Repairable Systems (Intervals)

Summary

The **Repairable Systems (Intervals)** procedure is designed to analyze data consisting of failure counts 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 *Life Tables* procedure in that it allows for a failure rate that changes as the system ages.

Sample StatFolio: *repairs1.sgp*

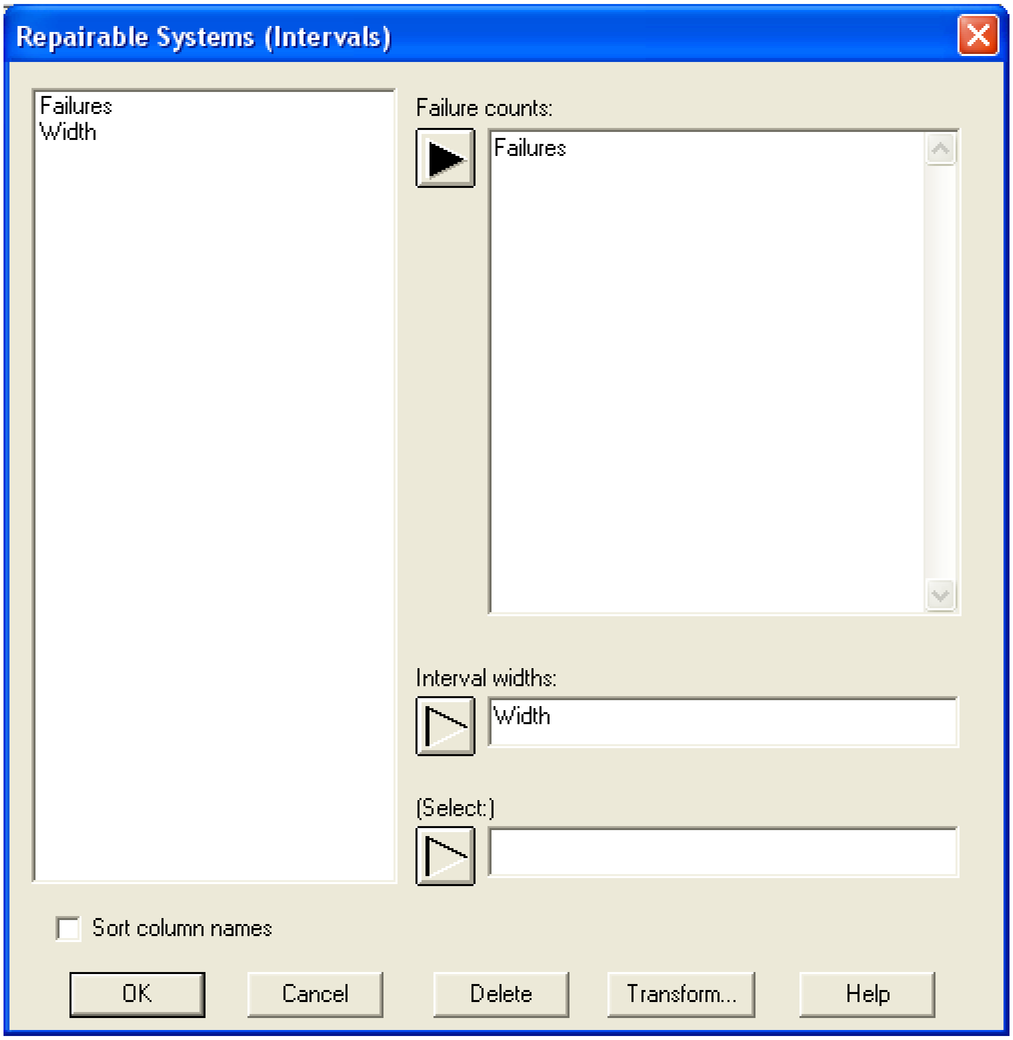
Sample Data:

The file *repairs1.sgd* contains the repair history of a computer system, reported by Cox and Lewis (1966) on page 11 of their much referenced text. The number of failures in consecutive intervals is shown in the following table:

Interval	*Failures*	*Width*
0-5000	23	5000
5000-10000	9	5000
10000-15000	12	5000
15000-20000	19	5000
20000-25000	23	5000
25000-30000	16	5000
30000-35000	4	5000
35000-40000	2	5000
40000-45000	8	5000
45000-50000	7	5000
50000-55000	8	5000
55000-60000	12	5000
60000-65000	9	5000
65000-70000	19	5000
70000-75000	17	5000
75000-80000	24	5000
80000-85000	13	5000
85000-90000	23	5000
90000-93500	9	3500

Data Input

The data input dialog box requests information about the failure counts and the width of the sampling intervals:



- **Failure counts**: one or more numeric variables containing the failure counts.

