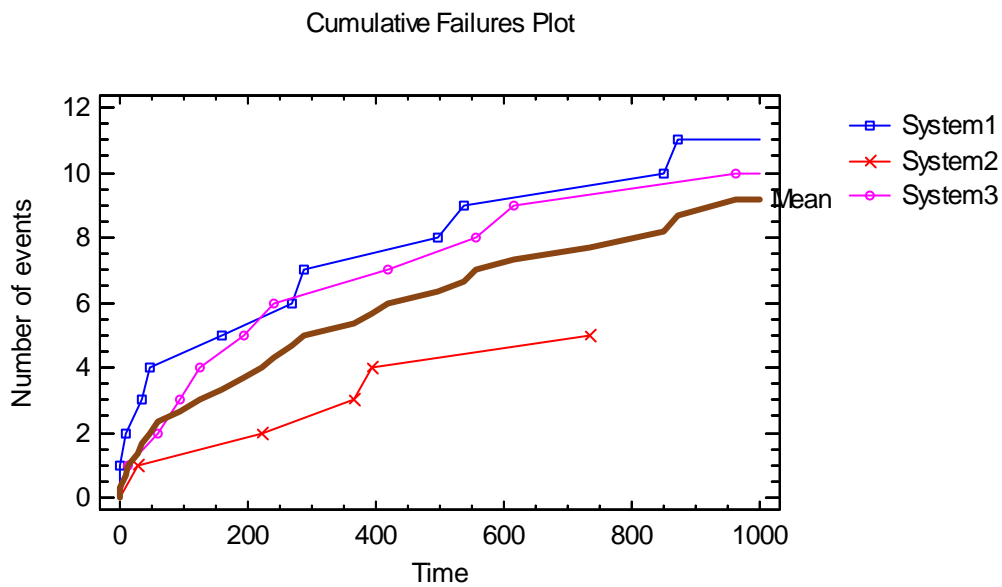
In the plot above, the number of failures that have occurred at each failure time is plotted and connected by straight lines.

Pane Options



- *Include event points* – check if point symbols should be plotted at each failure time.
- *Include mean* – check if a line should be placed at the estimated mean number of failures. The estimated mean may be determined in either of two ways:
 1. Using a *parametric* approach that depends on the currently selected model (see *Analysis Options* below).
 2. Using a *nonparametric* approach that essentially averages the observed curves (this is only available if more than one sample of failure times has been entered).

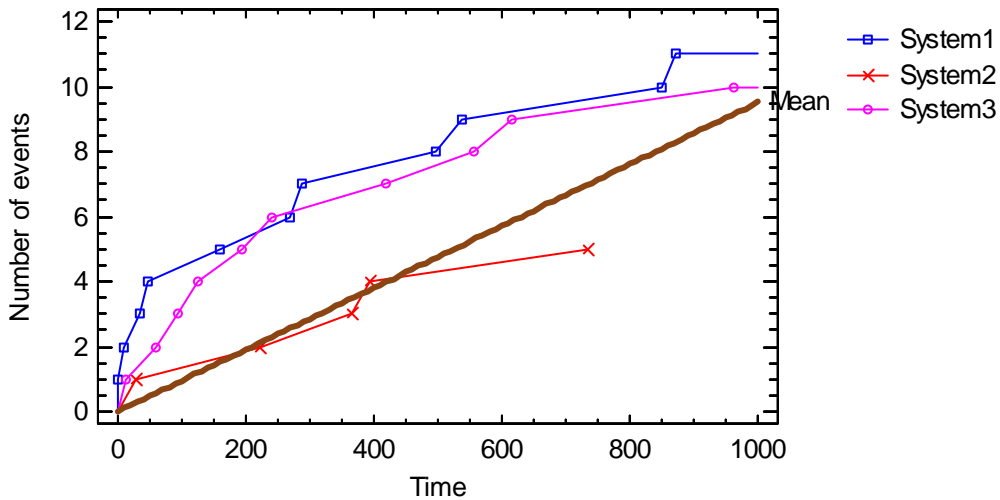
For example, the plot below adds a nonparametric estimate of the mean:



The nonparametric mean curve is determined using the “mean summation operator” discussed in Chapter 10 of Tobias and Trindade (1995).

The parametric approach uses the estimated failure rate model. For a homogenous Poisson process, the mean curve will be linear. For the sample data, the fitted curve equals $M(t) = 0.0083842t$ as shown below:

Cumulative Failures Plot



The fact that the observed curves do not match the fit is indicative of a time-dependent failure rate.

Interevent Time Statistics

This pane displays statistics for the times between consecutive failures, called “interevent times”. The top of the table shows statistics calculated for each sample:

Interevent Time Statistics				
Sample size: 28				
Number of censored values: 2				
Sample Statistics for Uncensored Values				
	Count	Mean	Standard dev.	C.V.
System1	11	79.2636	99.6159	1.25677
System2	5	146.58	130.44	0.88989
System3	10	96.16	101.315	1.05361
COMBINED	26	98.7077	104.949	1.06323

The table includes:

- *Sample size*: the total number of observed interevent times in all samples combined. This includes the times from the last event to the end of the observation period, if the observation period extended beyond the last failure.
- *Number of censored values*: the number of additional interevent times beyond the last failure, as described above. These values are said to be “censored” since the actual time from the last observed failure to the next failure is greater than the observed time from the last failure to the end of the sampling period.

