Sample Size Determination Using Statgraphics Centurion

Neil W. Polhemus, CTO, StatPoint Technologies, Inc.

Web site: www.statgraphics.com

STATPOINT TECHNOLOGIES, INC.

Preliminaries

One of the most important questions asked of statisticians is:

"How much data do I need?"

In general:

1. More data is better than less data (from a statistical perspective).
2. On the other hand, more data costs more money and takes longer to collect.
3. The information gained by each additional data value decreases as the sample size increases.
4. All data are not created equal.

2

STATPOINT TECHNOLOGIES, INC.

Statgraphics Procedures

Sample size determination procedures are available for:

1. Estimating or testing a single population parameter
2. Comparing estimates from 2 or more samples
3. Estimating capability indices
4. Constructing control charts
5. Developing statistical tolerance limits
6. Creating a screening experiment
7. Developing an acceptance sampling plan

3

STATPOINT TECHNOLOGIES, INC.

Two Basic Methods for Determining Sample Size

Method #1: based on the parameter to be estimated, find a sample size *n* such that the margin of error for the estimate is acceptably small.

Method #2: based on a statistical test to be performed, find a sample size *n* such that the power of the test against a specific alternative hypothesis is acceptably large.

4

STATPOINT TECHNOLOGIES, INC.

Example #1 – estimating a population proportion using the margin of error method

Suppose we wish to estimate the proportion of voters p who would select Mitt Romney over Barack Obama in a head-to-head election for U.S. president. We could ask that the sample size n be large enough so that the 95% confidence interval for p is no wider than:

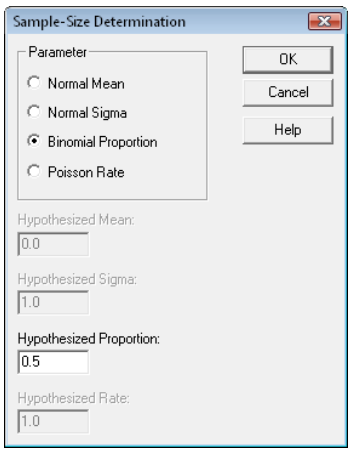
$$\hat{p} \pm 0.03$$

5

STATPOINT TECHNOLOGIES, INC.

Tools – Sample Size Determination – One Sample

- Select the parameter to be estimated.
- Specified the expected value of the parameter.

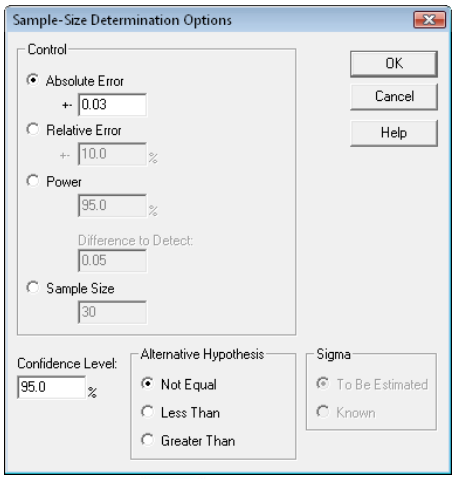


6

STATPOINT TECHNOLOGIES, INC.

Second Dialog Box

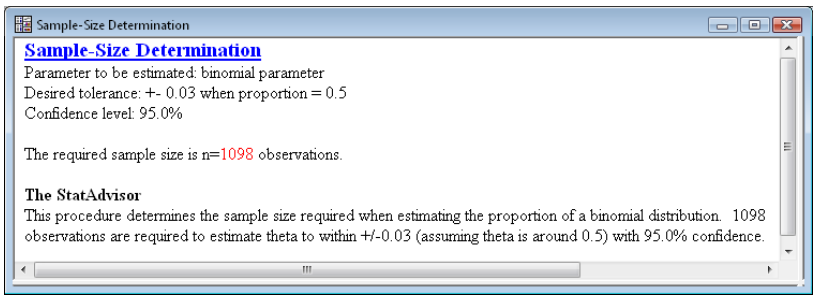
- Specify the desired width of the confidence interval and the level of confidence.
- "Alternative hypothesis" controls whether a two-sided confidence interval or a one-sided confidence bound will be constructed.



