

## Accelerated Life Testing



Summary.....	1
Data Input .....	3
Statistical Model .....	4
Analysis Options .....	7
Analysis Summary .....	8
Plot of Fitted Model .....	9
Predictions .....	12
Percentiles.....	14
Percentile Plot.....	16
Percentile Probability Plot.....	18
Observed versus Predicted .....	21
Residual Probability Plot.....	21
Unusual Residuals.....	23
Residual Plots .....	24
Trellis Plot .....	28
Failure Time Distribution.....	30
Save Results.....	32
Calculations .....	33

### Summary

When testing items to estimate their failure time characteristics, it is often impossible to observe enough failures under normal operating conditions to obtain reasonable estimates of the failure time distribution. Consequently, one or more factors are often used to stress the production process so as to obtain more failures in a sufficiently short time period.

The **Accelerated Life Testing** procedure fits various types of accelerating models and extrapolates the results to normal operating conditions. In addition to a primary accelerating variable, which must be quantitative, additional predictors may be included in the models and may be either quantitative or categorical. The distribution of failure times may take any of seven

different forms, including a Weibull, exponential, normal, lognormal, logistic, loglogistic, or smallest extreme value distribution. Failure times may be censored or uncensored.

The output of the procedure includes an estimate of the survivor function and failure time distribution. Predictions may be calculated from the fitted model and unusual residuals identified.

**Sample StatFolio:** *accellifetest.sgp*

**Sample Data:**

The *accellifetest.sgp* contains a dataset measuring the failure times of 25 items subjected to various levels of two accelerating factors: *temperature* and *voltage*. The data are shown below:

<i>Temperature</i>	<i>Voltage</i>	<i>Hours</i>	<i>Censored</i>	<i>CHours</i>
85	6	500	1	>500
85	6	500	1	>500
85	6	500	1	>500
85	6	480	0	480
85	6	475	0	475
85	8	350	0	350
85	8	325	0	325
85	8	315	0	315
85	8	330	0	330
85	8	310	0	310
45	12	500	1	>500
45	12	500	1	>500
45	12	475	0	475
45	12	495	0	495
45	12	450	0	450
65	12	250	0	250
65	12	230	0	230
65	12	245	0	245
65	12	210	0	210
65	12	200	0	200
85	12	60	0	60
85	12	55	0	55
85	12	70	0	70
85	12	65	0	65
85	12	55	0	55
25	4			

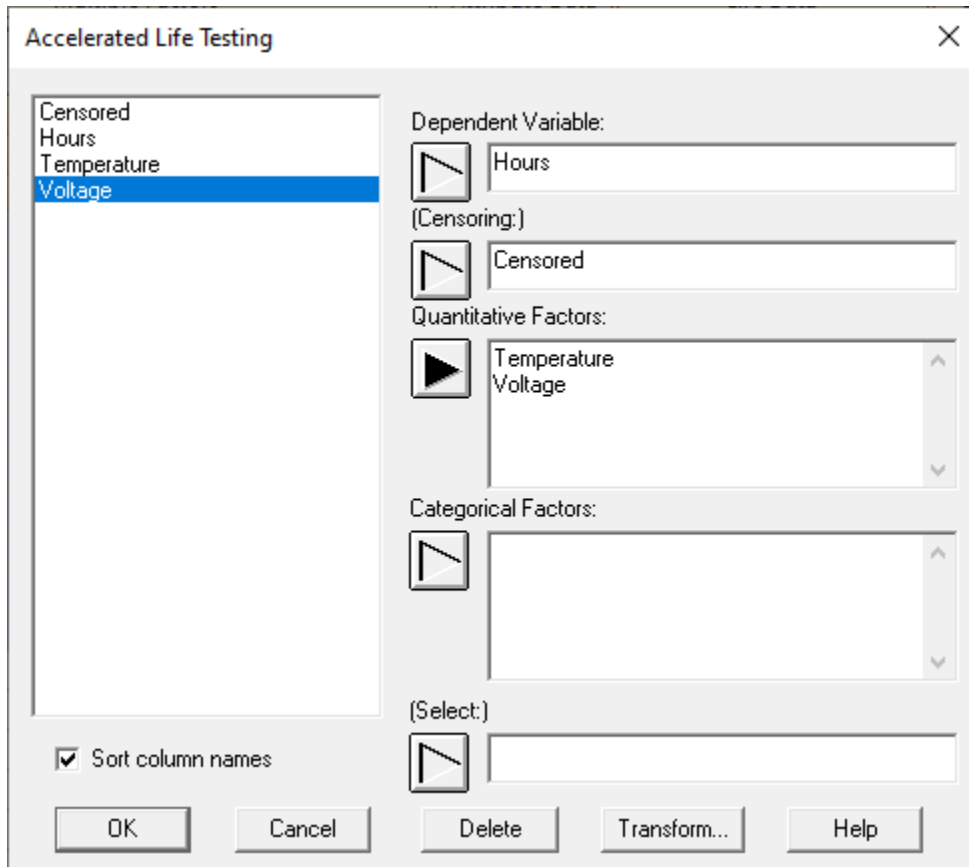
Data were collected at *Temperature* = 45, 65, and 85 degrees centigrade and at *Voltage* = 6, 8 and 12 volts. *Hours* represents either the time until each item failed or 500 if the item had not failed after 500 hours. *Censored* is set to 0 for actual failure times or to 1 if an item did not fail by 500 hours. An additional column named *CHours* is also shown, which is a special censored

numeric data column. In that column, right-censored values are represented by the notation  $>500$ .

The goal of the study is to estimate the failure time distribution at normal operating conditions where  $Temperature = 25$  and  $Voltage = 4$ . Note that these values have been added at the bottom of the datasheet, with the cell for  $Hours$  left empty. The procedure will automatically recognize that combination of factors as one for which predictions are desired.

## Data Input

The data input dialog box requests information about the failure times and the predictor variables:



- **Dependent Variable:** a variable containing Y, the failure times (for uncensored data) or censoring times (for censored data). It may be either of two types:
  1. A *numeric* column similar to  $Hours$  containing the failure times and censoring times. If censored data is present, it must be paired with a second column such as  $Censored$  containing censoring indicators.
  2. A special *censored numeric* column similar to  $CHours$  containing both failure times and the censoring information. If interval censoring is present, this option

must be used. Interval censored data is entered using notation such as  $[200,250]$  to indicate the interval containing the actual failure time.

- **(Censoring):** an optional column indicating whether or not each data value has been censored. Enter 0 if the value of the dependent variable represents an uncensored failure time. Enter 1 if the value has been right-censored (the true failure time is greater than the value entered) or -1 if the value has been left-censored (the true failure time is less than the value entered). This field should be left blank if the dependent variable is a censored numeric column.
- **Quantitative Factors:** numeric columns containing the values of any quantitative factors to be included in the model. There must be at least 1 quantitative factor. The first factor entered is considered to be the primary accelerating variable.
- **Categorical Factors:** numeric or non-numeric columns containing the levels of any categorical factors to be included in the model.
- **Select:** subset selection.

## Statistical Model

STATGRAPHICS fits two types of accelerating life data regression models: location-scale regression models and log-location-scale regression models. Each contains a term involving one of 7 common acceleration models for the primary accelerating variable, referred to as  $X_1$  in the equations below.

*Arrhenius model:*

$$Y = A \exp\left(-\frac{B}{kX_1}\right) \quad (1)$$

where  $X_1$  is temperature in degrees Kelvin ( $^{\circ}\text{C} + 273.15$ ),  $k = 0.00008617$  (Boltzmann's constant), and  $A$  and  $B$  are two unknown parameters. Taking logarithms of both sides:

$$\ln(Y) = \ln(A) - \frac{B}{kX_1} \quad (2)$$

showing that the log of  $Y$  is linearly related to the reciprocal of  $X_1$ .

*Inverse power rule model:*

$$Y = \frac{A}{X_1^B} \quad (3)$$

where  $X_1$  is the value of the accelerating variable and  $A$  and  $B$  are unknown parameters. The model can be linearized by taking logarithms of both sides:

$$\ln(Y) = \ln(A) - B \ln(X_1) \quad (4)$$

*Exponential model:*

$$Y = Ae^{-BX_1} \quad (5)$$

where  $X_1$  is the value of the accelerating variable and  $A$  and  $B$  are unknown parameters. The model can be linearized by taking logarithms of both sides:

$$\ln(Y) = \ln(A) - BX_1 \quad (6)$$

*Eyring model:*

$$Y = \frac{1}{X_1} e^{-\left(A - \frac{B}{X_1}\right)} \quad (7)$$

where  $X_1$  is the value of the accelerating variable and  $A$  and  $B$  are unknown parameters. It can be linearized as

$$\ln(Y) = -A + \frac{B}{X_1} - \ln(X_1) \quad (8)$$

*Linear model:*

$$Y = A + BX_1 \quad (9)$$

where  $X_1$  is the value of the accelerating variable and  $A$  and  $B$  are unknown parameters.

*Reciprocal model:*

$$Y = A + B/X_1 \quad (10)$$

### Location-Scale Models

If the selected acceleration model is either linear or reciprocal, a location-scale model is fit to the data. For this type of model, the failure times are related to the predictor variables through a linear function of the form

$$Y = \mu + \Phi^{-1}(p)\sigma \quad (11)$$

where  $\mu$  is a location parameter that depends on the predictor variables,  $\sigma$  is a scale parameter, and  $\Phi^{-1}(p)$  is the standardized inverse cdf of the lifetime distribution, i.e.,

$$F(Y) = \Phi\left(\frac{Y - \mu}{\sigma}\right) \quad (12)$$

The location parameter is related to the predictor variables by

$$\mu = m(X_1) + \beta_2 X_2 + \dots + \Phi^{-1}(p)\sigma \quad (13)$$

where  $m(X_1)$  is the selected accelerating model. For such a model, lifetimes may be assumed to follow either a normal, logistic, or smallest extreme value distribution.