- *Sample Statistics for Uncensored Values*: statistics calculated from the uncensored observations only. Included are the sample mean \bar{x} , sample standard deviation s , and coefficient of variation s/\bar{x} .

If failures occur according to a homogeneous Poisson process (a Poisson process with constant rate), then the interevent times follow an exponential distribution with coefficient of variation equal to 1.

A table is also displayed showing the estimated MTBF (mean time between failures) and associated confidence limits:

Confidence Limits for Mean				
	<i>Count</i>	<i>Estimate</i>	<i>95.0% lower confidence limit</i>	<i>95.0% upper confidence limit</i>
System1	12	90.9091	50.8077	182.111
System2	5	146.58	62.811	451.434
System3	11	100.0	54.3763	208.534
COMBINED	28	105.112	67.5327	160.91

These statistics are based on all of the data, including any censored observations. Letting

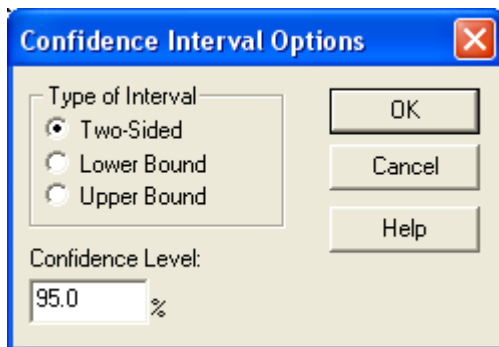
$$\mu_q = \text{MTBF for sample } q \tag{8}$$

the estimated mean time between failures is given by

$$\hat{\mu}_q = \frac{1}{\hat{\lambda}_q} \tag{9}$$

The confidence limits for the MTBF, assuming that events occur according to a homogeneous Poisson process, are given by the reciprocals of the confidence limits for the rates given by equation (4).

Pane Options



- *Type of Interval* – select either a two-sided confidence interval or a one-sided confidence bound.
- *Confidence level* – level of confidence for the interval or bound, usually 90%, 95%, or 99%.

Trend Test

An important assumption of the estimates calculated above is that events occur throughout the sampling period at a constant rate. This assumption can be tested using either of three tests:

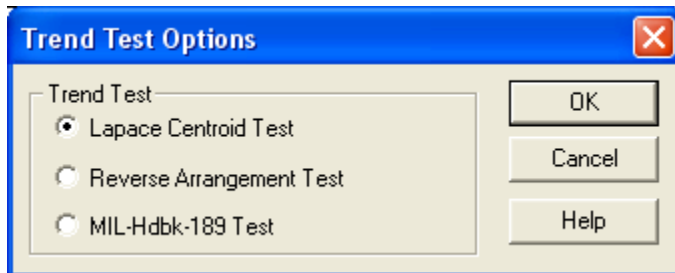
1. *Laplace Centroid test* – tests whether the failure times behave like the order statistics of a sample that is uniformly distributed over the sampling interval, as they should if they come from a homogenous Poisson process.
2. *Reverse arrangement test* – tests whether the magnitude of the interevent times is random versus the alternative that the times are increasing (or decreasing) over the sampling period. The null hypothesis is that of a stationary renewal process in which interevent times are randomly drawn from a single distribution (although not necessarily an exponential distribution).
3. *Mil-Hdbk-189 test* – estimates a growth parameter β and then tests the hypothesis that $\beta = 0$. The null hypothesis is that of a homogenous Poisson process.

For each test, a table similar to that shown below is displayed:

Laplace Centroid Test		
	<i>Test statistic</i>	<i>P-Value</i>
System1	-2.02856	0.0425027
System2	-1.07908	0.280553
System3	-1.88581	0.0593204
COMBINED	-2.96659	0.00301136

Small P-values (less than 0.05 if operating at the 5% significance level) lead to the conclusion that a trend is present. This is a clear suggestion that trend is present in the sample data.

Pane Options



- *Trend Test* – type of test to apply.