STATPOINT TECHNOLOGIES, INC.

7

Solution:



8

Example #2 – comparing two means using the power method

Suppose we wish to compare a new process to an old process. We intend to obtain n_1 measurements from the new process and n_2 measurements from the old process. We wish to be 90% certain that we will obtain a statistically significant result if the mean of the new process is at least 20% larger than the mean of the old process.

Null hypothesis: $\mu_{new} \leq \mu_{old}$

Alternative hypothesis: $\mu_{new} > \mu_{old}$

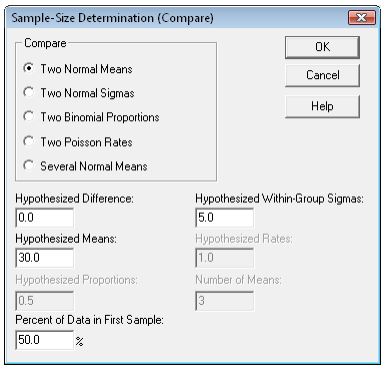
We desire large enough samples such that the probability of rejecting the null hypothesis is at least 0.9 when $\mu_{new} / \mu_{old} = 1.2$.

9

STATPOINT TECHNOLOGIES, INC.

Tools – Sample Size Determination – Two or More Samples

- Specify the parameters to be compared.
- Specify the difference indicated by the null hypothesis.
- Enter assumed means and common with-in group standard deviation.
- Enter percent of data in the first sample (usually 50%).

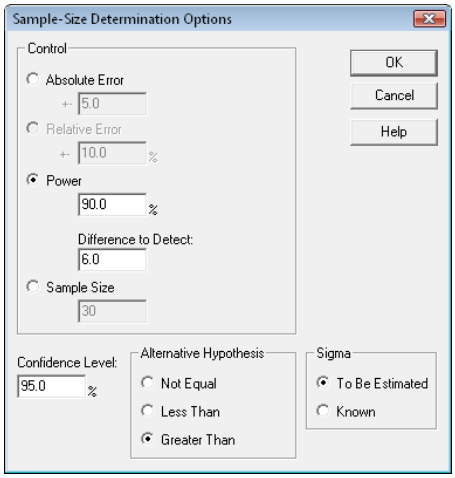


10

STATPOINT TECHNOLOGIES, INC.

Second Dialog Box

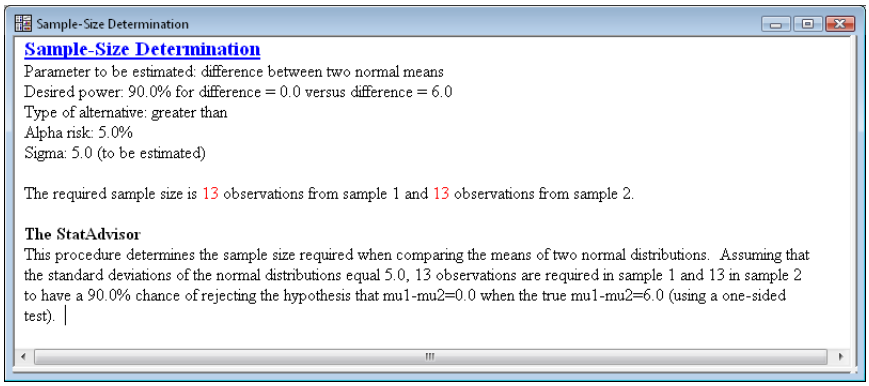
- Specify power (probability of rejecting the null hypothesis).
- Specify the difference to detect (alternative hypothesis).
- Specify confidence level as 1-α.
- Select a two-sided or one-sided test.
- Indicate whether sigma is known (will use z instead of t).