### Log-Location-Scale Models

If the selected acceleration model is either Arrhenius, inverse power, exponential, or Eyring, a log-location-scale model is fit to the data. For this type of model, the logarithms of the failure times are related to the predictor variables through a function of the form

$$\ln(Y) = \mu + \Phi^{-1}(p)\sigma \quad (14)$$

where

$$F(Y) = \Phi\left(\frac{\log(Y) - \mu}{\sigma}\right) \quad (15)$$

The location parameter is related to the predictor variables by

$$\mu = \ln(m(X_1)) + \beta_2 X_2 + \dots + \Phi^{-1}(p)\sigma \quad (16)$$

where  $m(X_1)$  is the selected accelerating model. For such a model, lifetimes may be assumed to follow either a lognormal, loglogistic, Weibull, or exponential distribution.

For the sample data, the following model will be fit:

$$Y = AV^c \exp\left(\frac{-B}{kT}\right) \quad (17)$$

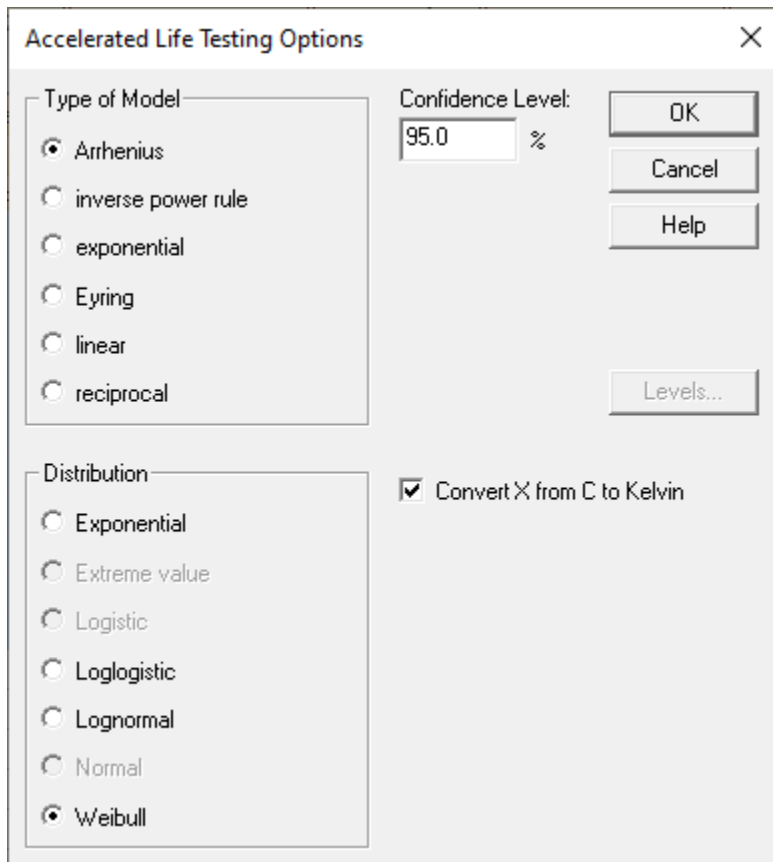
which when linearized is

$$\ln(Y) = \ln(A) + c \ln(V) - \frac{B}{kT} \quad (18)$$

The primary accelerating variable is *Temperature* (T) which follows an Arrhenius model. The second accelerating variable is *Voltage* (V) whose effect on *Hours* is based on a power rule. It will further be assumed that the logarithm of time to failure follows a Weibull distribution.

## Analysis Options

The *Analysis Options* dialog box specifies the accelerating model and other options:



- **Type of model:** the acceleration model for the primary accelerating variable  $m(X_1)$ .
- **Distribution:** the failure time distribution.
- **Confidence level:** the level used when calculating confidence limits.
- **Convert X from C to Kelvin:** whether the primary accelerating variable  $X_1$  should be converted from centigrade to Kelvin. This box should be checked when using models such as the Arrhenius model if the primary accelerating variable is temperature and the data have been entered in degrees centigrade.

## Analysis Summary

The *Analysis Summary* displays a table showing the estimated model and likelihood ratio tests for the significance of the model coefficients.

<b>Accelerated Life Testing - Hours</b>				
Dependent variable: Hours				
Censoring: Censored				
Factors:				
Temperature (degrees C)				
Voltage				
Number of uncensored values: 20				
Number of right-censored values: 5				
Number of left-censored values: 0				
Number of interval-censored values: 0				
Arrhenius model: Hours=A*exp(-B/(0.00008617*(Temperature+273.15)))				
Distribution: Weibull				
<b>Estimated Regression Model</b>				
<i>Parameter</i>	<i>Estimate</i>	<i>Standard Error</i>	<i>Lower 95.0% Conf. Limit</i>	<i>Upper 95.0% Conf. Limit</i>
<b>CONSTANT: log(A)</b>	-8.92567	1.46672	-11.8004	-6.05095
<b>Temperature: B</b>	-0.542142	0.0497102	-0.639572	-0.444712
<b>Voltage</b>	-0.361822	0.0258886	-0.412562	-0.311081
<b>SIGMA</b>	0.150018	0.0254383	0.107598	0.20916
Log likelihood = -102.784				
<b>Likelihood Ratio Tests</b>				
<i>Factor</i>	<i>Chi-Square</i>	<i>Df</i>	<i>P-Value</i>	
<b>Temperature</b>	65.0475	1	0.0000	
<b>Voltage</b>	67.8435	1	0.0000	

The table includes:

- **Data Summary:** a summary of the input data, including the number of observations  $n$  used to fit the model.
- **Model and distribution:** the selected primary acceleration model and lifetime distribution.
- **Estimated Regression Model:** estimates of the coefficients in the regression model, with standard errors and approximate confidence intervals.
- **Likelihood Ratio Tests:** tests run to determine whether or not the coefficients are significantly different from 0. Two-sided P-values are displayed. Small P-values (less than 0.05 if operating at the 5% significance level) correspond to statistically significant variables.

The above table shows the result of fitting model (18) to the sample data, assuming a Weibull distribution for the failure times at fixed values of the predictor variables. The estimated model has parameters:

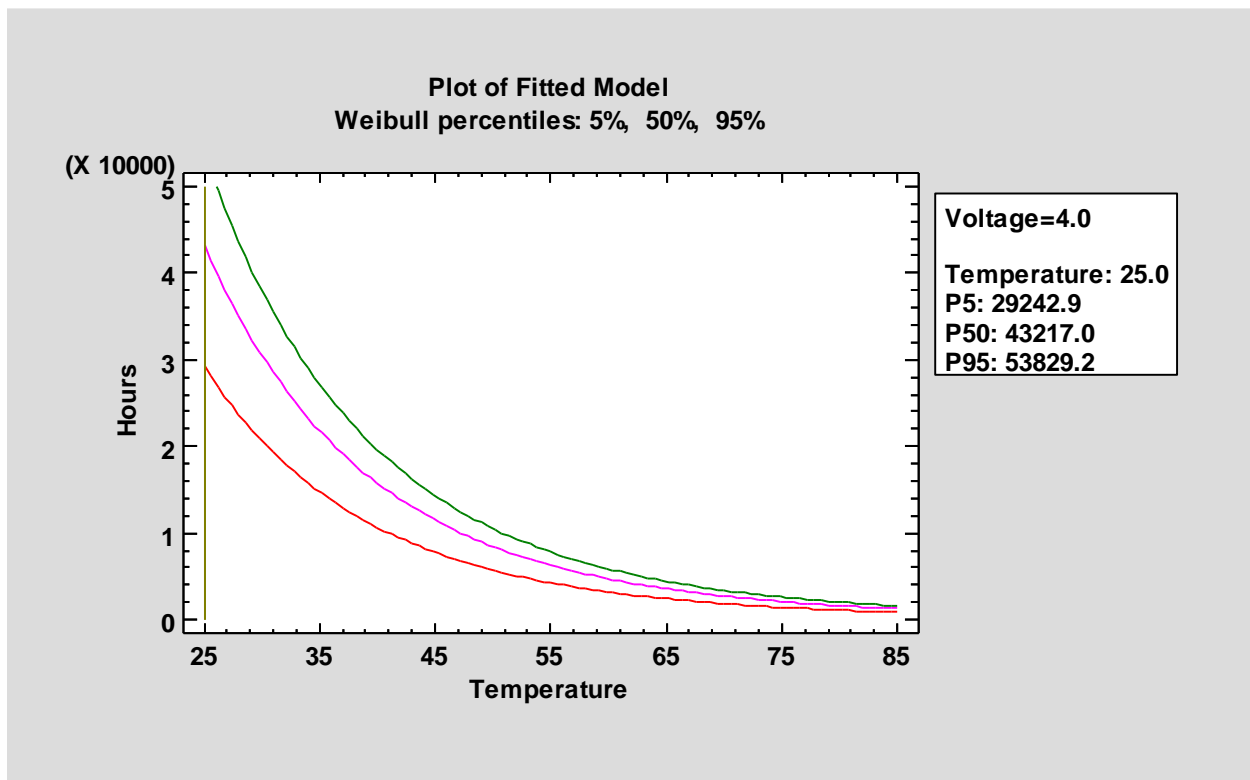
$$\mu = -8.92567 + 0.542142/(0.00008617*Temperature) - 0.361822*Voltage \quad (19)$$



based on a log-linear model with  $Y = \ln(\text{hours})$ . Both *Temperature* and *Voltage* have a statistically significant negative effect on item lifetimes.

## Plot of Fitted Model

The *Plot of Fitted Model* pane displays the percentiles as a function of any single variable X with all other variables set fixed at specified values.



For example, the above plot shows how the 5-th, 50-th, and 95-th percentiles vary as a function of *Temperature*, with *Voltage* set equal to 4. The percentiles of the failure time distribution decrease as the temperature increases, with the variability decreasing as well.