- **Interval width:** a column containing the width of each interval. If all interval widths are the same, a single number may be entered.

- **Select**: subset selection.

Analysis Summary

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

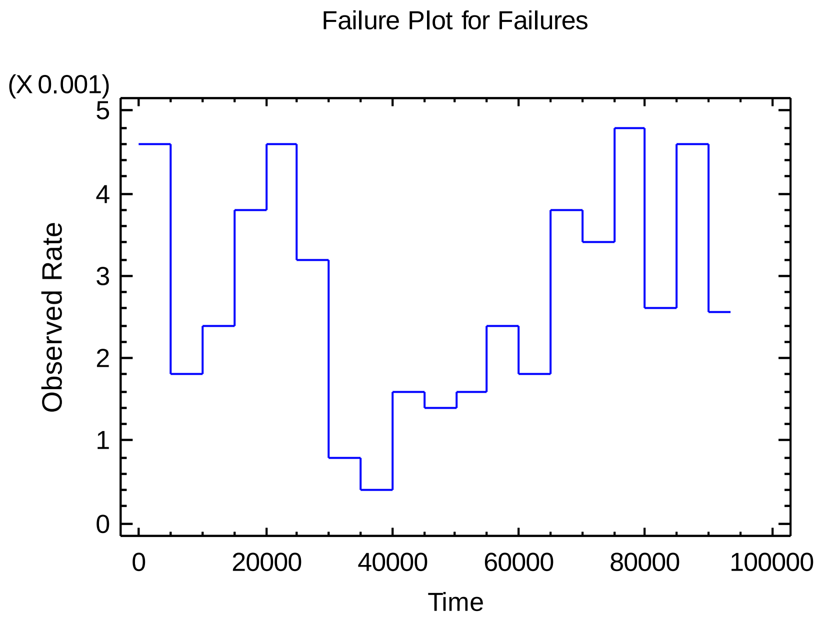
Repairable Systems (Intervals)

Sample	*Start*	*End*	*Events*
Failures	0.0	93500.0	257

The StatAdvisor
This procedure constructs models for the occurrence of failures in repairable systems. For Failures, there were 257 observed failures.

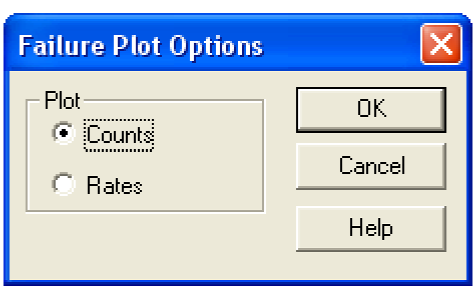
Failure Plot

The *Failure Plot* displays either the number of failures or failure rate in each interval:



If the width of the intervals varies, it may be preferable to plot the failure rates rather than the number of failures.

Pane Options



- *Plot*- whether to plot failure counts or failure rates.

Failure Rate Statistics

This table shows the estimated failure rate, with confidence limits:

Failure Rate Statistics

	Count	*Sample size*	*Observed rate*	*95.0% lower confidence limit*	*95.0% upper confidence limit*
Failures	257	93500.0	0.00274866	0.00242284	0.00310608

If

n_q = number of observed failures for sample q (1)

and

T_q = sample size (duration of observation) for sample q, (2)

then the observed rate equals

$$\hat{\lambda}_q = \frac{n_q}{T_q} \quad (3)$$

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

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

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, similar to that shown below for data that came from several samples:

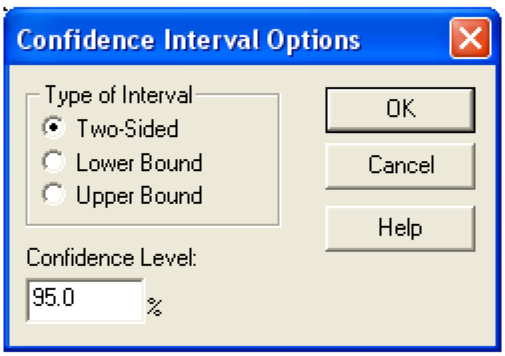
Group Comparisons
Likelihood ratio statistic = 0.105356
P-value = 0.948686

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. The test is constructed by calculating

$$H = 2\left\{ \sum_{i=1}^{q} n_i \log(n_i / T_i) - \left(\sum_{i=1}^{q} n_j \right) \log\left(\sum_{i=1}^{q} n_i / \sum_{i=1}^{q} T_i \right) \right\} \quad (5)$$

and comparing it to a chi-squared distribution with $q - 1$ degrees of freedom. If the corresponding P-value is greater than or equal to a selected value of α (such as 0.05), then there is not a statistically significant difference amongst the observed failure rates.

Pane Options

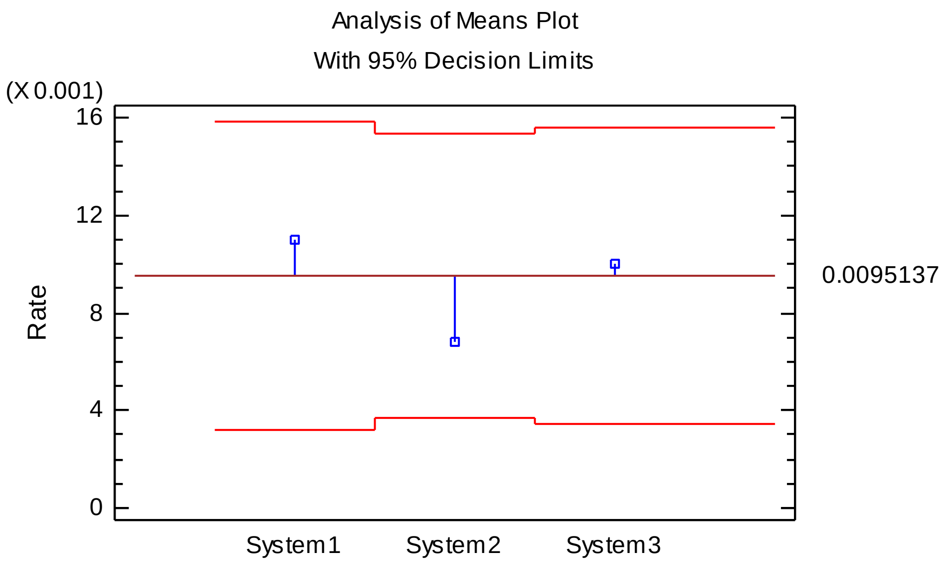


- *Type of Interval* – select either a two-sided confidence interval or a one-sided confidence bound.

- *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 the above plot, there are no significant differences amongst the failure rates in the 3 systems.