STATPOINT TECHNOLOGIES, INC.

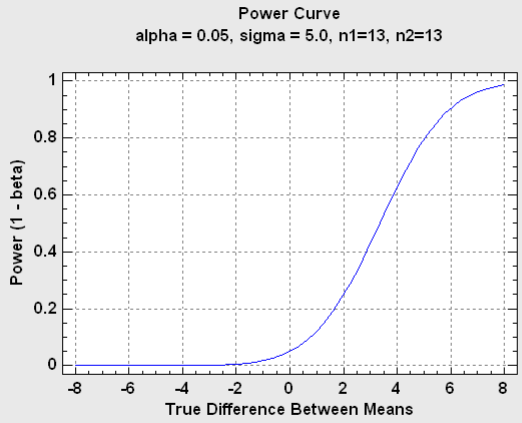
11

Solution:



12

Power Curve



13

Example #3: Capability Indices

A capability index measures how "capable" a process is of meeting established specification limits.

A popular index is

$$C_{pk} = \min\left(\frac{\mu - LSL}{3\sigma}, \frac{USL - \mu}{3\sigma}\right)$$

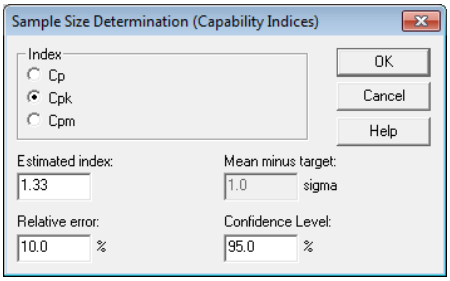
Often, a process is deemed "capable" if $C_{pk} \geq 1.33$

14

STATPOINT TECHNOLOGIES, INC.

Tools – Sample Size Determination – Capability Indices

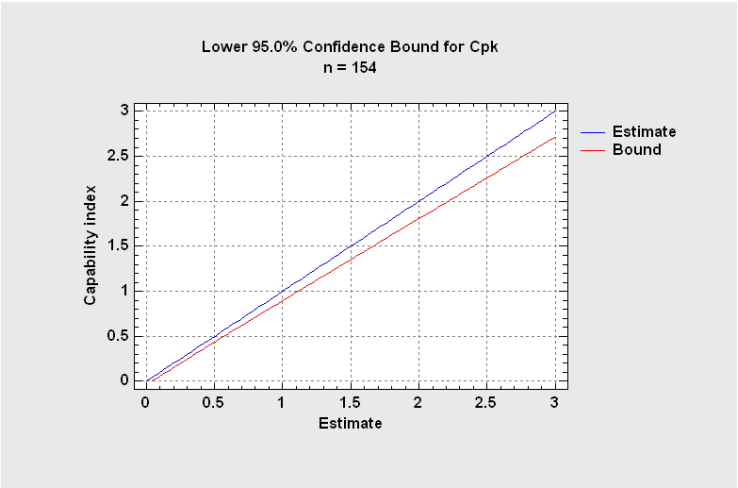
- Select the index to be estimated.
- Specify the expected value of the index.
- Indicate the desired relative error (distance from estimated C_{pk} to its lower confidence bound).



15

STATPOINT TECHNOLOGIES, INC.

Solution:



16

Example #4: Control Charts

For control charts based on grouped data (such as taking *k* samples from repeated batches and constructing X-bar and R charts), determines a suitable value for the group size *k*.