### *Pane Options*

Plot of Fitted Model Options

	Low	High	Hold
<input checked="" type="radio"/> Temperature	25.0	85.0	25.0
<input type="radio"/> Voltage	4.0	12.0	4.0
<input type="radio"/>			
<input type="radio"/>			
<input type="radio"/>			
<input type="radio"/>			
<input type="radio"/>			
<input type="radio"/>			
<input type="radio"/>			
<input type="radio"/>			
<input type="radio"/>			
<input type="radio"/>			
<input type="radio"/>			
<input type="radio"/>			
<input type="radio"/>			
<input type="radio"/>			
<input type="radio"/>			
<input type="radio"/>			
<input type="radio"/>			
<input type="radio"/>			
<input type="radio"/>			
<input type="radio"/>			

Percentiles

5.0 %

50.0 %

95.0 %

%

%

Plot mean

Use Arrhenius scaling

Show density function

Show results at:

25.0

OK

Cancel

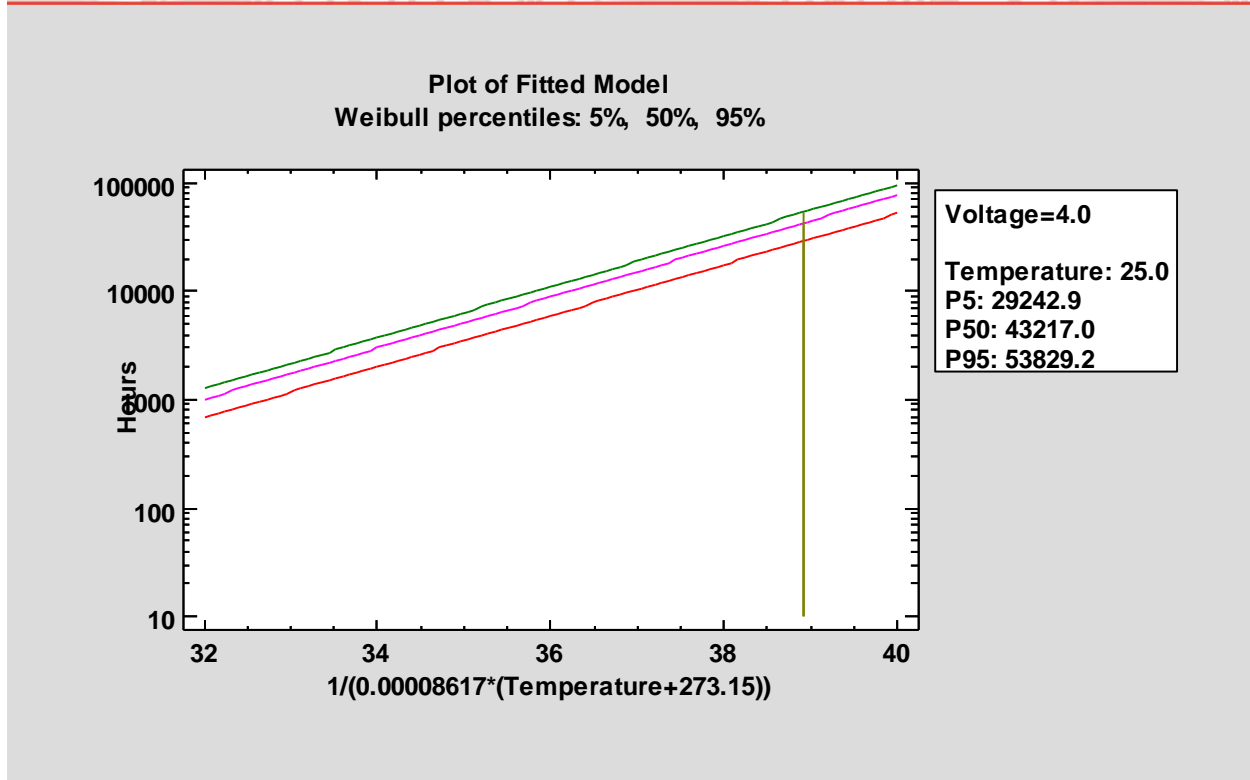
Help

Next

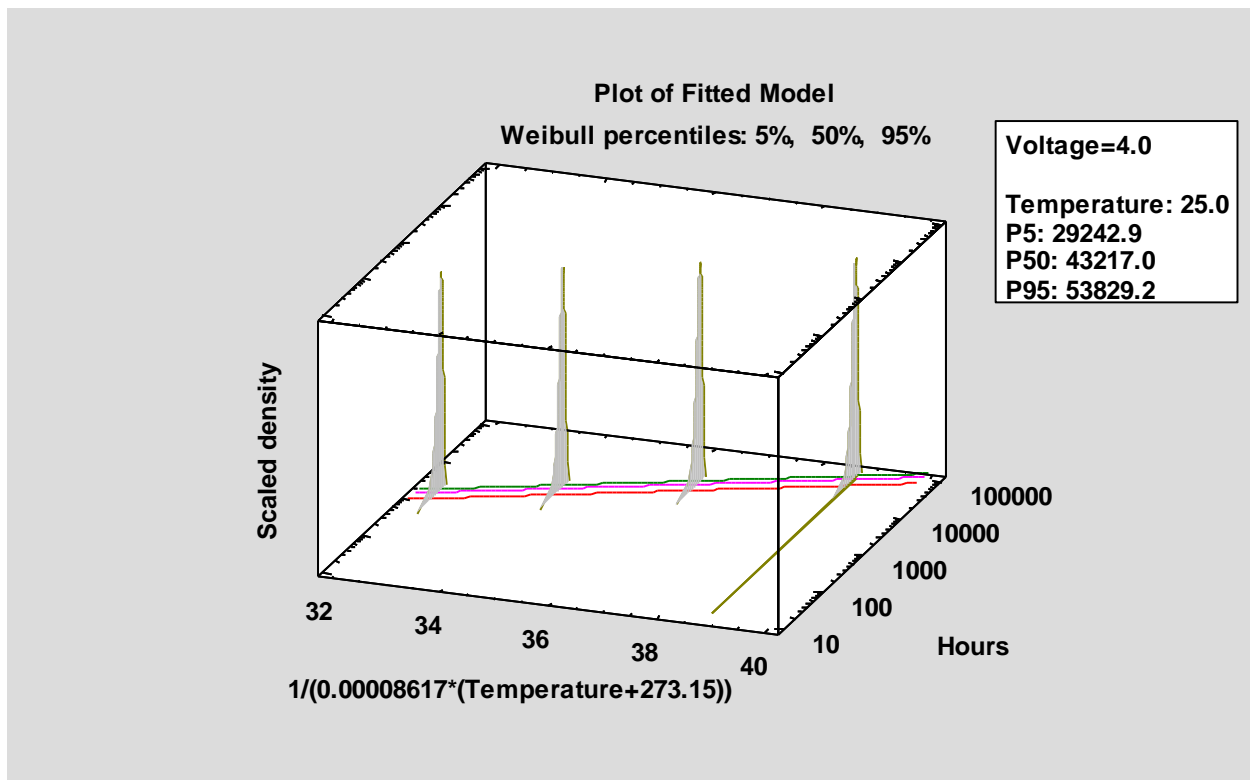
Back

- **Factors:** select one factor to plot on the horizontal axis, with lower and upper limits for the plot. For all other factors, specify values at which they should be held fixed.
- **Percentiles:** percentages corresponding to the desired percentiles.
- **Plot Mean:** if checked, a line is drawn at the estimated mean failure time MTTF.
- **Use Arrhenius scaling:** scale the plot so that the horizontal axis displays  $1/(.00008617 * X)$  and the vertical scale is logarithmic.
- **Show density function:** displays a scaled version of the probability density function at selected values of X.
- **Show results at:** if checked, an additional line is added at the indicated value of X.
- **Next and Back:** used to display other factors when more than 16 are present.

The plot below displays the model using Arrhenius scaling with an additional line at *Temperature* = 25 degrees C.



If “Show density function” is selected, a 3-dimensional plot will be created which adds an illustration of the probability density function at selected values of X. An example is shown below:



## Predictions

The *Predictions* pane creates predictions using the fitted model. By default, the table includes a line for each row in the datasheet that has complete information on the X variables and a missing value for the Y variable. This allows you to add columns to the bottom of the datasheet corresponding to levels at which you want predictions without affecting the fitted model.

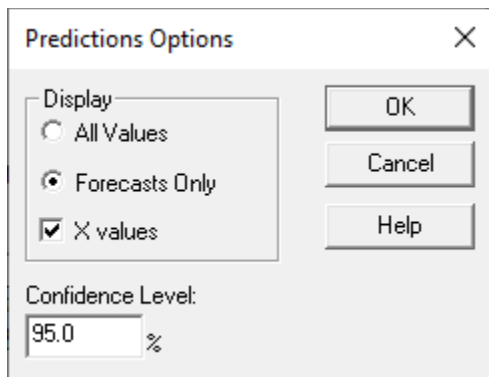
In the current example, a row has been added to the bottom of the datasheet containing the normal operating levels of *Temperature* and *Voltage*. The value of the fitted failure time model for that combination is shown below:

Predictions for Hours						
Row	Temperature	Voltage	Fitted Value	Standard Error	Lower 95.0% CL for Fitted Value	Upper 95.0% CL for Fitted Value
26	25.0	4.0	45659.8	18849.9	20329.6	102551.

Included in the table are:

- **Row** - the row number in the datasheet.
- **X values** – values of the predictor variables.
- **Observed Value** - the observed values,  $Y_i$ , if any
- **Fitted Value** - the fitted values, given by  $\hat{\mu}_i$  for location-scale models and  $\exp(\hat{\mu}_i)$  for log-location-scale models.
- **Standard Error** – the standard errors corresponding to  $\hat{\mu}_i$  for location-scale models and  $\exp(\hat{\mu}_i)$  times the standard errors corresponding to  $\hat{\mu}_i$  for log-location-scale models.
- **Confidence Limits** – approximate confidence limits for the *Fitted Values*.

### Pane Options



The dialog box titled "Predictions Options" contains the following settings:

- Display:**
  - All Values
  - Forecasts Only
  - X values
- Confidence Level:** 95.0 %
- Buttons:** OK, Cancel, Help

- **Display All Values or Forecasts Only** – whether to display all rows or only rows with missing values for Y.
- **X values** – whether to display values of the predictor variables.
- **Confidence Level** - level of confidence for the interval estimates.

## Percentiles

The *Percentiles* pane displays a table of estimated percentiles at a selected combination of the predictor variables. It also displays the estimated mean time to failure (MTTF).