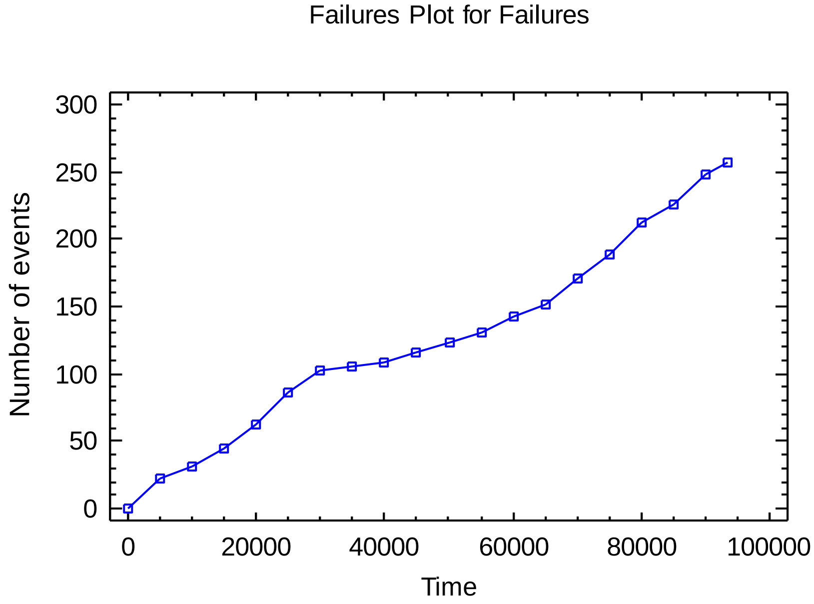
Cumulative Failures Plot

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

$$M(t) = \int_0^t \lambda(u) du \tag{6}$$

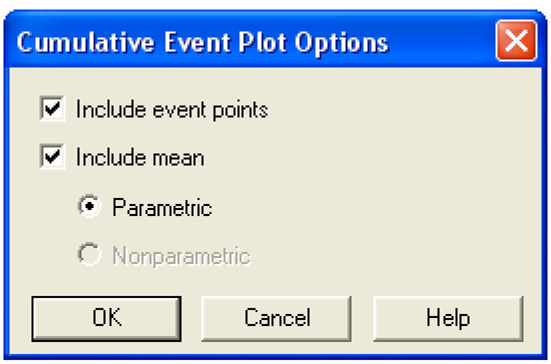
If the failure rate is a constant value λ, then $M(t) = \lambda t$, a straight line. If the failure rate increases or decreases over time, then *M(t)* will curve.

The *Cumulative Failures Plot* will plot *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 the end of each interval is plotted and connected by straight lines.

Pane Options



- *Include event points* – check if point symbols should be plotted at the end of each interval.

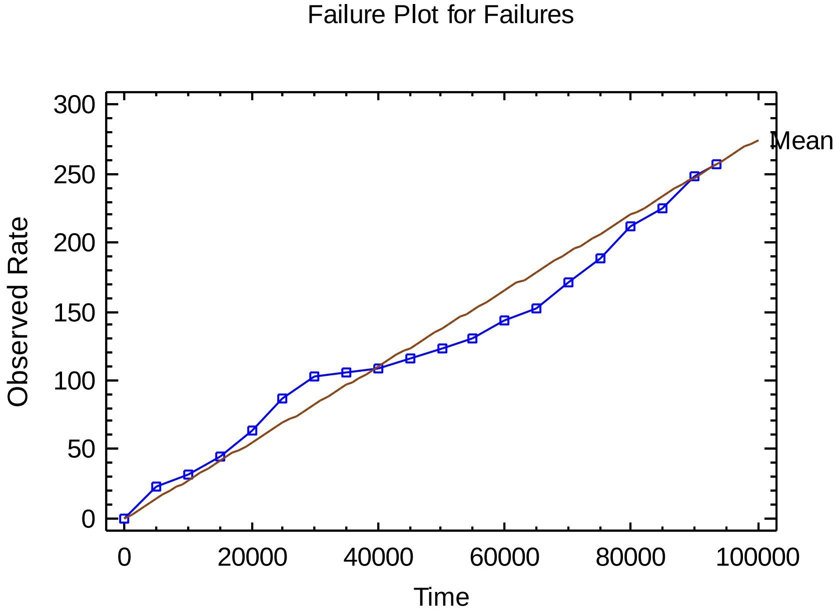
- *Include mean* – check if a line should be placed at the estimated mean number of failures. The estimated mean may be determined in either of two ways:

  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).

The parametric approach uses the estimated failure rate model. For a nonhomogenous Poisson process, the parametric curve will be linear. For the sample data, the fitted curve equals

$M(t)$ = 0.00274866*t



The fact that the observed curve does not match the fit well is indicative of a nonstationary failure rate.

Trend Test

An important assumption of the estimates calculated above is that events occur throughout the sampling period at a constant rate. One way to test this hypothesis is to compare the observed count in each interval to the expected count assuming that the rate is constant. This can be done using a chi-squared test. The results for the sample data are shown below:

Trend Test

Chi-Square Test

	Test statistic	*P-Value*
Failures	34.5063	0.0108991

A small P-value such as that observed leads to the conclusion that a trend is present.

