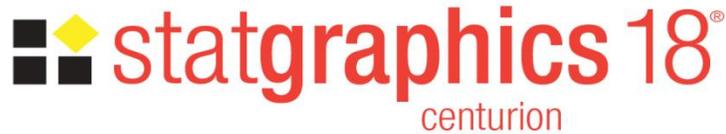


Attribute Capability Analysis Statlet



Revised: 10/10/2017



Summary	1
Analysis Window	2
Classical Analysis	3
Bayesian Analysis	5
Sample Size Determination.....	6
Nonconformities Per Unit.....	7

Summary

This Statlet performs a capability analysis based on attribute data. The data may consist of either the number of nonconforming items in a sample or the total number of nonconformities if an item can have more than one nonconformity. The analysis is based on either the binomial or the Poisson distribution.

The Statlet will calculate:

1. Parameter estimates and confidence limits or upper confidence bounds.
2. Capability indices (at both the best estimate and the upper bound).
3. DPM (defects per million).

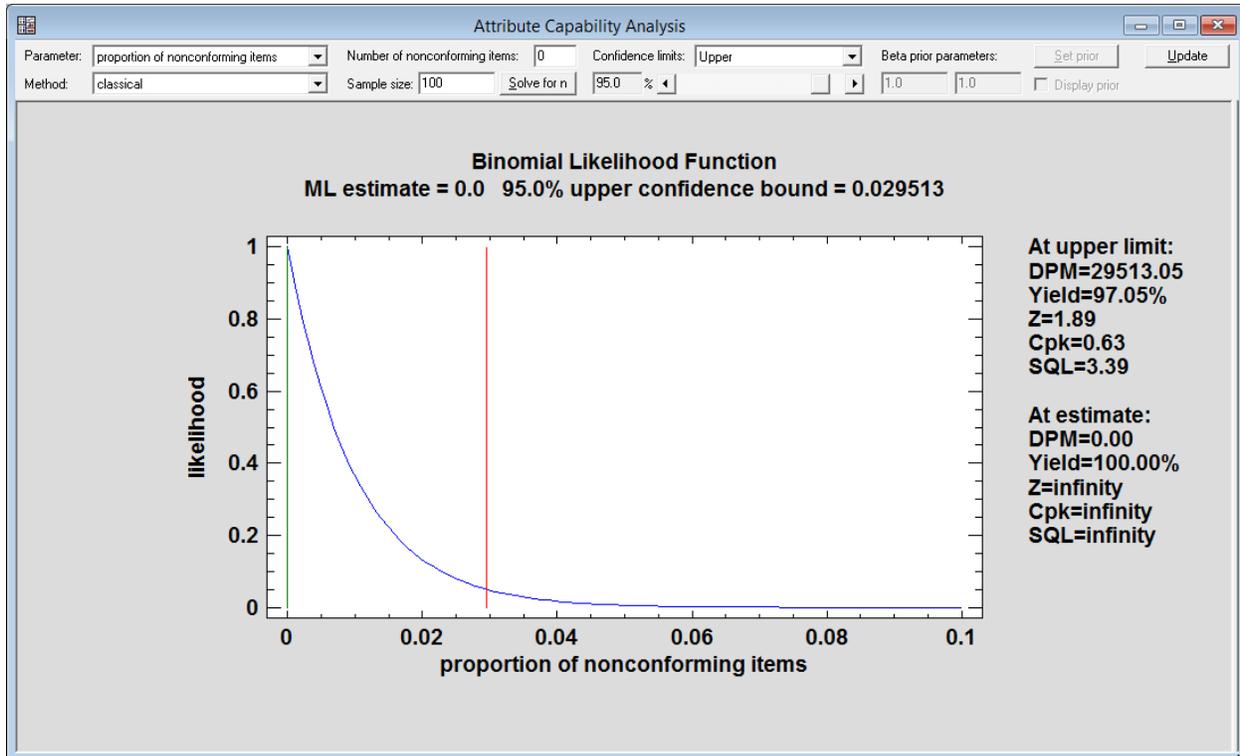
The analysis may be based on either a classical or Bayesian approach.

Sample StatFolio

The file *attribute_capability.sgp* contains a typical example.