Null hypothesis: process mean = target value

Alt. hypothesis: process mean = target value + Δ

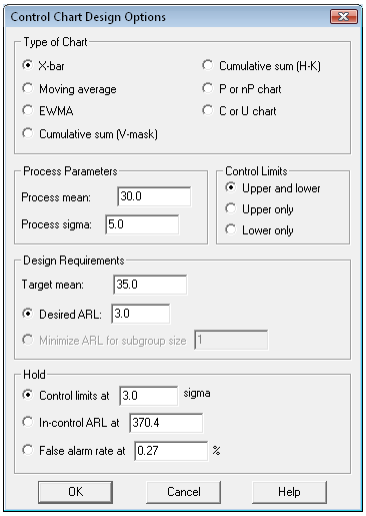
The value of *k* is based on the average run length (ARL), which is the average number of batches that will need to be inspected before getting an out-of-control signal if the mean shifts from the null hypothesis to the alternative.

17

STATPOINT TECHNOLOGIES, INC.

Tools – Sample Size Determination – Control Charts

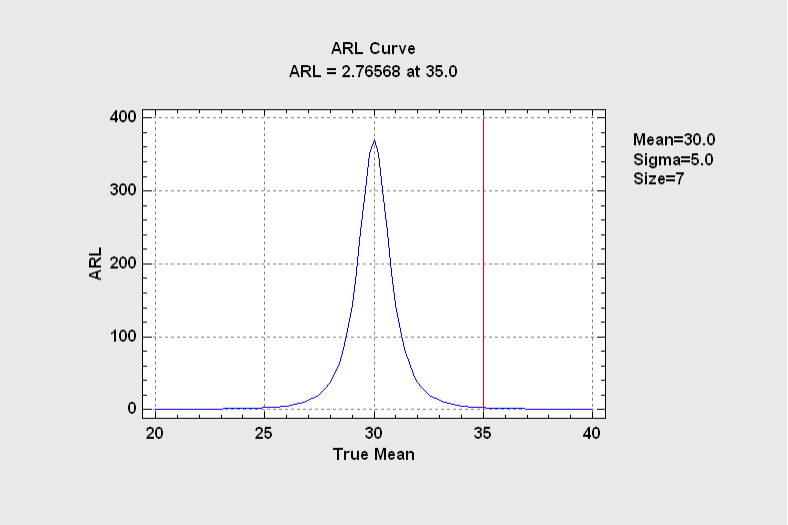
- Select type of chart
- Enter in-control process mean and sigma
- Enter target out-of-control mean
- Enter desired ARL at the target mean



18

STATPOINT TECHNOLOGIES, INC.

ARL Curve for $k = 7$



19

Example #5: Statistical Tolerance Limits

Statistical tolerance limits bound a specified proportion of a population at a specified level of confidence.

Suppose I wish to demonstrate that I am 95% confident that at least 99% of measurements obtained from items I am producing will fall within my specification limits.

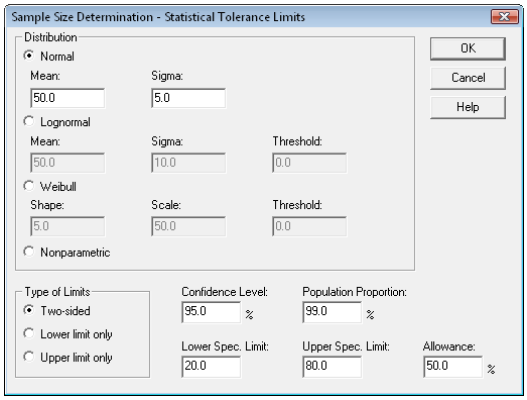
I could ask for a sufficiently large sample size n such that the expected width of my tolerance interval is no more than 50% of the distance between the specification limits.

20

STATPOINT TECHNOLOGIES, INC.