<b>Failure Time Percentiles for Hours</b>				
Temperature=25.0				
Voltage=4.0				
MTTF: 42602.2 (18973.7,95655.9)				
<b>Percent</b>	<b>Percentile</b>	<b>Standard Error</b>	<b>Lower 95.0% Conf. Limit</b>	<b>Upper 95.0% Conf. Limit</b>
0.1	16199.9	7198.1	6781.09	38701.5
0.5	20630.1	8887.91	8867.17	47997.5
1.0	22899.5	9756.98	9934.5	52784.4
2.0	25428.2	10731.0	11119.8	58147.7
3.0	27043.6	11357.0	11874.1	61592.7
4.0	28258.1	11829.9	12439.4	64193.0
5.0	29242.9	12214.7	12896.5	66308.2
6.0	30077.4	12541.9	13283.0	68105.6
7.0	30805.5	12828.2	13619.6	69677.4
8.0	31454.0	13083.9	13918.8	71080.4
9.0	32040.5	13315.7	14188.9	72351.9
10.0	32577.5	13528.3	14435.9	73517.9
15.0	34766.1	14399.5	15438.4	78290.6
20.0	36459.4	15078.8	16209.6	82006.3
25.0	37875.7	15650.4	16851.6	85129.6
30.0	39117.1	16154.0	17412.0	87878.8
35.0	40240.8	16612.0	17917.4	90376.8
40.0	41282.8	17038.7	18384.5	92701.5
45.0	42268.5	17443.8	18824.9	94907.6
50.0	43217.0	17835.2	19247.4	97036.9
55.0	44144.2	18219.2	19659.2	99124.8
60.0	45064.9	18601.9	20066.8	101204.
65.0	45994.0	18989.5	20477.0	103308.
70.0	46949.1	19389.3	20897.4	105478.
75.0	47952.9	19811.2	21337.8	107765.
80.0	49038.7	20269.3	21812.6	110248.
85.0	50263.5	20788.4	22346.2	113058.
90.0	51745.5	21419.7	22989.1	116472.
91.0	52094.0	21568.7	23139.8	117277.
92.0	52468.5	21729.0	23301.7	118144.
93.0	52875.6	21903.5	23477.3	119086.
94.0	53324.5	22096.2	23670.7	120127.
95.0	53829.2	22313.2	23887.9	121299.
96.0	54412.5	22564.6	24138.4	122656.
97.0	55115.7	22868.3	24439.9	124294.
98.0	56027.9	23263.4	24829.9	126425.
99.0	57415.9	23867.2	25421.2	129679.
99.5	58636.4	24400.5	25939.1	132550.
99.9	61016.8	25447.2	26943.4	138180.

Confidence intervals are included based on a large-sample normal approximation. For example, at a temperature = 25 and a voltage = 4, it is estimated that 50% of the items will have failed after approximately 43,217 hours. The 95% confidence interval for the 50-th percentile ranges from 19,247 hours to 97,037 hours.

*Pane Options*

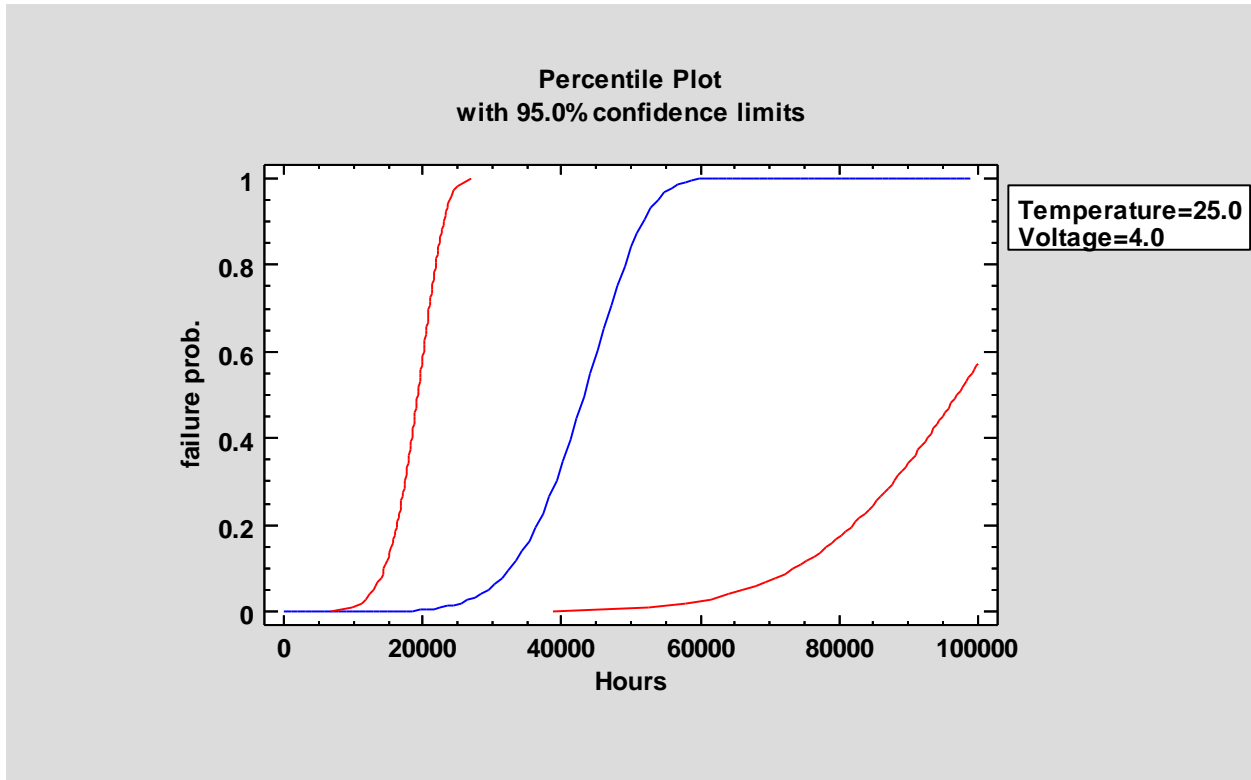
The dialog box titled 'Factor Level Options' contains the following elements:

- Factor**: A list of predictor variables. 'Temperature' and 'Voltage' are currently visible.
- Level**: Input fields for each factor. '25.0' is entered for Temperature, and '4.0' is entered for Voltage. There are 14 empty input fields below.
- Confidence Level**: A field containing '95.0' followed by a percentage symbol (%).
- Buttons**: 'OK', 'Cancel', and 'Help' are on the right side. 'Next' and 'Back' are below them.

- **Level**: values of the predictor variables at which the percentiles are to be estimated.
- **Confidence Level**: percentage confidence for the interval estimates.
- **Next** and **Back**: used to display other factors when more than 16 are present.

## Percentile Plot

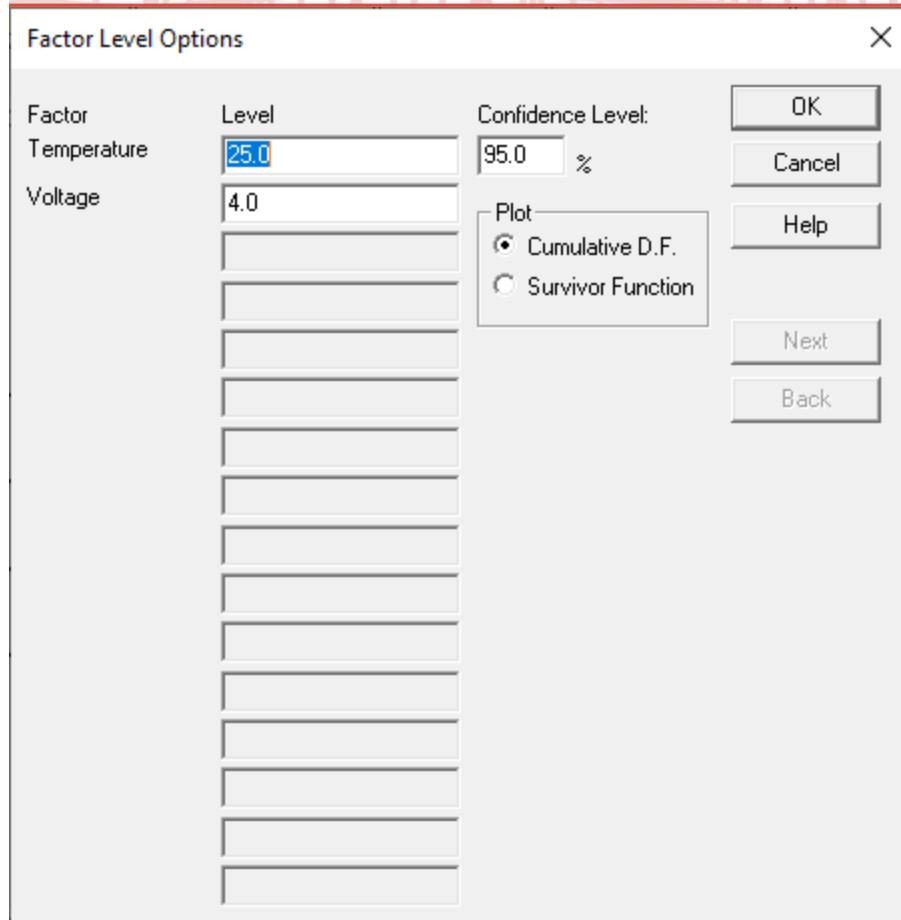
The *Percentile Plot* graphs the estimated percentiles at a selected combination of the predictor variables.



Confidence intervals are included based on a large-sample normal approximation.