Analysis Window

When *Attribute Capability Analysis* is selected from the main Statgraphics Centurion window, an analysis window similar to that shown below will appear:



The analysis displayed in the window is based on the settings selected in the Statlet toolbar:

- **Parameter:** Select either *proportion of nonconforming items* if each item in your sample is classified as either conforming or nonconforming, or *mean rate of nonconformities* if items may have more than one nonconformity.
- **Method:** Select either *classical* if the analysis is to be based only on the likelihood function or *Bayesian* if prior information about the unknown parameter will be supplied.
- **Number of nonconforming items** or **Number of nonconformities:** If each item is classified as either conforming or nonconforming, enter the number of nonconforming items X out of the total number of items n . If items may have more than 1 nonconformity, enter the total number of nonconformities X summed across all of the n items.
- **Sample size:** Enter n , the total number of items inspected.
- **Confidence limits:** Select either *Two-sided* to calculate confidence intervals for the parameter or *Upper* to calculate an upper confidence bound. Use the slider to set the level of confidence for the interval or bound.

- **Beta prior parameters** or **Gamma prior parameters**: If performing a Bayesian analysis, enter the parameters to be used for the prior distribution.

Classical Analysis

As a first example, suppose that $X = 1$ nonconforming item was found in a sample of $n = 800$ items. To analyze the data using a classical approach, you would set the fields to:

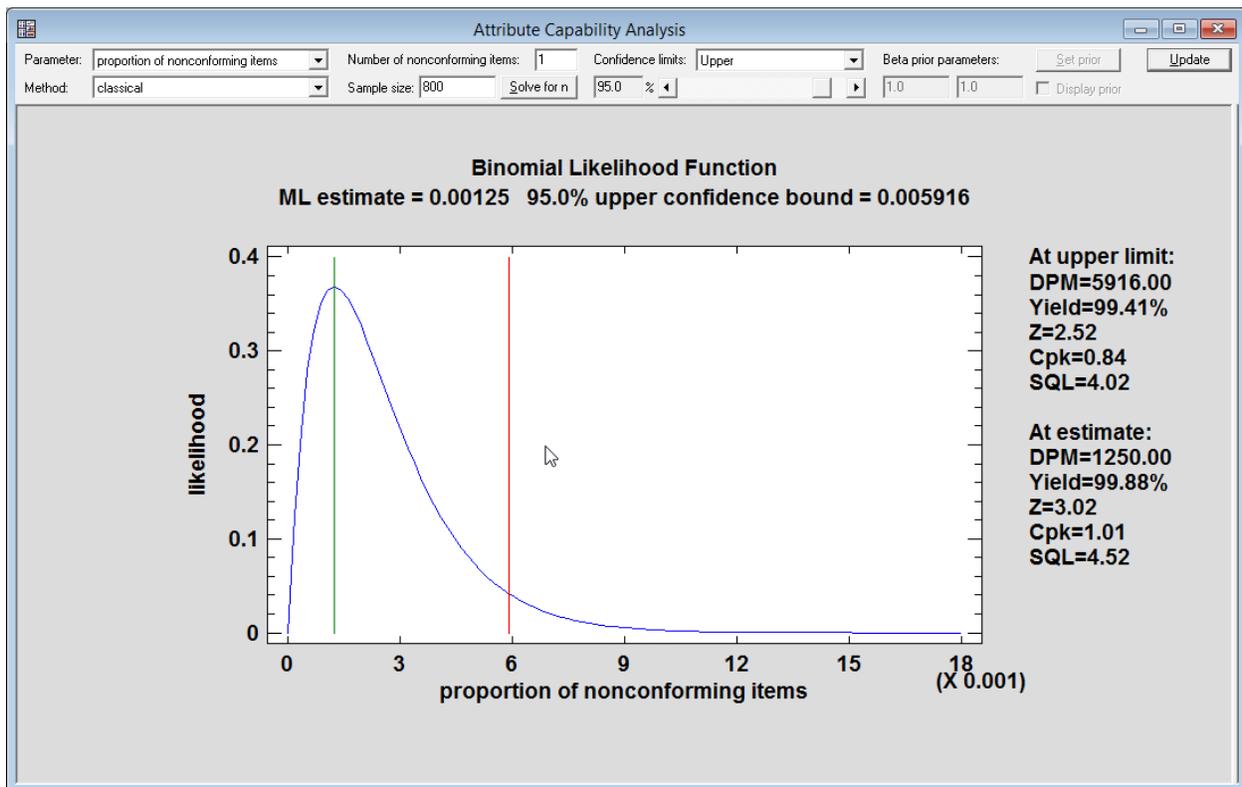
Parameter: proportion of nonconforming items