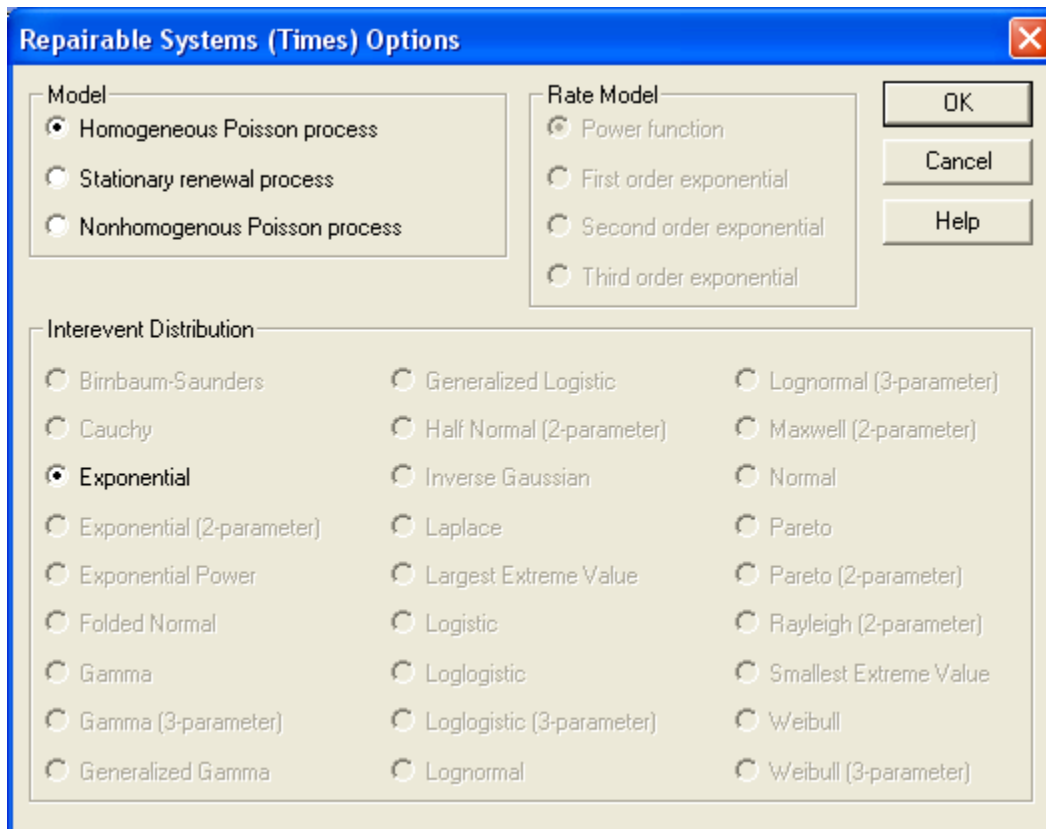
Point Process Model

The times at which failures occur form a stochastic *point process*. To model a point process, it is necessary to estimate:

1. the rate at which failures occur.
2. the distribution of the times between failures.

In addition, it is important to determine whether or not consecutive interevent times are independent.

In STATGRAPHICS, the *Analysis Options* dialog box is used to select the type of point process model to use:



- **Model** – the type of model to estimate. The following model types are available:
 - *Homogenous Poisson process* – a process in which events occur at a constant rate and the interevent times are independently distributed occurring to an exponential distribution.

- *Stationary renewal process* – a process in which events occur at a constant rate and the interevent times are independently distributed occurring to a probability distribution that is not necessarily exponential.
- *Nonhomogenous Poisson process* – a generalization of the homogenous Poisson process in which the rate changes over time.
- **Interevent Distribution** – for a stationary renewal process, the distribution of the time between events.
- **Rate Model** – for a nonhomogenous Poisson process, the type of function that characterizes changes in the failure rate over time. The following functions are available:

Model	Rate function
Power function	$\lambda(t) = at^b$
First order exponential	$\lambda(t) = e^{a+bt}$
Second order exponential	$\lambda(t) = e^{a+bt+ct^2}$
Third order exponential	$\lambda(t) = e^{a+bt+ct^2+dt^3}$
IBM Model	$\lambda(t) = a + bce^{-ct}$

Example #1 – Homogeneous Poisson Process

If the data follow a homogenous Poisson process, the rate is assumed to be constant and the interevent times follow an exponential distribution. Fitting this model to the sample data gives the following results:

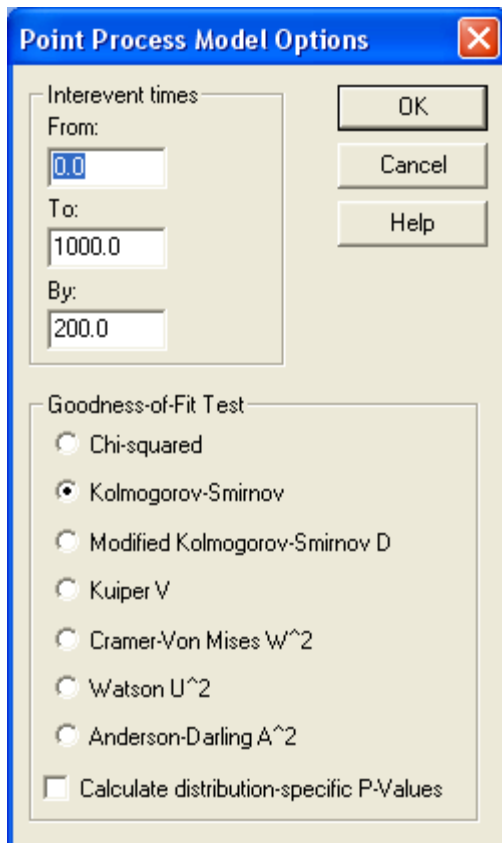
Point Process Model			
Model: homogeneous Poisson process			
Rate model: 0.0095137			
Mean cumulative events model: 0.0095137*t			
Interevent distribution: Exponential			
mean = 105.112			
<i>t</i>	<i>Rate</i>	<i>Mean cum events</i>	<i>Mean interevent time</i>
0.0	0.0095137	0.0	105.112
200.0	0.0095137	1.90274	105.112
400.0	0.0095137	3.80548	105.112
600.0	0.0095137	5.70822	105.112
800.0	0.0095137	7.61096	105.112
1000.0	0.0095137	9.5137	105.112
Goodness-of-Fit Test			
<i>Kolmogorov-Smirnov D</i>	<i>P-value</i>		
0.1413	0.631059		

The output includes:

- *Rate model*: the estimated rate is $\lambda(t) = 0.0095137$ for all *t*.