Point Process Model

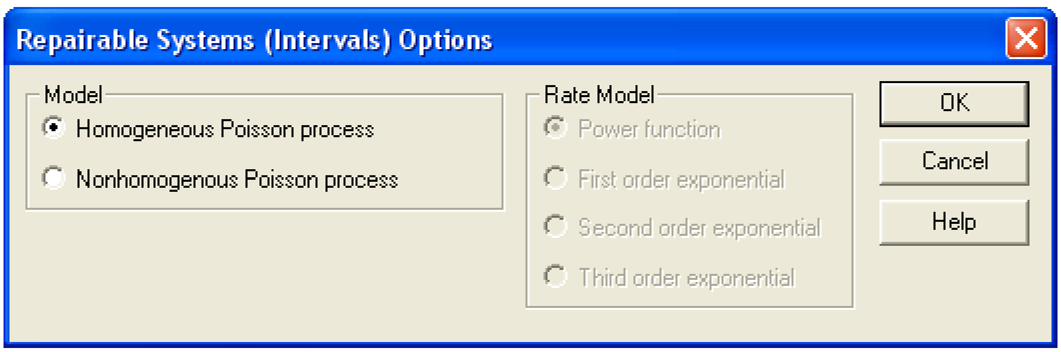
The times at which failures occur form a stochastic *point process*. To model a point process, it is necessary to estimate:

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. For interval data, 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.

    - *Nonhomogenous Poisson process* – a generalization of the homogenous Poisson process in which the rate changes over time.

- **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: 0.00274866

Mean cumulative events model: 0.00274866*t

Interevent distribution: Exponential
mean = 363.813

t	*Rate*	*Mean cum events*	*Mean interevent time*
0.0	0.00274866	0.0	363.813
20000.0	0.00274866	54.9733	363.813
40000.0	0.00274866	109.947	363.813
60000.0	0.00274866	164.92	363.813
80000.0	0.00274866	219.893	363.813
100000.	0.00274866	274.866	363.813

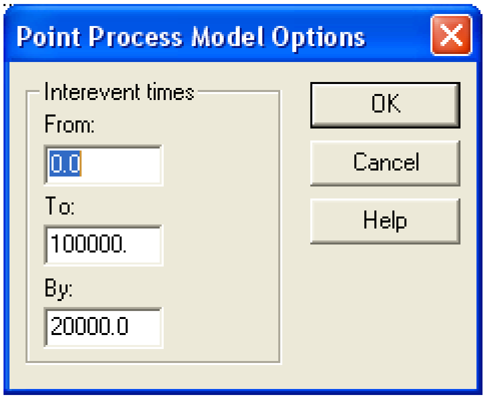
Goodness-of-Fit Test

	Test statistic	*P-Value*
Failures	34.5063	0.0108991

The output includes:

- *Rate model*: the estimated rate is $\lambda(t)$ = 0.00274866 for all *t*.
- *Mean cumulative events model*: the mean cumulative events function increases linearly according to $M(t)$ = 0.00274866*t*.
- *Interevent distribution*: the distribution of time between failures is exponential with an estimated MTBF of 363.813.
- 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* – performs a chi-square test to determine whether the selected model fits the data adequately. The observed numbers of events in the intervals are compared to the expected values using the fitted model. A small P-value, such as that shown for the sample data, implies that the selected model is not adequate.

Pane Options



- *Interevent Times From, To and By* – settings used to define the values of *t* at which estimates are displayed.

Example #2 – Nonhomogenous Poisson Process

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

Point Process Model
Model: nonhomogeneous Poisson process
Rate model: exp(-5.47206-0.0000327072*t+3.67983E-10*t^2)

Interevent distribution: Exponential
Mean = exp(5.47206+0.0000327072*t-3.67983E-10*t^2)

t	*Rate*	*Mean cum events*	*Mean interevent time*
0.0	0.00420256	0.0	237.95
20000.0	0.00253133	64.3457	395.049
40000.0	0.0020466	108.852	488.614
60000.0	0.00222111	150.474	450.226
80000.0	0.00323561	203.106	309.061
100000.	0.00632692	293.089	158.055