Method: classical

Number of nonconforming items: 1

Sample size: 800

To estimate the capability of the process using both the maximum likelihood estimate and the 95% upper confidence bound, set the window as shown below:



The analysis begins by obtaining the maximum likelihood estimate of θ , the parameter of the binomial distribution

$$p(x) = \binom{n}{x} \theta^x (1 - \theta)^{n-x} \quad (1)$$

The estimate is given by

$$\hat{\theta} = \frac{X}{n} \quad (2)$$

which in the current example equals 0.00125 and is shown by the green vertical line. The upper confidence bound for the proportion of nonconforming items, shown as a red vertical line, is calculated from

$$\hat{\theta}_U = \frac{v_3 F_{\alpha, v_3, v_4}}{v_4 + v_3 F_{\alpha, v_3, v_4}} \quad (3)$$

where F is the value of the cumulative distribution function for Snedecor's F distribution with the following degrees of freedom:

$$v_3 = 2(X + 1), \quad v_4 = 2(n - X) \quad (4)$$

The blue curve is the likelihood function given the input data.

The margin of the graph displays capability indices and other statistics, calculated at both the upper confidence bound for θ and at the maximum likelihood estimate. The statistics include:

$$\text{DPM (defects per million)} = 1000000 \hat{\theta} \quad (5)$$

$$\% \text{ yield} = 100 (1 - \hat{\theta})\% \quad (6)$$

It also displays an equivalent Z-score which indicates the value of a standard normal distribution which is exceeded with probability equal to $\hat{\theta}$. Specifically, if $\Phi(z)$ is the cumulative standard normal distribution (tabulated in a Z-table in most statistics textbooks), then the Z-score is the value of Z that satisfies

$$\Phi(z) = 1 - \hat{\theta} \quad (7)$$

Given Z, it is then possible to compute both C_{pk}

$$C_{pk} = \frac{Z}{3} \quad (8)$$

and the sigma quality level (SQL) given by

$$SQL = Z + 1.5 \quad (9)$$

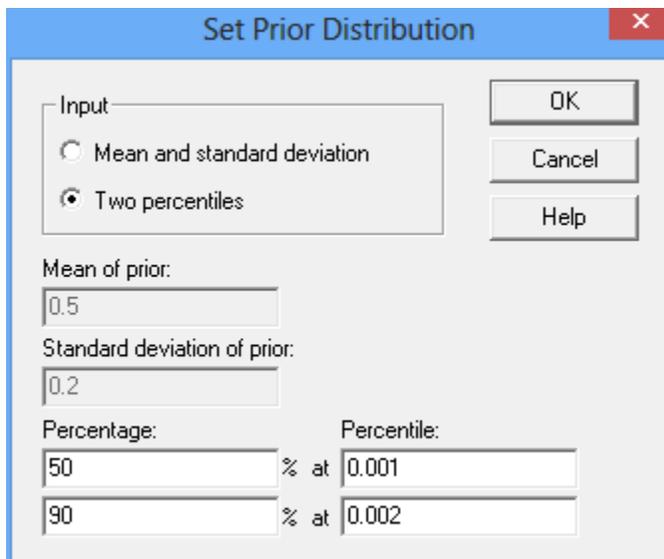
Based on the upper confidence bound in the example above, we can state with 95% confidence that the process has a C_{pk} value of no less than 0.84.