- *Mean cumulative events model*: the mean cumulative events function increases linearly according to $M(t) = 0.0095137t$.
- *Interevent distribution*: the distribution of time between failures is exponential with an estimated MTBF of 105.112.
- *Tabled values*: the failure rate, mean cumulative events function, and mean interevent time are given at selected values of t . For this model, the values are the same for all t .
- *Goodness-of-fit test*: the results of a goodness-of-fit test used to determine whether the interevent times appear to follow an exponential distribution. A small P-value would indicate that the exponential distribution does not fit the observed interevent times.

Pane Options



- *Interevent Times From, To and By* – settings used to define the values of t at which estimates are displayed.
- *Goodness-of-Fit Test* – used to test the adequacy of the selected model for describing the interevent time distribution. The available tests depend on whether the data contain any censored values. If using the chi-squared test, the data will be grouped into classes with equal expected frequencies. If *Calculate distribution-specific P-values* is checked, the P-Values will be based on tables or formulas specifically developed for the distribution being tested. Otherwise, the P-Values will be based on a general table or formula that

applies to all distributions. For more details, refer to the document *Distribution Fitting (Censored Data)*.

For the sample data, the goodness-of-fit test suggests that the exponential distribution is adequate.

Example #2 – Stationary Renewal Process

A stationary renewal process has a constant failure rate, but the interevent distribution is not necessarily exponential. For example, the following table shows the results of fitting a gamma distribution to the interevent times:

Point Process Model			
Model: stationary renewal process			
Rate model: 0.00946612			
Mean cumulative events model: 0.00946612*t			
Interevent distribution: Gamma			
shape = 0.886563			
scale = 0.00839231			
(Mean = 105.64)			
<i>t</i>	<i>Rate</i>	<i>Mean cum events</i>	<i>Mean interevent time</i>
0.0	0.00946612	0.0	105.64
200.0	0.00946612	1.89322	105.64
400.0	0.00946612	3.78645	105.64
600.0	0.00946612	5.67967	105.64
800.0	0.00946612	7.57289	105.64
1000.0	0.00946612	9.46612	105.64
Goodness-of-Fit Test			
<i>Kolmogorov-Smirnov D</i>	<i>P-value</i>		
0.119066	0.822196		

The gamma distribution is defined by shape and scale parameters and contains the exponential distribution as a special case when the shape parameter equals 1. Notice that the rate model, mean cumulative events curve, and mean interevent time are slightly different than if an exponential distribution is used.

NOTE: the distribution parameters are determined using maximum likelihood estimation including any censored data as described in the document *Distribution Fitting (Censored Data)*.

Example #3 – Nonhomogenous Poisson Process

This model allows for a time-dependent failure rate. Fitting a power function for the rate yields the following results:

Point Process Model			
Model: nonhomogeneous Poisson process			
Rate model: $0.137601 * t^{-0.490687}$			
Mean cumulative events model: $0.27017 * t^{0.509313}$			
Interevent distribution: Exponential			
Mean = $7.26737 * t^{0.490687}$			
<i>t</i>	<i>Rate</i>	<i>Mean cum events</i>	<i>Mean interevent time</i>
0.0	0.137601	0.0	7.26737
200.0	0.010222	4.01405	97.8279
400.0	0.00727488	5.71348	137.459
600.0	0.00596239	7.02403	167.718
800.0	0.00517743	8.13241	193.146
1000.0	0.00464047	9.11123	215.496
Goodness-of-Fit Test			
<i>Kolmogorov-Smirnov D</i>	<i>P-value</i>		
0.139279	0.649084		

Note the following:

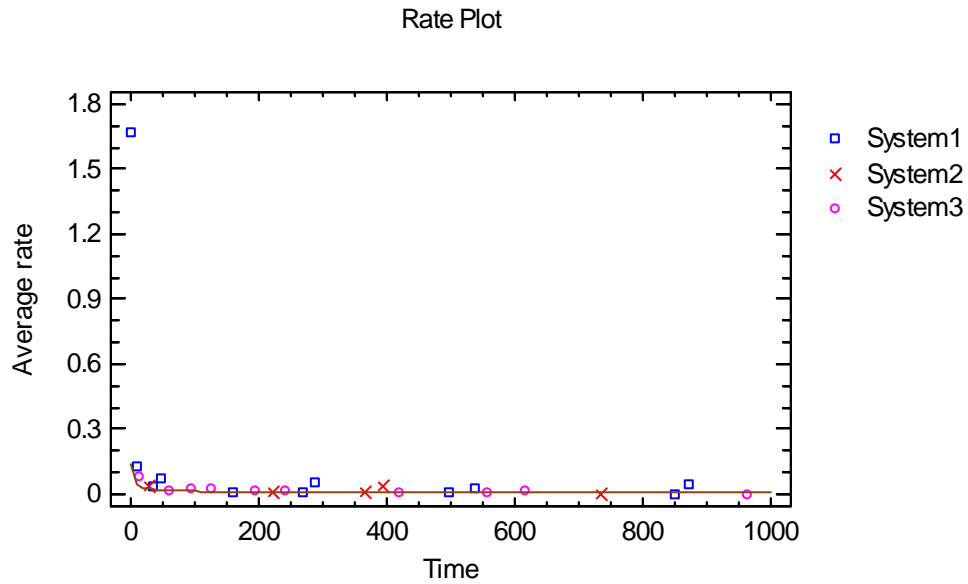
- *Rate model*: the estimated rate is $\lambda(t) = 0.137601 t^{-0.490687}$, which decreases over time.
- *Mean cumulative events model*: the mean cumulative events function is given by $M(t) = 0.270t^{0.509}$, which has a decreasing slope over time.
- *Mean interevent time*: the MTBF increases as the product ages.
- *Goodness-of-fit test*: this tests whether the normalized interevent times follow a standard exponential distribution. The normalized interevent times are calculated according to