Goodness-of-Fit Test

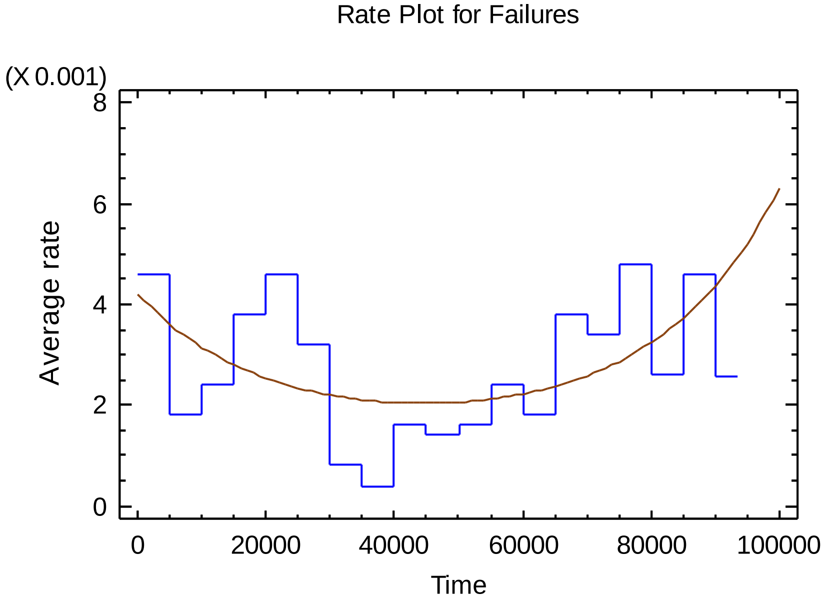
	Test statistic	*P-Value*
Failures	8.88652	0.962314

Note the following:

- *Rate model*: the estimated rate is $\lambda(t)$ = exp(-5.47 - 0.0000327t + 0.000000000370t^2).
- *Mean interevent time*: the MTBF increases as the product ages and then decreases.
- *Goodness-of-Fit Test* – the large P-value implies that the model fits the data adequately.

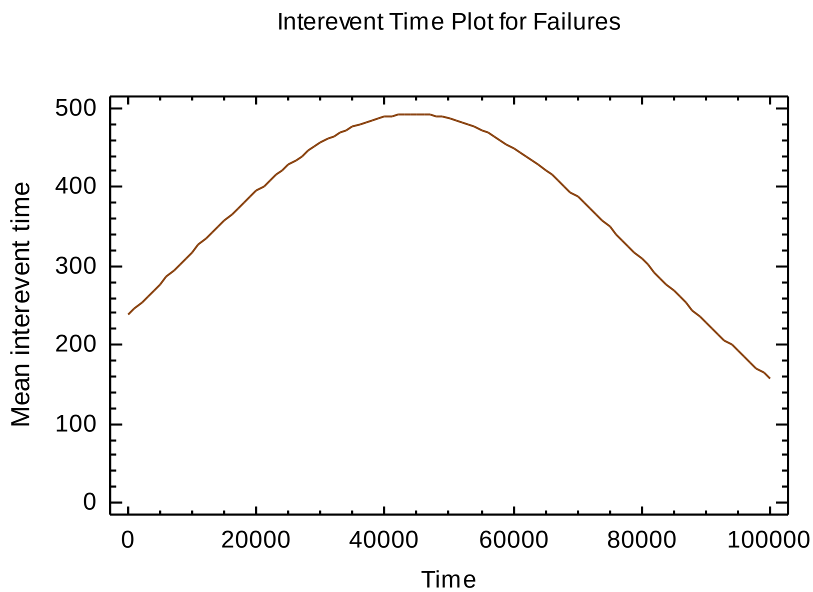
Failure Rate Plot

This plot shows the estimated failure rate, together with the observed failure rate of all samples combined:



Interevent Time Plot

This plot displays the estimated mean time between failures:



Model Percentiles

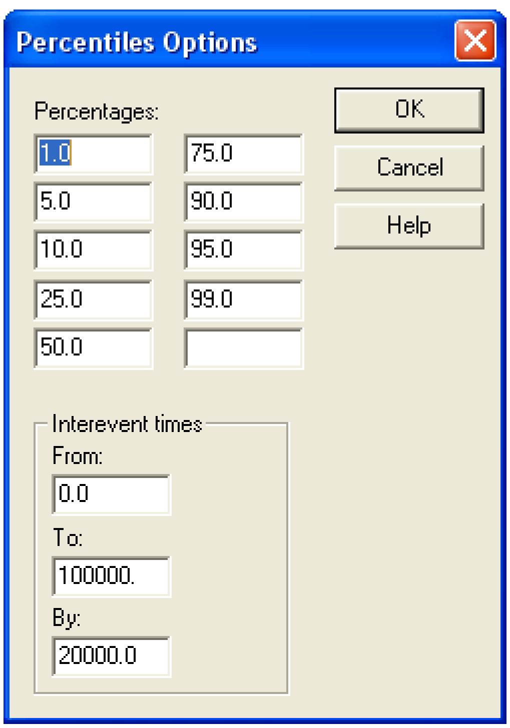
This table shows estimated percentiles for the distribution of time between consecutive failures:

Model Percentiles
Model: nonhomogeneous Poisson process
Rate model: exp(-5.47206-0.0000327072*t+3.67983E-10*t^2)

t	*1.0%*	*5.0%*	*10.0%*	*25.0%*	*50.0%*	*75.0%*	*90.0%*	*95.0%*	*99.0%*
0.0	2.39148	12.2052	25.0705	68.454	164.934	329.869	547.9	712.835	1095.8
20000.0	3.97038	20.2634	41.6226	113.649	273.827	547.655	909.635	1183.46	1819.27
40000.0	4.91074	25.0626	51.4807	140.566	338.682	677.363	1125.08	1463.76	2250.15
60000.0	4.52492	23.0936	47.4361	129.522	312.073	624.146	1036.68	1348.76	2073.37
80000.0	3.10617	15.8528	32.5628	88.9114	214.225	428.45	711.64	925.865	1423.28
100000.	1.5885	8.10715	16.6527	45.4695	109.555	219.111	363.935	473.49	727.869

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 68.45 or less. The 25th percentile changes as the system ages, reaching as high as 140.57 after t = 40000 hours.

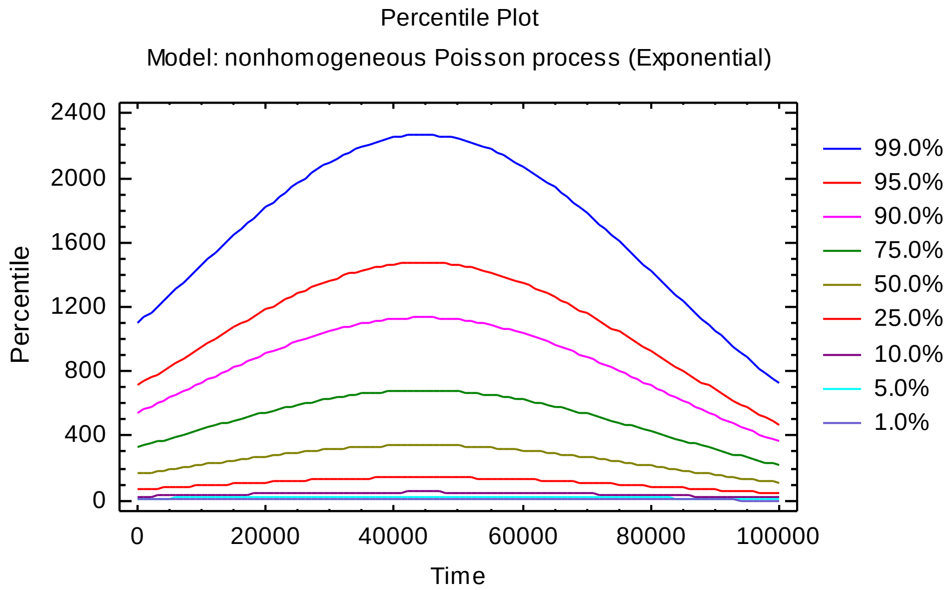
Pane Options



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

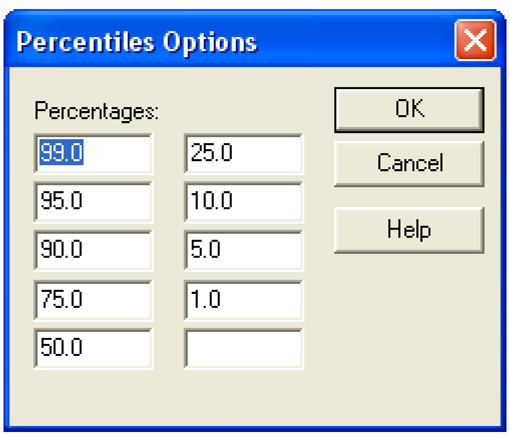
Interevent Time Percentile Plot

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.