Bayesian Analysis

If prior information exists about the most likely values of θ , this information can be combined with the data using a Bayesian approach. The Bayesian approach begins by specifying a prior distribution for θ . In the case of a binomial distribution, the most common prior distribution to use is the beta distribution, which is defined over the range $0 \leq \theta \leq 1$. The beta distribution has the form

$$p_0(\theta) = \frac{\Gamma(v+w)}{\Gamma(v)\Gamma(w)} \theta^{v-1} (1-\theta)^{w-1} \quad (10)$$

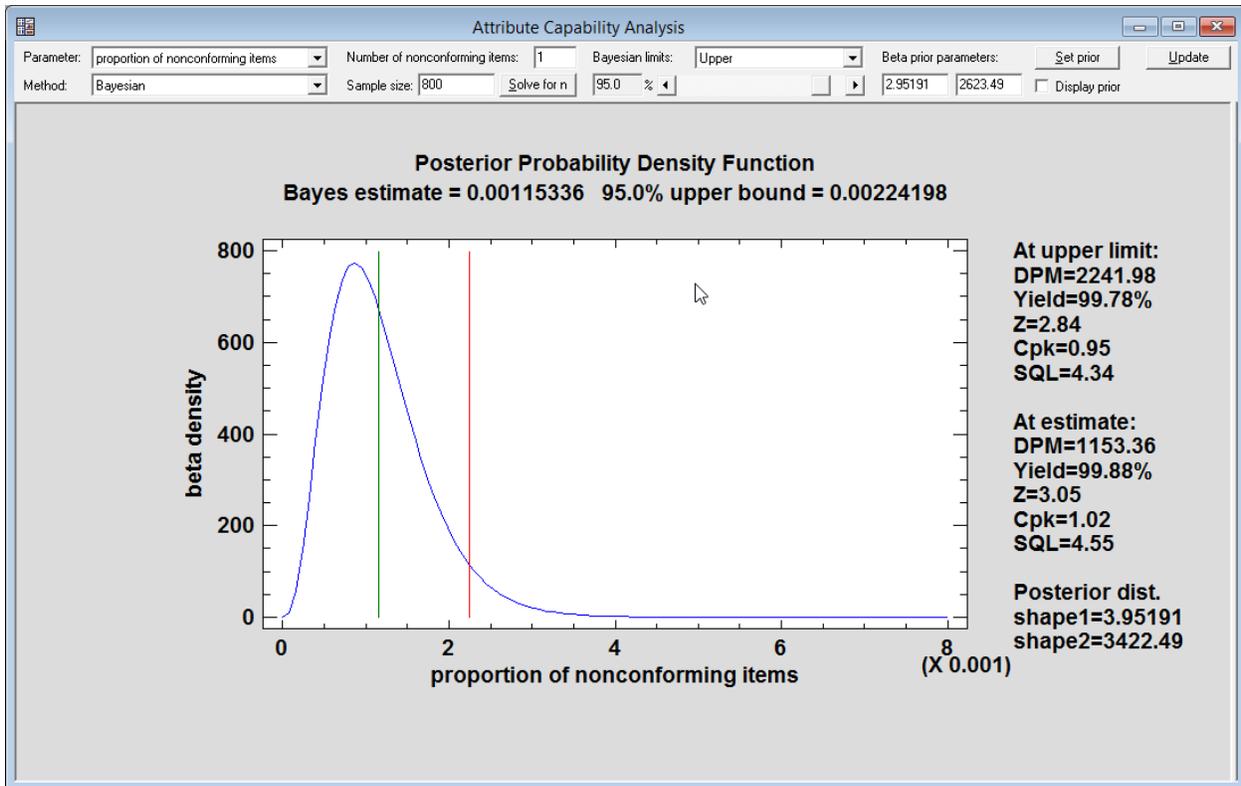
with parameters v and w . To help obtain a reasonable prior, the Statlet has a button labeled *Set prior* which displays the following dialog box:



You can set the prior in either of 2 ways:

1. By specifying the mean and standard deviation of the distribution of your prior knowledge about θ .
2. By specifying 2 percentiles of that distribution.

The above dialog box indicates that the analyst is 50% sure that $\theta \leq 0.001$ and 90% sure that $\theta \leq 0.002$. When *OK* is pressed, the main window will update using a beta prior with the indicated percentiles:



The 95% lower bound on C_{pk} is now 0.95, calculated from the 95th percentile of the posterior beta distribution (displayed as a blue curve) which combines the prior distribution with the likelihood function given the data.