$$e_i = \int_{t_{i-1}}^{t_i} \hat{\lambda}(u) du = \hat{M}(t_i) - \hat{M}(t_{i-1}) \tag{10}$$

For this data, the goodness-of-fit test for the power function indicates that the normalized interevent times are adequately represented by a standard exponential distribution.

Failure Rate Plot

This plot shows the estimated failure rate, together with the reciprocals of the times between consecutive failures:

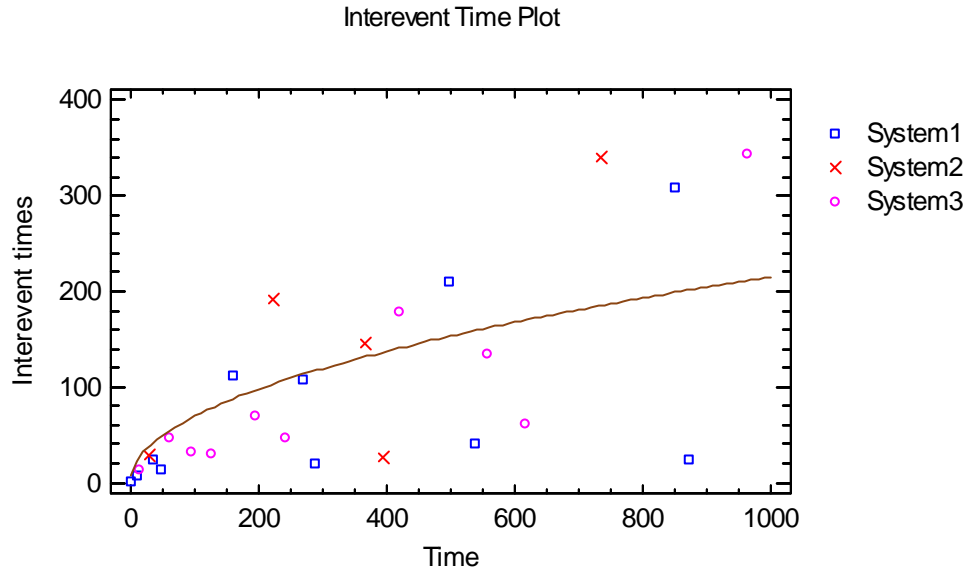


Since the mean time between consecutive failures is inversely proportional to the failure rate, the reciprocals of the interevent times are indicative of the average failure rate at that point in time.

Notice that the power function model shown above indicates a high initial failure rate which decreases rapidly and then levels off.

Interevent Time Scatterplot

This plot shows the interevent times (times between consecutive failures) together with the estimated mean time between failures:



The fitted power function shown above indicates that the mean time between failures increases as the system ages.

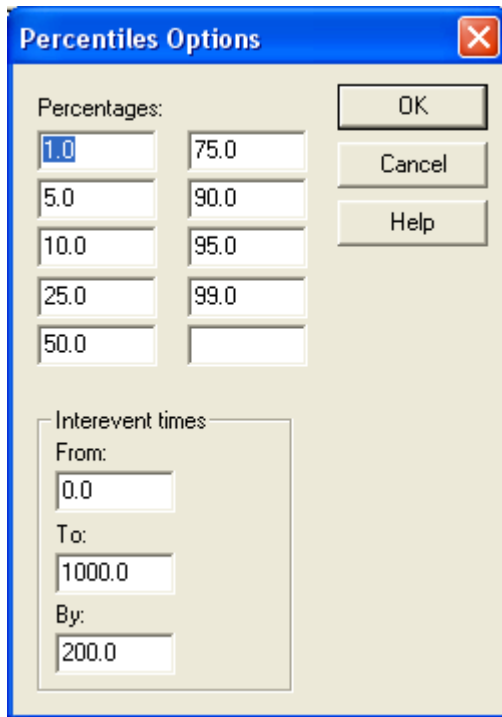
Model Percentiles

This table shows estimated percentiles for the distribution of time between consecutive failures:

Model Percentiles									
Model: nonhomogeneous Poisson process									
Rate model: $0.137601 * t^{-0.490687}$									
<i>t</i>	1.0%	5.0%	10.0%	25.0%	50.0%	75.0%	90.0%	95.0%	99.0%
0.0	0.0730396	0.372768	0.765694	2.09069	5.03736	10.0747	16.7337	21.7711	33.4675
200.0	0.983203	5.01792	10.3072	28.1433	67.8091	135.618	225.257	293.066	450.514
400.0	1.38151	7.05074	14.4828	39.5446	95.2796	190.559	316.512	411.791	633.024
600.0	1.68562	8.60281	17.6709	48.2495	116.253	232.507	386.185	502.439	772.37
800.0	1.94118	9.90709	20.35	55.5646	133.879	267.757	444.735	578.614	889.47
1000.0	2.1658	11.0535	22.7047	61.9942	149.37	298.74	496.197	645.567	992.393

A percentile is a value below which a given percentage of the interevent times are estimated to lie. For example, the table above shows that when the system is new, 25% of the times between failures are estimated to be 2.09 or less. The 25th percentile increases as the system ages, reaching as high as 61.99 after $t = 1000$ hours.