### *Pane Options*





The dialog box is titled "Factor Level Options" and contains the following elements:

Factor	Level
Temperature	25.0
Voltage	4.0

Confidence Level: 95.0 %

Plot:

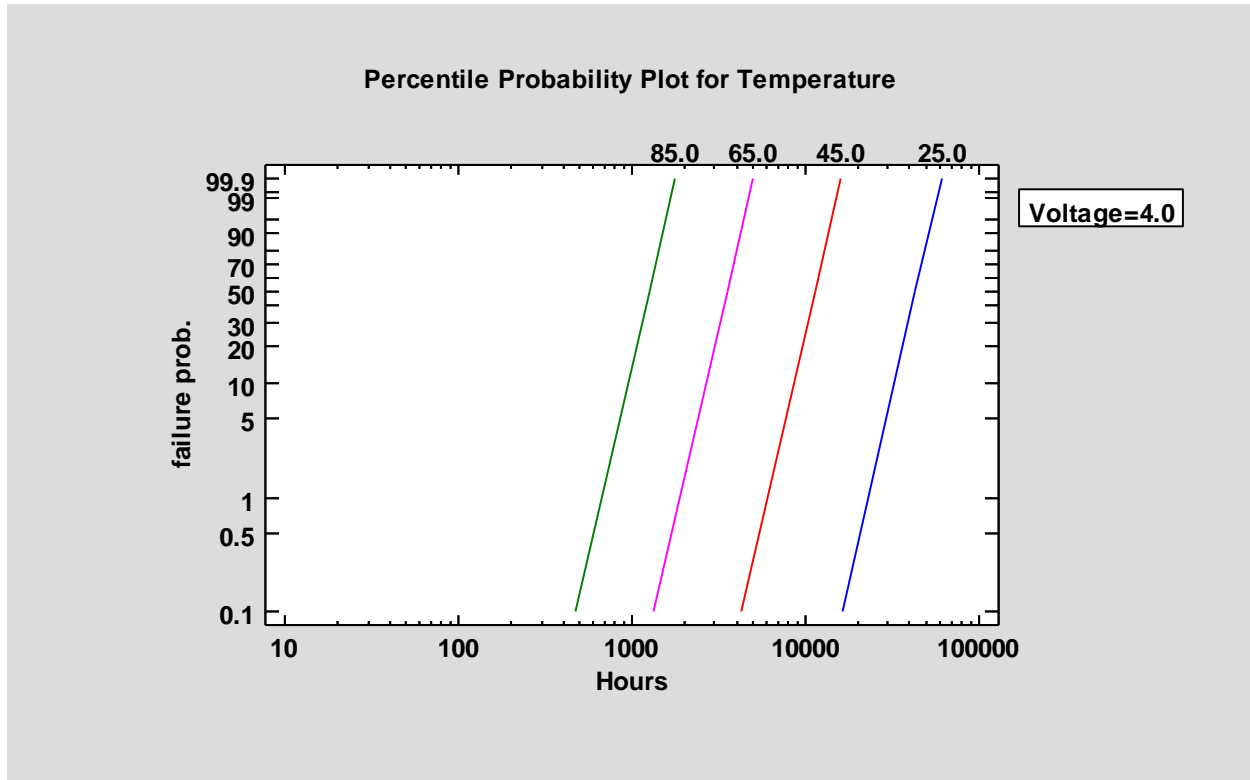
- Cumulative D.F.
- Survivor Function

Buttons: OK, Cancel, Help, Next, Back

- **Level:** value of the predictor variables at which the percentiles are to be estimated.
- **Confidence Level:** percentage confidence for the interval estimates.
- **Plot:** select *Cumulative D.F.* to plot the estimated failure probabilities or *Survivor Function* to plot the estimated survival probabilities.
- **Next** and **Back:** used to display other factors when more than 16 are present.

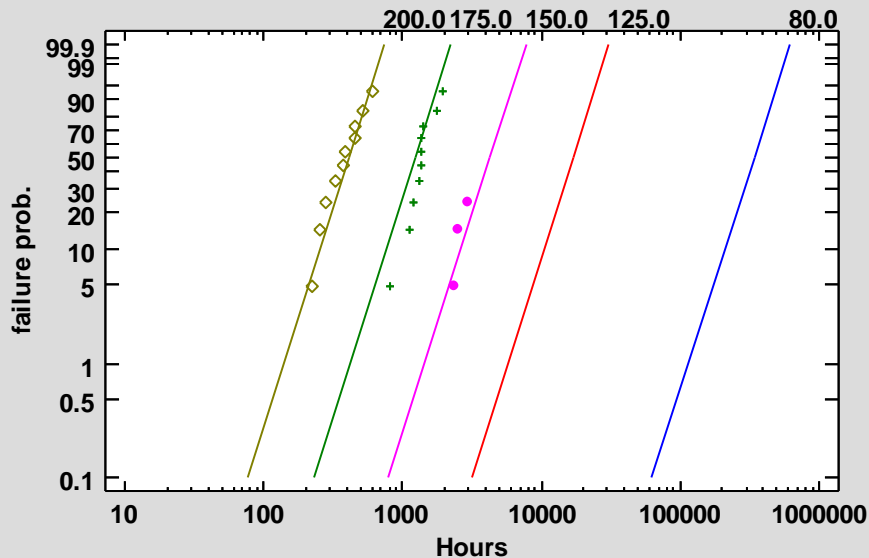
## Percentile Probability Plot

This plot graphs the estimated percentiles on a chart scaled so that the cumulative distribution function is a straight line. Lines are drawn at each observed value of a selected variable.



If the model contains only a single variable, the values of any uncensored observations are also displayed. A typical example (for a different dataset) is show below:

Percentile Probability Plot for Temperature



The  $m$  observed values for each level of the variable are sorted from lowest to highest and plotted at vertical locations equal to  $100(i+0.5)/(m+0.25)$  where  $i$  is the index of each observation after sorting. If the assumed distribution fits the observed data well, the observations should fall close to the cumulative distributions.

### Pane Options

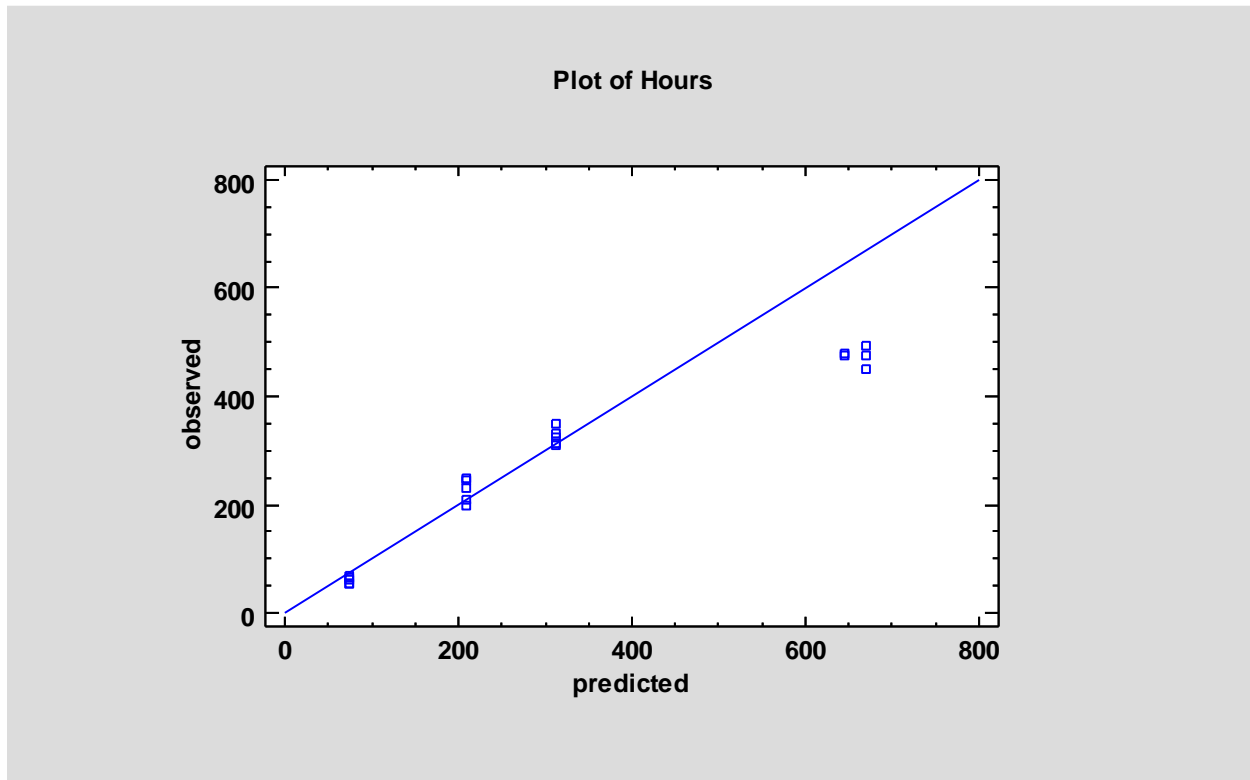
For models containing more than 1 variable, select the variable to be used to define the lines in the plot and values at which to hold the other variables fixed.

Percentile Probability Plot Options ✕

	Hold	
<input checked="" type="radio"/> Temperature	25.0	OK
<input type="radio"/> Voltage	4.0	Cancel
<input type="radio"/>		Help
<input type="radio"/>		
<input type="radio"/>		Next
<input type="radio"/>		Back
<input type="radio"/>		
<input type="radio"/>		
<input type="radio"/>		
<input type="radio"/>		
<input type="radio"/>		
<input type="radio"/>		
<input type="radio"/>		
<input type="radio"/>		
<input type="radio"/>		
<input type="radio"/>		
<input type="radio"/>		
<input type="radio"/>		
<input type="radio"/>		
<input type="radio"/>		
<input type="radio"/>		
<input type="radio"/>		

## Observed versus Predicted

The *Observed versus Predicted* pane plots any observed uncensored failure times  $Y_i$  versus  $\hat{\mu}_i$  for location-scale models and versus  $\exp(\hat{\mu}_i)$  for log-location-scale models.



If the model fits well, the points should be randomly scattered around the diagonal line. There is some suggestion in the above graph that the model is overpredicting the longest failure times, although interpretation is complicated by the absence of any censored values.