Sample Size Determination

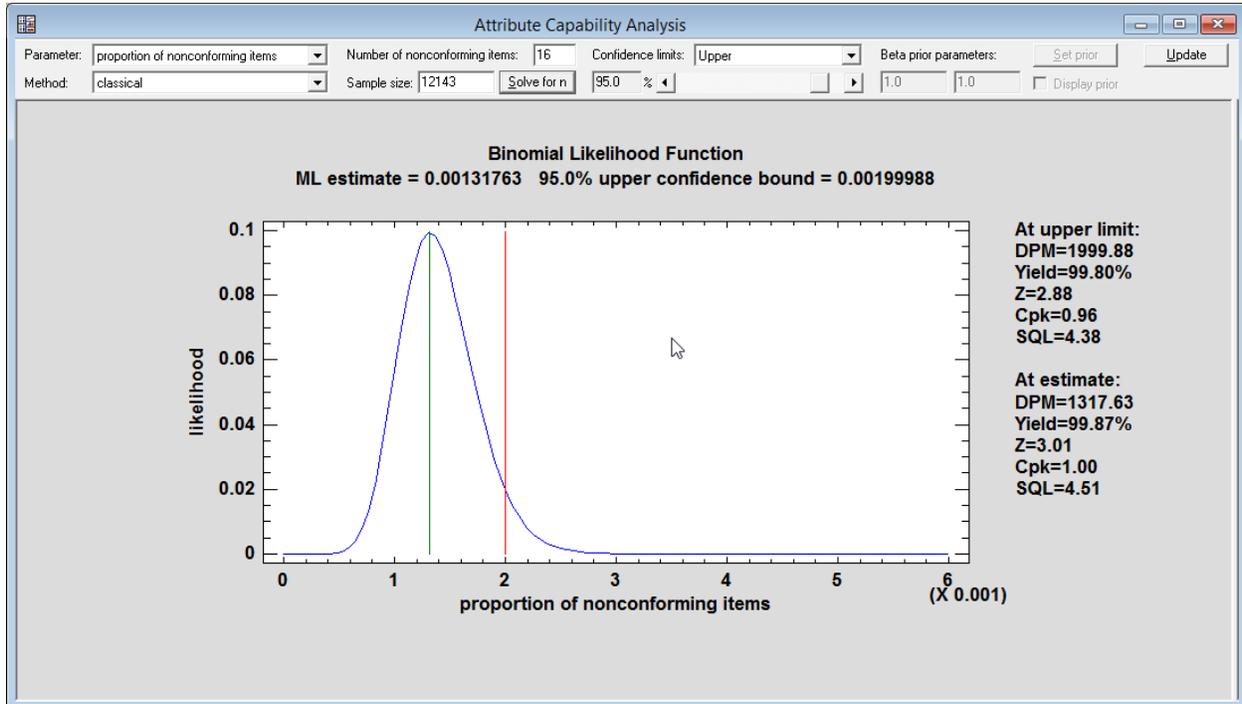
The Statlet can also determine what sample size is needed to obtain a specific upper bound for θ . If the *Solve for n* button is pressed, the following dialog box will appear:

You specify:

1. A target upper bound for θ .

- The quantity that is assumed to remain fixed as the sample size changes, either X or X/n .

When *OK* is pressed, the program will determine the smallest sample size n that yields an upper bound for θ that is no greater than that requested:



Using the classical approach, the required $n = 12,143$.

Nonconformities Per Unit

If the data consist of the total number of nonconformities amongst the n items, the analysis is based on the Poisson distribution given by

$$p(x) = \frac{(\lambda n)^x e^{-\lambda n}}{x!}, \quad x = 0, 1, 2, \dots \quad (11)$$

The fundamental parameter of interest is λ , the mean number of nonconformities per item. The maximum likelihood estimate is

$$\hat{\lambda} = \frac{X}{n} \quad (12)$$

The upper confidence bound is based on the chi-square distribution:

$$\hat{\lambda}_U = \frac{\chi_{\alpha, 2(X+1)}^2}{2n} \quad (13)$$

The conjugate prior is a gamma distribution given by

$$p_0(\lambda) = \frac{\beta^\alpha \lambda^{\alpha-1} e^{-\beta\lambda}}{\Gamma(\alpha)}, \quad 0 \leq \lambda \quad (14)$$

The use of the Statlet parallels that for the proportion of nonconforming items.