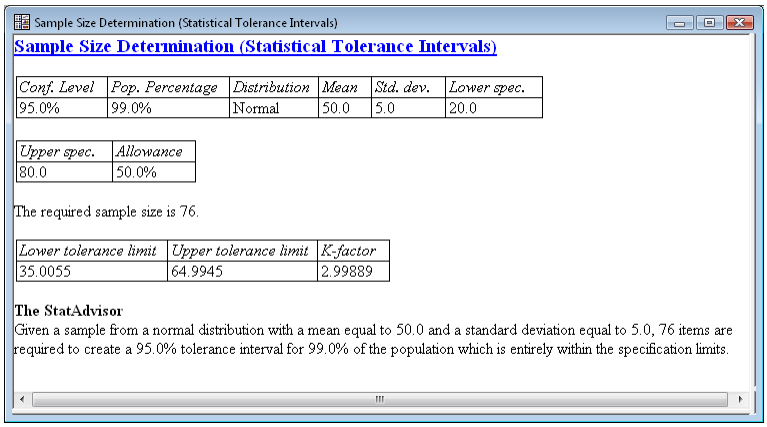
Tools – Sample Size Determination – Statistical Tolerance Limits

- Select assumed distribution
- Specify parameters
- Select type of limits
- Enter confidence level and population proportion
- Enter specification limits
- Enter allowance (percentage of distance between spec limits)

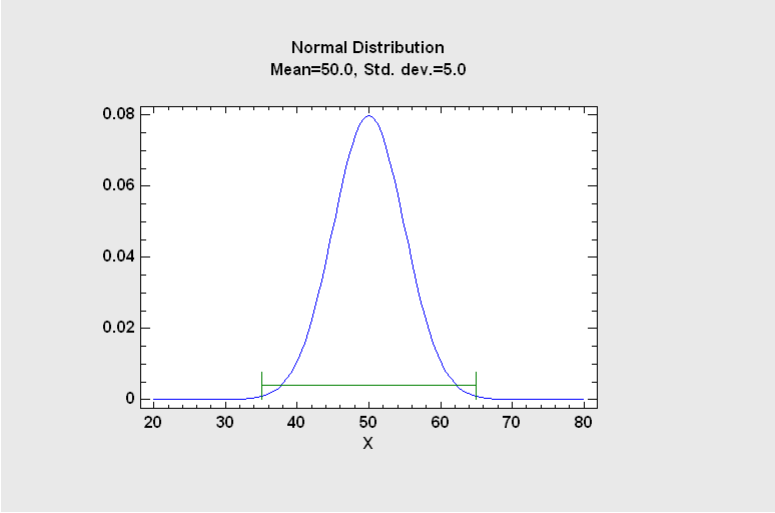


STATPOINT TECHNOLOGIES, INC.

Solution:



Tolerance Plot



23

Example #6 – Screening Designs

A screening design is a statistically designed experiment in which many factors are varied with the goal of identifying the factors that have the biggest impact on a response variable.

Most common screening designs contain:

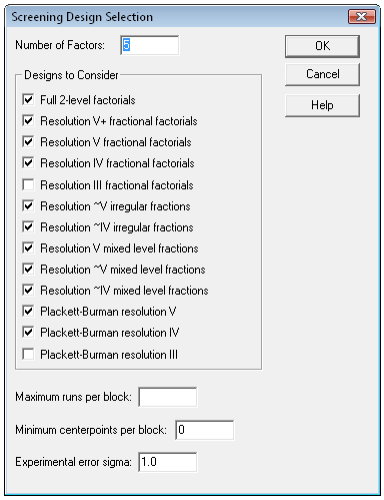
- Experiments at 2 levels of each factor (low and high).
- Additional experiments at a central level of all of the factors.

24

STATPOINT TECHNOLOGIES, INC.

Tools – Sample Size Determination – Screening Designs

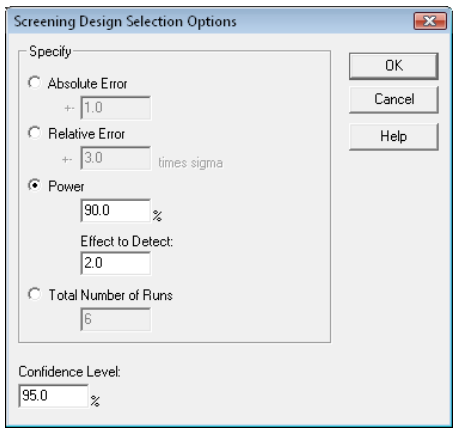
- Select the type of designs to be considered.
- Specify the maximum runs per block.
- Specify the minimum centerpoints per block.
- Enter the assumed sigma for the experimental error.