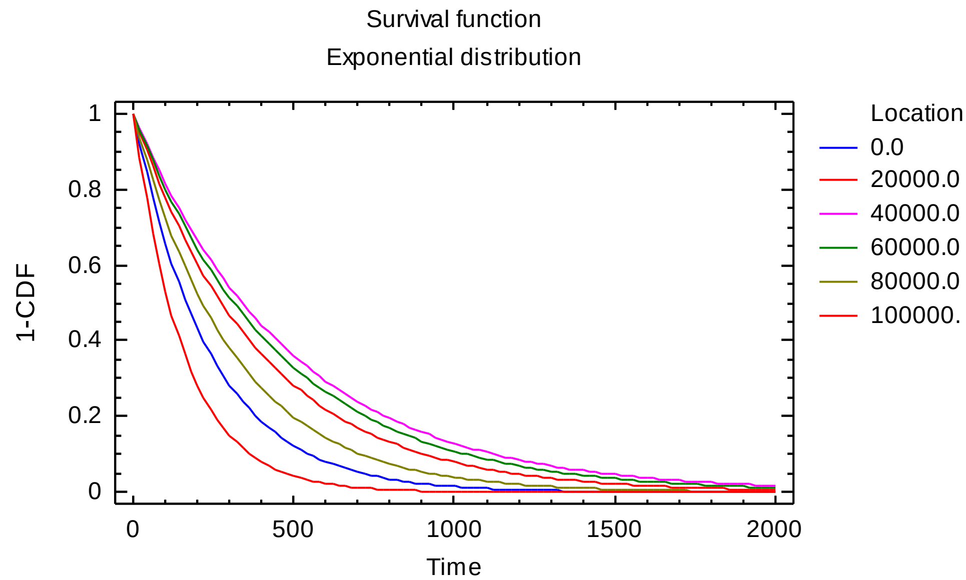
Pane Options



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

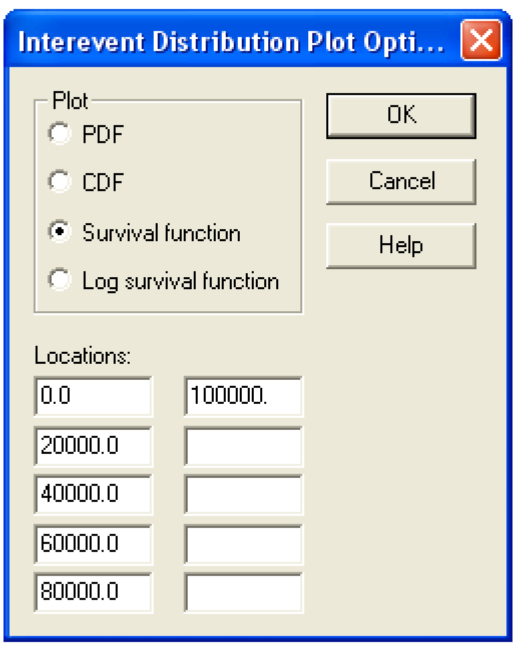
Interevent Time Distribution Plots

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 the position 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.

Calculations

Chi-Square Test for Trend

Let n_{iq} = number of failures in the i-th interval for system q for i=1,2,…,k. Let T_{iq} = width of the i-th interval for system q. For sample q, the test statistic is calculated by:

$$\chi_q^2 = \sum_{i=1}^{k} \frac{\left(n_{iq} - n_q \frac{T_{iq}}{T_q} \right)^2}{n_q \frac{T_{iq}}{T_q}}$$

which is compared to a chi-square distribution with k-1 degrees of freedom. (Note: intervals with expected counts less than 5 will be grouped together.)

For the combined samples, the counts in each interval are totaled and then the test is applied.