## Residual Probability Plot

In all regression applications, it is important to calculate and plot the residuals. The *Accelerated Life Testing* procedure creates three different types of residuals:

1. *Ordinary residuals*:

$$\text{for location-scale models: } r_i = y_i - \hat{\mu}_i \quad (21)$$

$$\text{for log-location-scale models: } r_i = y_i - \exp(\hat{\mu}_i) \quad (22)$$

2. *Standardized residuals*:

for location-scale models:  $e_i = \frac{y_i - \hat{\mu}}{\hat{\sigma}}$  (23)

for log-location-scale models:  $e_i = \exp\left(\frac{\ln(y_i) - \hat{\mu}_i}{\hat{\sigma}}\right)$  (24)

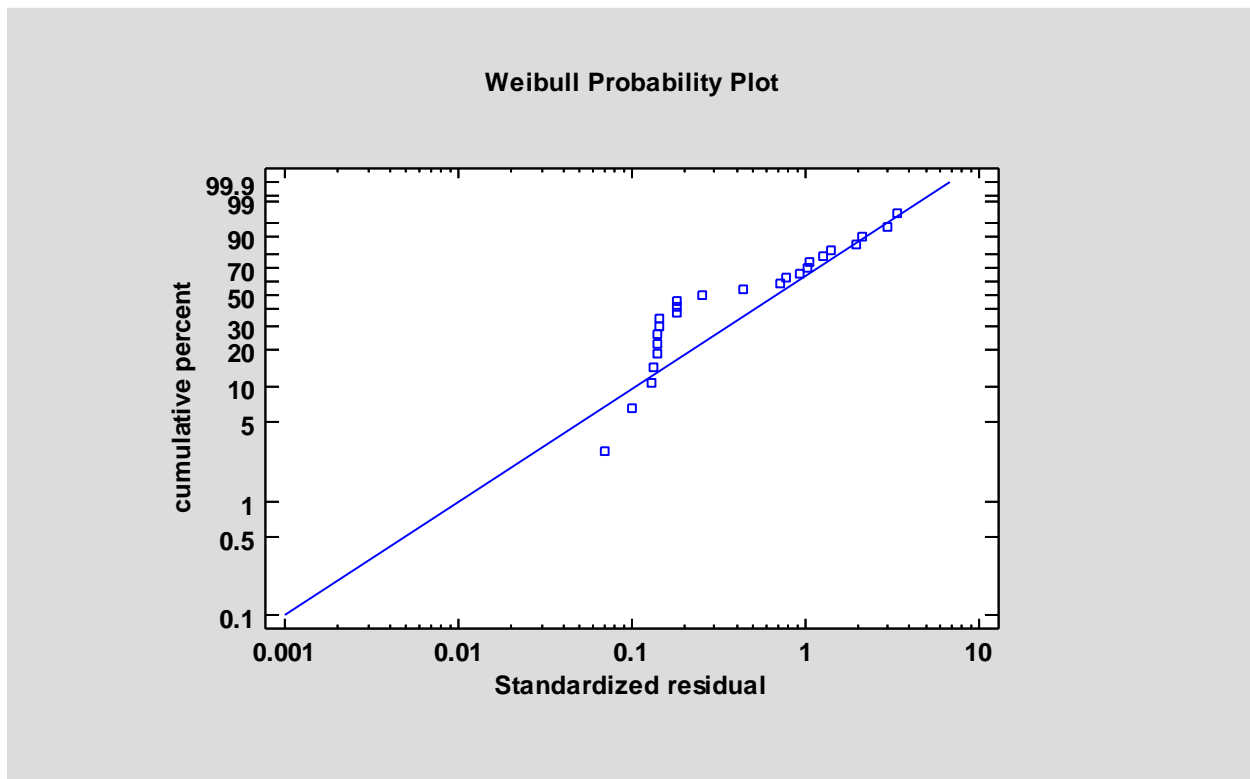
3. *Cox-Snell Residuals* – a type of Cox-Snell residuals constrained to lie between 0 and 1, defined by:

$$\hat{u}_i = \hat{F}(Y_i) \quad (25)$$

which is the estimated cumulative lifetime distribution evaluated at the observed failure time.

The *ordinary residuals* quantify the difference between the observed data values and the fitted values. The *standardized residuals* are scaled so that they should follow a standardized form of the assumed failure time distribution. The Cox-Snell residuals can be useful in identifying outliers.

The *Residual Probability Plot* displays the standardized residuals on a plot designed to help determine whether the assumed distribution of lifetimes is reasonable for the data:



If the selected distribution is adequate for the data, the points should lie along the diagonal reference line.

## Unusual Residuals

The *Unusual Residuals* pane lists all observations that have unusually large residuals.

Row	Y	Predicted Y	Residual	Standardized Residual	Cox-Snell Residual
6	350.0	313.103	36.897	2.10	0.8777
16	250.0	208.146	41.8542	3.39	0.9663
18	245.0	208.146	36.8542	2.96	0.9484

The table displays:

- *Row* – the row number in the data sheet.
- *Y* – the observed failure time.
- *Predicted Y* - the fitted values, given by  $\hat{\mu}_i$  for location-scale models and  $\exp(\hat{\mu}_i)$  for log-location-scale models.
- *Residual* – the ordinary residuals.
- *Standardized Residuals* – the standardized  $e_i$ .
- *Cox-Snell Residuals* – the Cox-Snell residuals  $\hat{u}_i$ .

A row is added to the list corresponding to all Cox-Snell residuals that are less than 0.025 or greater than 0.975, i.e., any residuals outside of the central 95% of the estimated lifetime distribution. Particular attention should be given to any residuals outside of the interval

$$0.00135 \leq \hat{u}_i \leq .99865$$

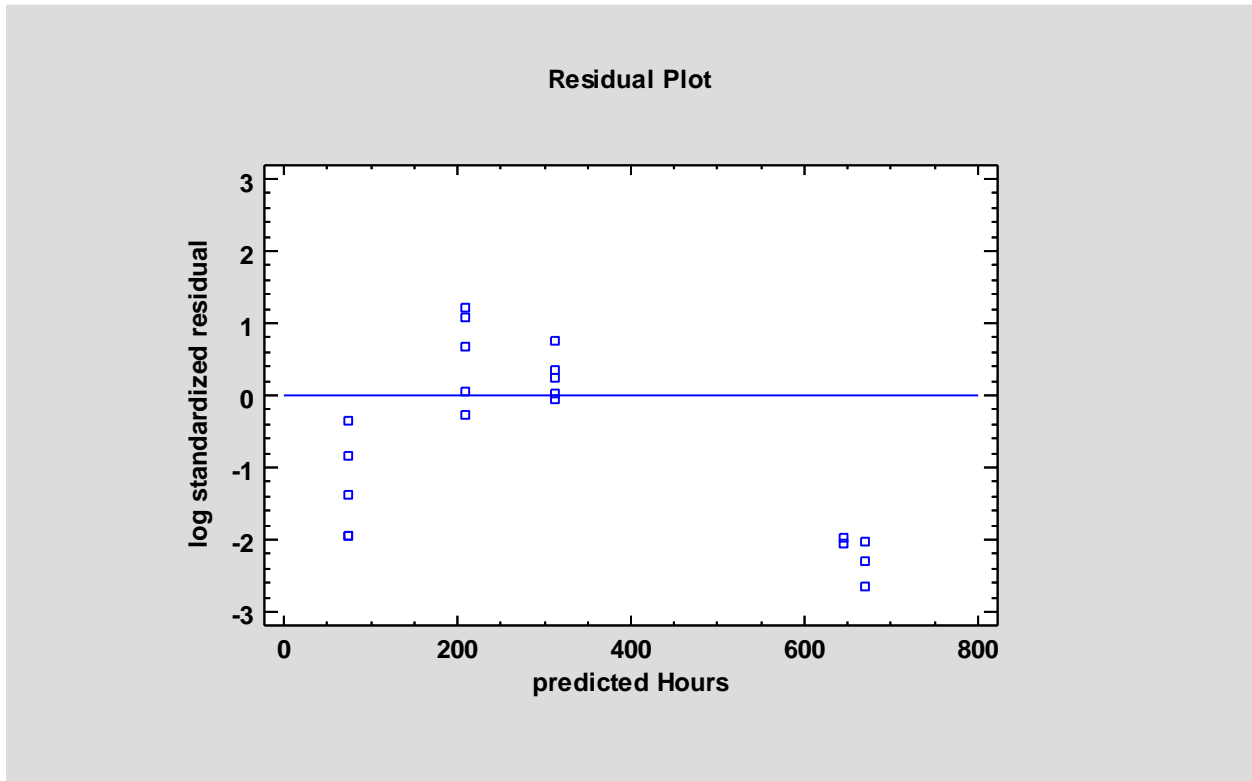
since that would be equivalent to being beyond 3 standard deviations if the distribution was Gaussian.

## Residual Plots

Several other types of residual plots can be created:

### Scatterplot versus Predicted Value

This plot is helpful in visualizing whether the variability is constant or varies according to the magnitude of Y.

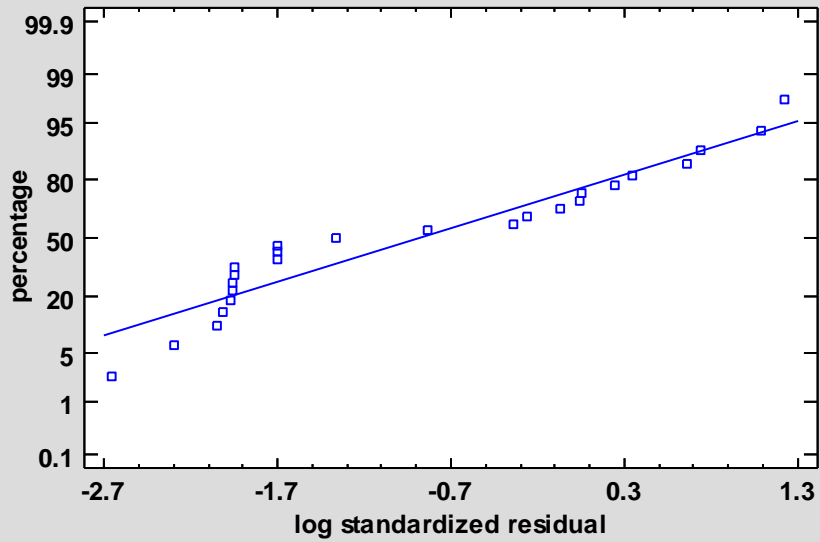


### Normal Probability Plot

This plot can be used to determine whether or not the deviations around the line follow a normal distribution.