STATPOINT TECHNOLOGIES, INC.

Second Dialog Box

- Specify either the margin of error for estimated effects or the power of t-tests against an effect of a given magnitude.
- "Effect" is defined as the change in the response as a factor is moved from its low level to its high level.



STATPOINT TECHNOLOGIES, INC.

Output

Output shows selected designs that meet the specified requirements.

Screening Design Selection

Screening Design Selection

Input

Number of Factors	*Min. Centerpoints per Block*	*Max. Runs per Block*	*Exp. Error Sigma*	*Desired Power*	*Target Effect*	*Confidence Level*
5	0		1.0	90.0%	2.0	95.0%

Selected Designs

Design	*Runs*	*Resol.*	*Corner Points*	*Center Points*	*Error D.F.*	*Reps.*	*Blocks*	*Power (%)*
Factorial 2^5	32	V+	32	0	16	1	1	99.863
Half fraction 2^5-1	22	V	16	6	6	1	1	91.4633
Half fraction in 2 blocks	22	IV*	16	6	6	1	2	91.4633
Irregular fraction 2^5*3/8	24	~IV	24	0	13	2	1	99.1553
Mixed level fraction 3*2^4-1	24	~V	16	0	7	1	1	92.7197
Mixed level fraction 3*2^4-2	24	~IV	16	0	13	2	1	95.5635

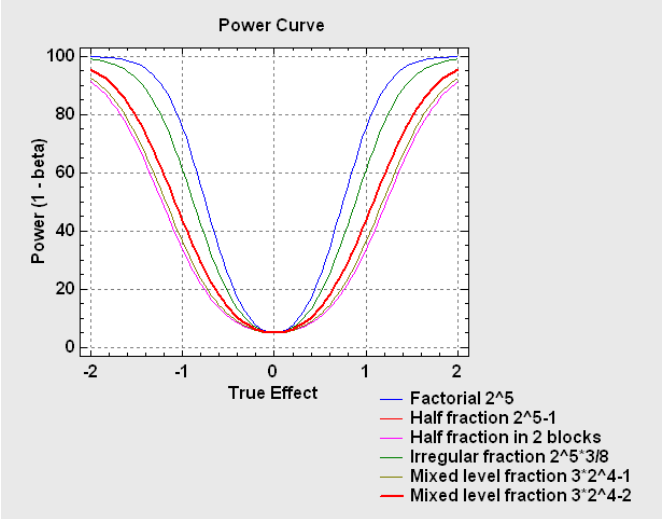
The StatAdvisor

The table shows 6 experimental designs which have at least a 90.0% chance of detecting an effect of magnitude 2.0.

27

STATPOINT TECHNOLOGIES, INC.

Power Curve



28

Example #7 – Acceptance Sampling

Sample *n* items from a lot of size *N*. Based on the sample, either accept or reject the lot.

Case #1 (attributes): Inspect each item and classify it as good or bad. If the number of bad items is less than or equal to *c*, accept the lot.

Case #2 (variables): Measure each item. If the mean of the *n* items is far enough from the nearest specification limit, accept the lot.

29

STATPOINT TECHNOLOGIES, INC.