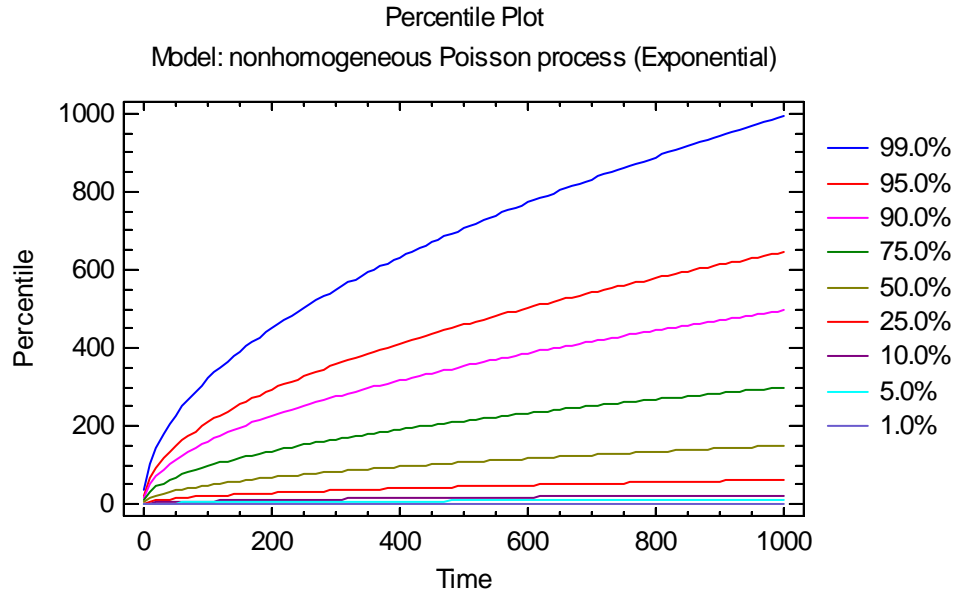
Pane Options



- **Percentages** – up to 10 percentages at which the percentiles will be calculated.
- **Interevent times** – range of values for t at which percentiles will be tabulated.

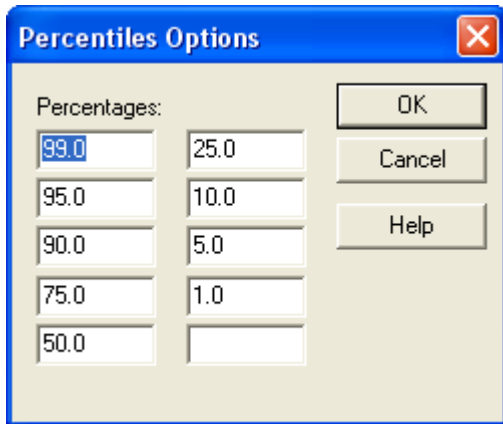
Interevent Time Percentile Plot

This plot displays estimated percentiles for the distribution of time between consecutive failures:



This plot corresponds to the values in the table referred to above.

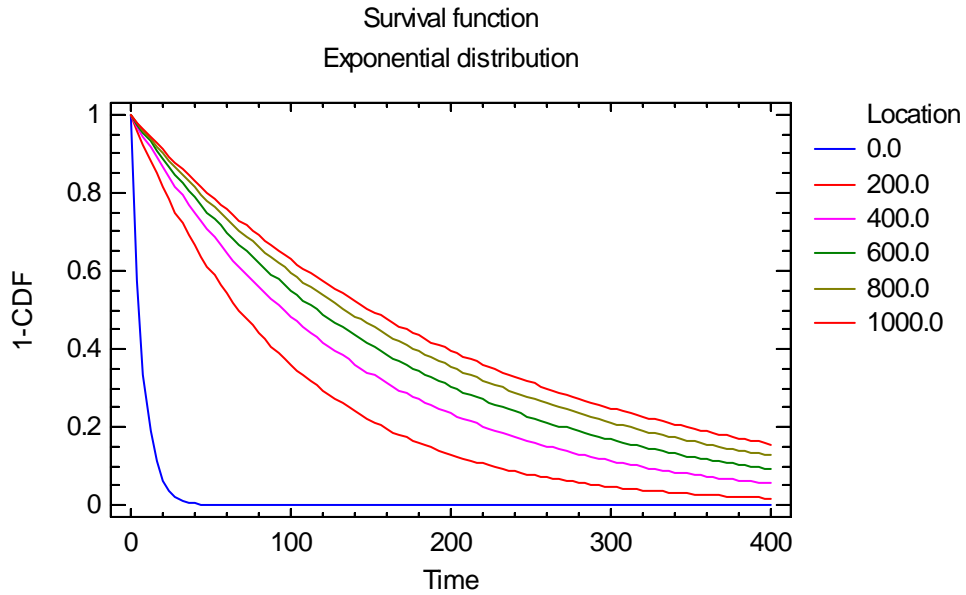
Pane Options



- **Percentages** – up to 10 percentages at which the percentiles will be plotted.