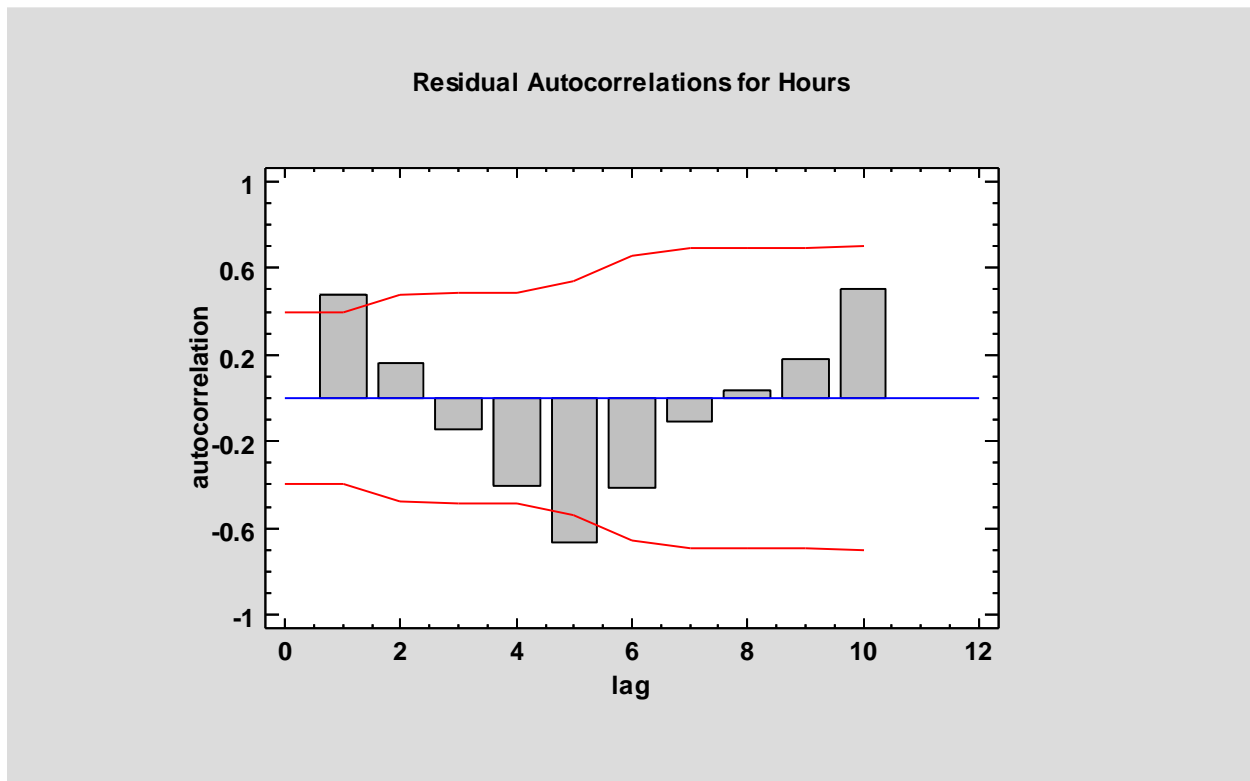
Normal Probability Plot for Hours



Although this plot is created in all regression procedures, the special *Residual Probability Plot* described earlier is more useful for life data residuals.

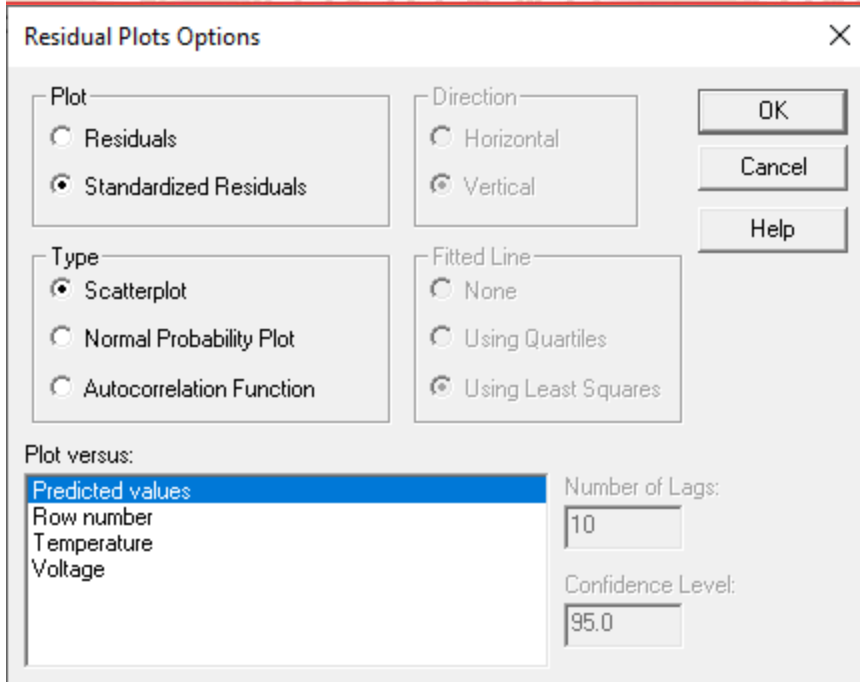
### Residual Autocorrelations

This plot calculates the autocorrelation between residuals as a function of the number of rows between them in the datasheet.



It is only relevant if the data have been collected sequentially. Any bars extending beyond the probability limits would indicate significant dependence between residuals separated by the indicated “lag”.

### *Pane Options*



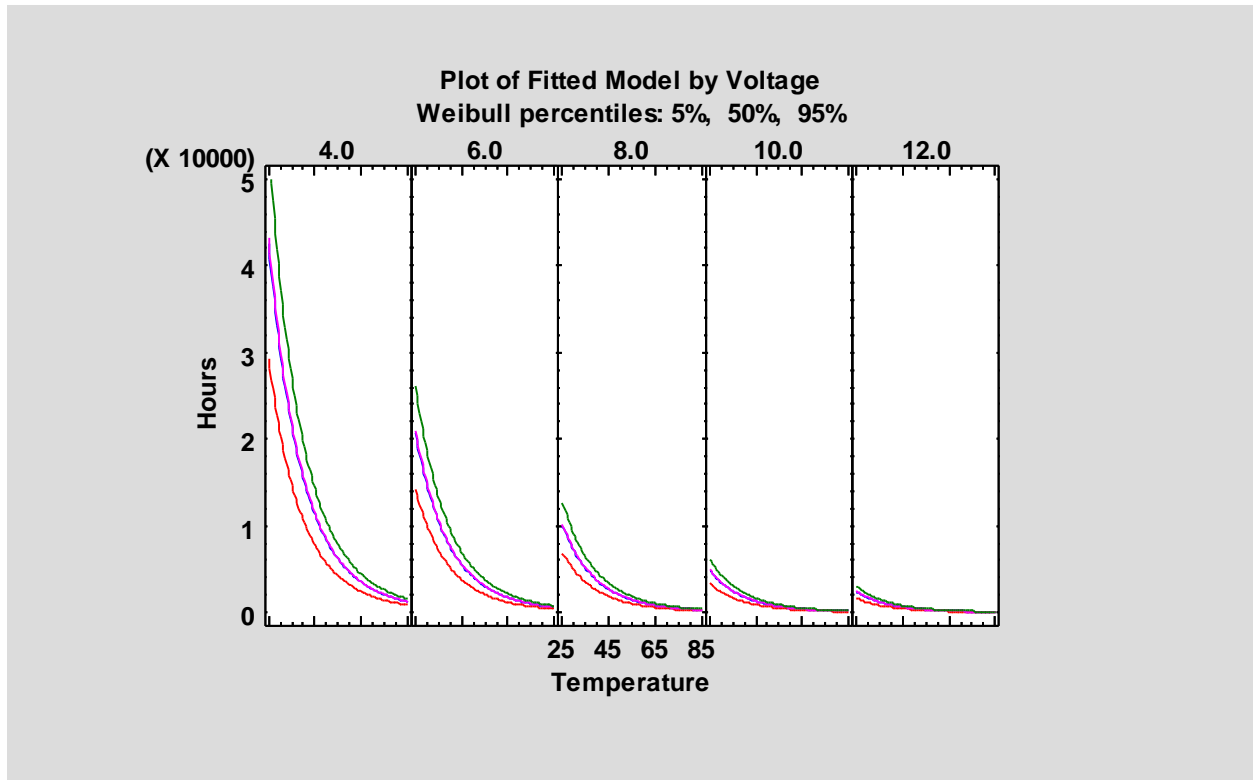
The dialog box 'Residual Plots Options' contains the following settings:

- Plot:**  Residuals,  Standardized Residuals
- Direction:**  Horizontal,  Vertical
- Type:**  Scatterplot,  Normal Probability Plot,  Autocorrelation Function
- Fitted Line:**  None,  Using Quartiles,  Using Least Squares
- Plot versus:** Predicted values (selected), Row number, Temperature, Voltage
- Number of Lags:** 10
- Confidence Level:** 95.0

- **Plot:** the type of residuals to plot.
- **Type:** the type of plot to be created. A *Scatterplot* is used to test for curvature. A *Normal Probability Plot* is used to determine whether the model residuals come from a normal distribution. An *Autocorrelation Function* is used to test for dependence between consecutive residuals.
- **Plot Versus:** for a *Scatterplot*, the quantity to plot on the horizontal axis.
- **Number of Lags:** for an *Autocorrelation Function*, the maximum number of lags. For small data sets, the number of lags plotted may be less than this value.
- **Confidence Level:** for an *Autocorrelation Function*, the level used to create the probability limits.