Formulation

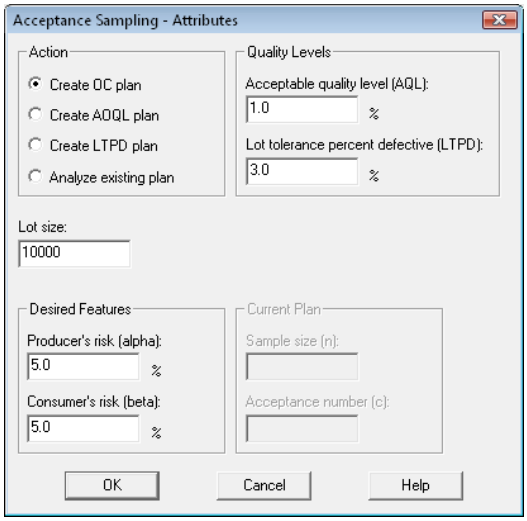
1. Null hypothesis: The lot is good: % defective items ≤ AQL (acceptable quality limit)
2. Alt. hypothesis: The lot is bad: % defective items ≥ LTPD (lot tolerance percent defective)

- Producer's risk α: probability of rejecting a good lot
- Consumer's risk β: probability of accepting a bad lot

30

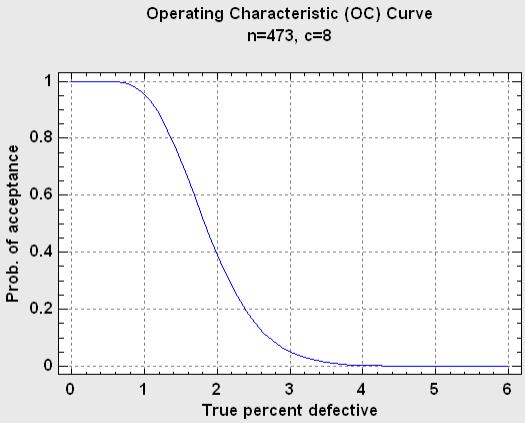
STATPOINT TECHNOLOGIES, INC.

SPC – Acceptance Sampling - Attributes



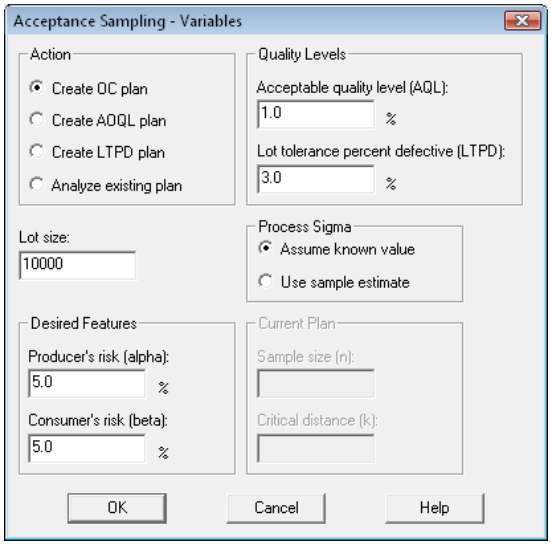
31

Solution – accept lot if after inspecting 473 items, no more than 8 items are defective.



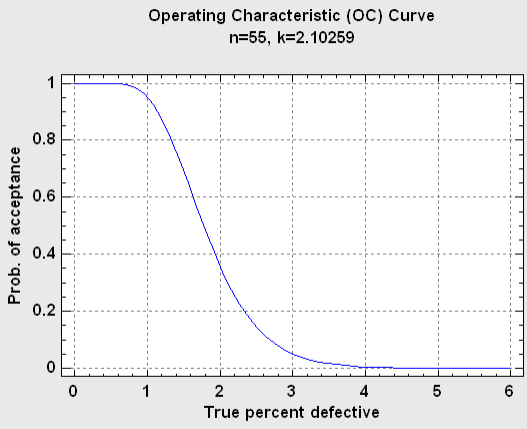
32

SPC – Acceptance Sampling - Variables



33

Solution – accept lot if after measuring 55 items, the sample mean is no closer than 2.10259 standard deviations to the nearest spec limit.



34

More Information

Go to www.statgraphics.com

Or send e-mail to info@statgraphics.com

STATPOINT TECHNOLOGIES, INC.