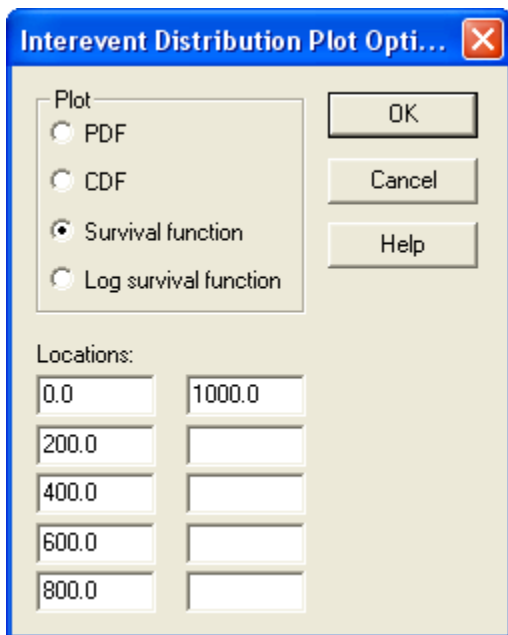
Interevent Time Distribution Plots

This plot displays the distribution of time between consecutive failures:



For a nonhomogeneous Poisson process, the distribution is exponential with a mean that depends upon the system age t .

Pane Options



- **Plot** – the type of function to be plotted. Choices include:

- **PDF** – probability density function. This function shows the relative probability of obtaining interevent times in the vicinity of X , where X is plotted along the horizontal axis.
 - **CDF** – cumulative distribution function. This function shows the probability that the time between failures is less than or equal to X .
 - **Survival function** - This function shows the probability that the time between failures is greater than X .
 - **Log survival function** – natural logarithms of the survival function.
- **Locations** – values of t at which to plot the distribution.

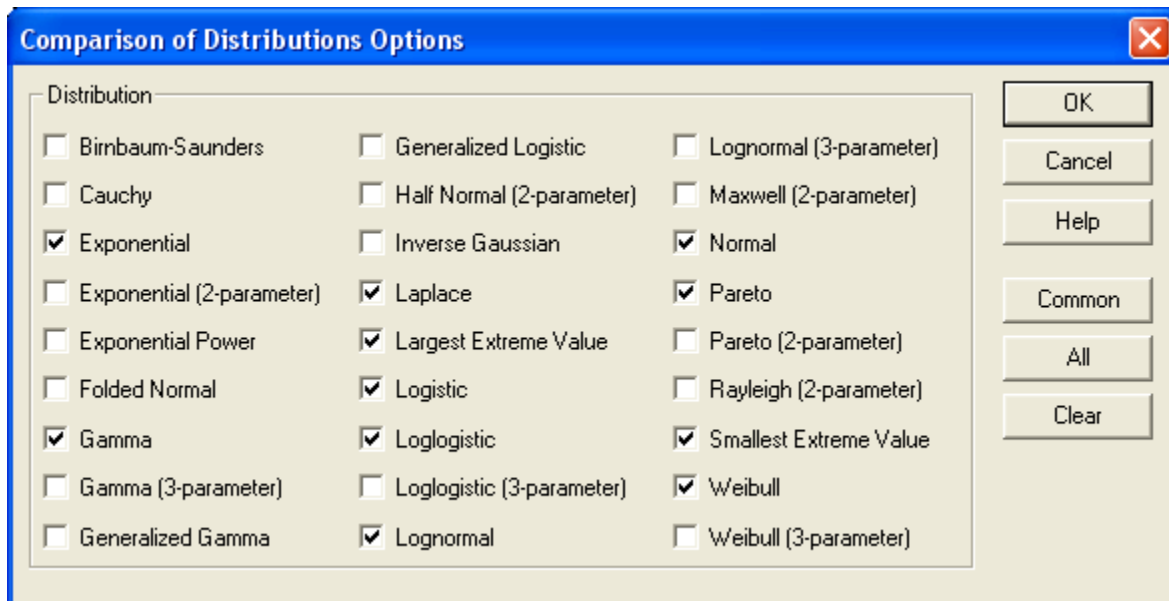
Comparison of Alternative Distributions

This table compares different distributions for the time between consecutive failures, if the selected point process model is a stationary renewal process:

Comparison of Alternative Distributions		
Distribution	Est. Parameters	Kolmogorov-Smirnov D
Lognormal	2	0.102256
Loglogistic	2	0.107455
Weibull	2	0.115722
Gamma	2	0.119066
Exponential	1	0.1413
Logistic	2	0.176582
Largest Extreme Value	2	0.19937
Normal	2	0.215728
Smallest Extreme Value	2	0.226478
Laplace	2	0.272333
Pareto	1	0.364945

The distributions are ranked from best to worst based on the goodness-of-fit statistic selected for the *Point Process Model* pane shown earlier. For the sample data, the lognormal distribution provides a somewhat better model than the exponential distribution, assuming that the failure rate is constant over time.