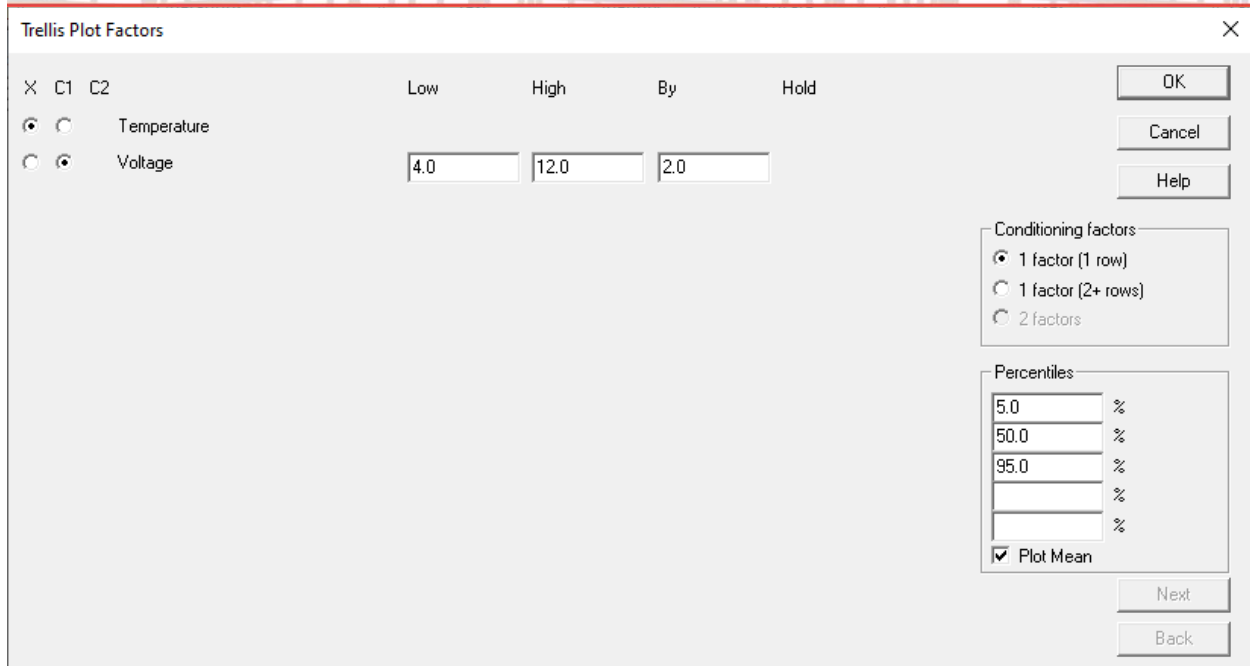
## Trellis Plot

For models containing 2 or more independent variables, a trellis plot may be used to display the estimated percentiles with respect to either 2 or 3 variables. The plot below shows how the percentiles vary by voltage and temperature.



Note that the lowest level of voltage corresponds to the highest percentiles.

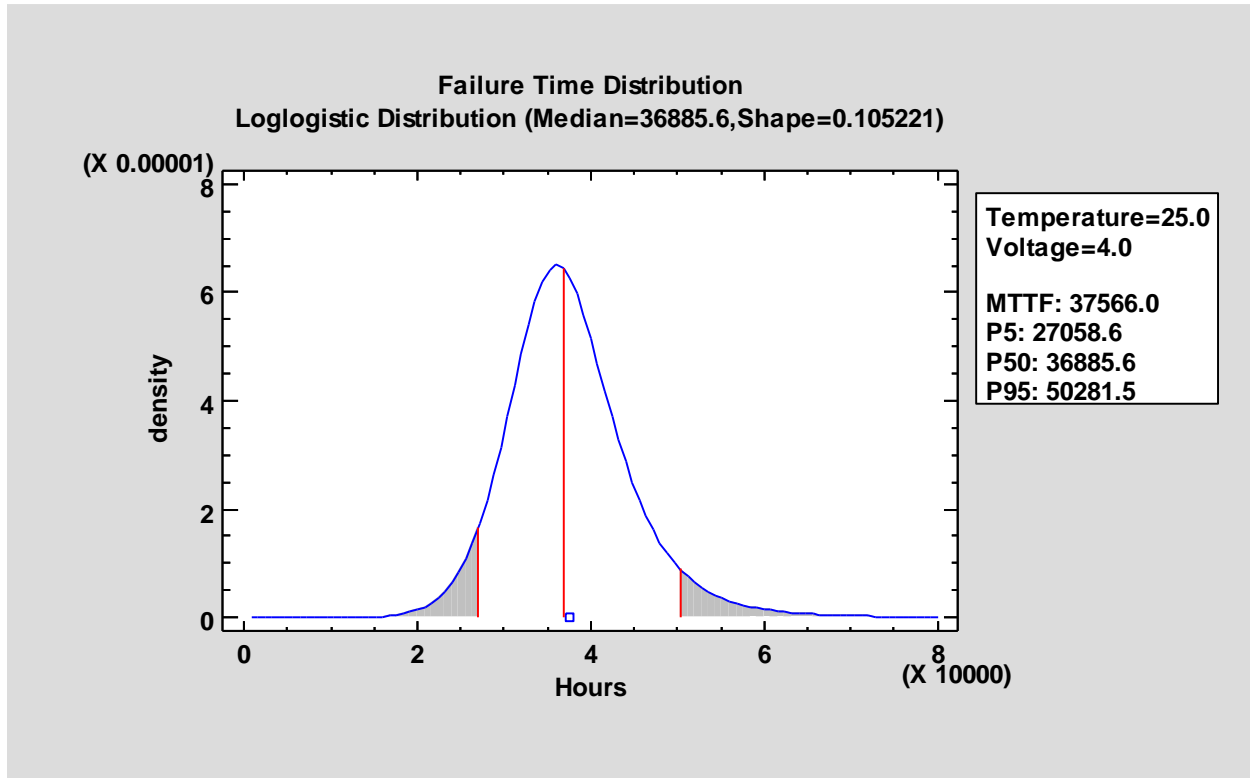
### *Pane Options*



- **X:** the factor used to define the multiple functions within each cell of the plot.
- **C1:** a conditioning factor whose values define the columns of the trellis plot.
- **C2:** a conditioning factor whose values define the row of the trellis plot.
- **Conditioning factors:** the number of conditioning factors in the plot. If only 1 factor is selected, multiple rows may be used to display each level if desired.
- **Low:** for a quantitative conditioning factor, the low end of the range for varying the factor.
- **High:** for a quantitative conditioning factor, the high end of the range for varying the factor.
- **By:** for a quantitative conditioning factor, the increment between the factor levels.
- **Hold:** for factors not selected, the value at which the factor should be held constant.
- **Percentiles:** percentages of the desired percentiles.
- **Plot Mean:** include a line at the estimated mean failure time.
- **Next** and **Back:** used to display other factors when more than 16 are present.

## Failure Time Distribution

The distribution of failure times may be plotted at the normal operating conditions or any other combination of the factors. A typical plot is shown below:



The type of distribution and estimated distribution parameters are shown at the top of the plot. The estimated mean time to failure and selected percentiles are displayed at the right.

### *Pane Options*

Factor Level Options

Factor	Level	Lower percentile:	
Temperature	25.0	5.0	%
Voltage	4.0	Upper percentile:	
		95.0	%
		<input checked="" type="checkbox"/> Plot mean	
		<input checked="" type="checkbox"/> Plot median	
		Shading	
		<input type="radio"/> No shading	
		<input type="radio"/> Shade below lower percentile	
		<input type="radio"/> Shade above upper percentiles	
		<input type="radio"/> Shade between percentiles	
		<input checked="" type="radio"/> Shade beyond percentiles	
		Resolution:	
			101

Buttons: OK, Cancel, Help, Next, Back

- **Level:** value of the predictor variables at which the distribution is displayed.
- **Lower percentile:** lower percentile for inclusion on the plot (if any).
- **Upper percentile:** upper percentile for inclusion on the plot (if any).
- **Plot mean:** whether the MTTF should be displayed.
- **Plot median:** whether the median failure time should be displayed (labeled P50).
- **Shading:** whether and how shading should be added to the plot.
- **Resolution:** number of locations at which to plot the density function.

## Save Results

The following results may be saved to the datasheet:

1. *Fitted Values* – the fitted values  $\hat{\mu}_i$  corresponding to each of the  $n$  observations.
2. *Standard Errors* – the standard errors for the  $n$  fitted values.
3. *Lower Limits for Fitted Values* – the lower confidence limits for the fitted values.
4. *Upper Limits for Fitted Values* – the upper confidence limits for the fitted values.
5. *Residuals* – the  $n$  residuals  $r_i$ .
6. *Standardized Residuals* – the  $n$  standardized residuals  $e_i$ .
7. *Cox-Snell Residuals* - the  $n$  Cox-Snell residuals  $\hat{u}_i$ .
8. *Coefficients* – the estimated model coefficients.
9. *Percentages* – the percentages at which percentiles were calculated.
10. *Percentiles* – the estimated percentiles.
11. *Std. Error of Percentiles* – the standard errors of the estimated percentiles.
12. *Lower Percentile Conf. Limits* – lower confidence limits for the percentiles.
13. *Upper Percentile Conf. Limits* – upper confidence limits for the percentiles.



## Calculations

### Standardized Distributions

Logistic, loglogistic:  $\Phi(z) = \exp(z) / [1 + \exp(z)]$  (13)

Normal, lognormal:  $\Phi(z) = \int_{-\infty}^z (1/\sqrt{2\pi}) \exp(-z^2/2)$  (14)

Smallest extreme value, Weibull, exponential:  $\Phi(z) = 1 - \exp[-\exp(z)]$  (15)

### Likelihood Functions

Let  $\delta_i = 1$  for an exact failure time and 0 for a right-censored observation.

Location-Scale Models:  $L(\beta, \sigma) = \prod_{i=1}^n \left[ \frac{1}{\sigma} \phi\left(\frac{y_i - \mu_i}{\sigma}\right) \right]^{\delta_i} \left[ 1 - \Phi\left(\frac{y_i - \mu_i}{\sigma}\right) \right]^{1-\delta_i}$  (16)

Log-Location-Scale Models:  $L(\beta, \sigma) = \prod_{i=1}^n \left[ \frac{1}{\sigma} \phi\left(\frac{\log(y_i) - \mu_i}{\sigma}\right) \right]^{\delta_i} \left[ 1 - \Phi\left(\frac{\log(y_i) - \mu_i}{\sigma}\right) \right]^{1-\delta_i}$  (17)

### Standard Errors for Coefficients

Determined from the partial derivatives evaluated at the maximum likelihood estimates. Confidence intervals are based on a large-sample normal approximation.

### Mean Failure Times

Distribution	$E(Y)$
Normal	$\mu$
Lognormal	$\exp(\mu + \sigma^2/2)$
Logistic	$\mu$
Loglogistic	$\exp(\mu)\Gamma(1+\sigma)\Gamma(1-\sigma)$
Smallest extreme value	$\mu-0.5772\sigma$
Weibull	$\exp(\mu)\Gamma(1+\sigma)$
Exponential	$\exp(\mu)$