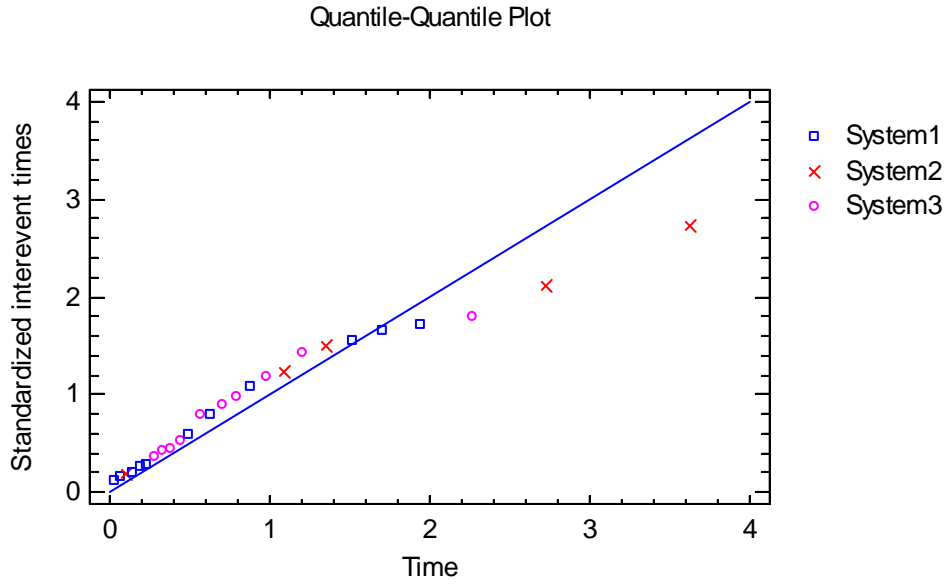
Pane Options



- **Distribution** – check all distributions to be fit.

Interevent Time Quantile-Quantile Plot

This plots displays the quantiles of the data and the fitted distribution:



If the distribution fits the data well, the points should lie close to the diagonal line.

NOTE: when fitting a nonhomogenous Poisson process, the interevent times are normalized using equation (10) before being plotted.

Save Results

The following results may be saved to the datasheet:

1. *Interevent Times* – the times between consecutive failures
2. *Censoring indicators* – a column containing a 0 for each uncensored time and a 1 for each censored time.
3. *Interevent locations* – the locations t corresponding to the end of each interevent times.
4. *Variable numbers* – an integer indicating the sample number of each interevent time.
5. *Normalized interevent times* – the normalized times as defined by equation (10).

Calculations

Laplace Test

Let t_{qj} = failure time of j th failure for system q . For a single sample, the test statistic is calculated by:

$$L_q = \frac{\sum_{j=1}^{n_q} t_{qj} - \frac{n_q T_q}{2}}{T_q \sqrt{\frac{n_q}{12}}} \tag{11}$$

which is compared to a standard normal distribution. For k samples, a pooled statistic is calculated from

$$L = \frac{\sum_{j=1}^{n_1} t_{1j} + \sum_{j=1}^{n_2} t_{2j} + \dots + \sum_{j=1}^{n_k} t_{kj} - \frac{1}{2}(n_1 T_1 + n_2 T_2 + \dots + n_k T_k)}{\sqrt{\frac{1}{12}(n_1 T_1^2 + n_2 T_2^2 + \dots + n_k T_k^2)}} \tag{12}$$

which is also compared to a standard normal distribution.

Reverse Arrangement Test

For sample q , let X_i be the i -th interevent time. Let R_q be the number of times that $X_i < X_j$ for $i=1, \dots, n_q-1$ and $j = i+1, \dots, n_q$. The test statistic

$$Z_q = \frac{R_q + \frac{1}{2} - \frac{n_q(n_q - 1)}{4}}{\sqrt{\frac{(2n_q + 5)(n_q - 1)n_q}{72}}} \tag{13}$$

is then compared to a standard normal distribution and a P-value P_q is calculated. To combine the test results, Fisher’s composite test is used which calculates

$$\chi^2 = \sum_{q=1}^k -2 \ln(P_q) \tag{14}$$

which is compared to a chi-square distribution with $2k$ degrees of freedom.

Mil-Hdbk-189 test

For a time truncated sample, the test statistic

$$\chi_q^2 = 2 \sum_{i=1}^{n_q} \ln \left(\frac{T_q}{t_{qi}} \right) \quad (15)$$

is compared to a chi-square distribution with $2n_q$ degrees of freedom. For a failure truncated system, the test statistic

$$\chi_q^2 = 2 \sum_{i=1}^{n_q-1} \ln \left(\frac{T_q}{t_{qi}} \right) \quad (16)$$

is compared to a chi-square distribution with $2(n_q-1)$ degrees of freedom.

For the combined samples, the separate test statistics are summed and compared to a single chi-square distribution.

Power Function Model

Parameter estimates are obtained using a modified MLE procedure.

$$\hat{b} = \frac{\sum_{q=1}^k N_q - 1}{\sum_{q=1}^k \sum_{i=1}^{N_q} \ln(T_q / t_{qi})} \quad (17)$$

$$\hat{a} = \frac{\sum_{q=1}^k n_q}{\sum_{q=1}^k T_q \hat{b}} \quad (18)$$

where $N_q = n_q$ if the data for system q are time truncated, while $N_q = n_q - 1$ if the data for system q are failure truncated.