COGENT
QC SYSTEMS

Sample Examples – The Calculator in Action

Cogent's Statistical Sample Size Calculator allows you to estimate the random sample size required to achieve a specified level of statistical precision, or sampling error, in making a sample inference about a population. There are four required input parameters:

(1) the population count
(2) the estimated incidence rate of the attribute
(3) the desired statistical precision, expressed as a percentage of the population
(4) the desired confidence level.

For definitions of these terms, see the "Cogent Quality Trend Reports Demystified – Infographic".

The following examples show different situations in which the Calculator might be used, and the appropriate ways to use it. We welcome any questions or comments about the Calculator and its application to your situation. Contact Cogent at info@cogentqc.com or use the Calculator Feedback form on the webpage (http://cogentqc.com/tools-resources/statistical-calculator/).

> **PLEASE NOTE:** The Calculator provides an *estimated* sample size based on the input parameters, which include an *estimated* defect rate. The period you choose to achieve the specified precision and confidence may make a significant difference, as described below. In order to confirm that the sample size is adequate, after completing your sample reviews, calculate the confidence interval using the actual sample defect rate. (The confidence interval calculation is beyond the scope of the online Calculator. If you would like assistance with this final step, please contact Cogent.)

Example 1

How many transaction records (e.g., loan originations) do you need to randomly sample *in a single sampling* in order to achieve 2% precision at 95% confidence (one-sided) over one year? (This has become the industry precision standard in mortgage quality control.) Assume that 12,000 loans are originated in the calendar year and your estimated incidence rate (frequently known as 'defect rate') is 5%.

Population Count	12,000
Expected Incidence Rate (%)	5
Desired Precision Level (%)	2
Confidence Level	95 (one-sided)
Sample Size	**311**

311 records must be randomly sampled.[1] The resulting sample defect rate can be inferred to the annual population from which the sample was drawn, with 95% confidence that the defect rate of the population will not exceed the sample defect rate by more than 2%.

This kind of calculation might be used for an annual retrospective audit, which is typical in internal audits, where the entire population has already been established and a random selection of loans can be expected to reflect varying production levels at different periods over the year. More often in the QC world, sampling needs to be prospective and more frequent than annually. This requires a different calculation.

Example 2

How many transaction records do you need to randomly sample *each month* in order to achieve 2% precision at 95% confidence (one-sided) over one year? Assume 1,000 loans are originated in the first month of the year and your estimated incidence rate at the beginning of the year is 5%.

The Calculator is designed to provide a sample size based on the period over which you wish to achieve the desired precision and confidence levels. In this case, that is a full year. But you only have population data for the first month of the year. So first, you estimate a population for the whole year by extrapolating the current month's population. Then, you divide the suggested sample size by 12 in order to derive the current month's sample size.[2]

Population Count	12,000 *[estimate = 12x current month's population]*
Expected Incidence Rate (%)	5
Desired Precision Level (%)	2
Confidence Level	95 (one-sided)
Sample Size	**311** *[annual]*
	26 *[monthly = 311/12]*

You must sample 26 loans every month, all else being equal, in order to achieve 2% precision and 95% confidence over one year. However, all else is not always equal.

Example 3

After 3 months of audits, you discover that your defect rate is closer to 10%, not the 5% you assumed. How many transaction records do you need to randomly sample *each month for the*

[1] Do not introduce bias into your selections. Once you have a good estimate of the number of records to be sampled, make sure that they are sampled randomly.
[2] The statistical sample size calculations in Cogent QC Systems perform this extrapolation and derivation automatically.

remainder of the year in order to achieve 2% precision at 95% confidence (one-sided) over one year? Assume 1,000 loans per month are originated for the remainder of the year.

Population Count	12,000
Expected Incidence Rate (%)	10
Desired Precision Level (%)	2
Confidence Level	95 (one-sided)
Sample Size	**576** *[annual]*

You have already sampled and audited 78 loans based on the earlier estimated 5% defect rate (3 months x 26 loans per month). Therefore, you must sample 498 loans over the next 9 months of the year (576 - 498 = 498), which translates to 55 loans per month.[3]

Example 4

How many transaction records do you need to randomly sample *each month* in order to achieve 2% precision at 95% confidence (one-sided) over *each month*? Assume 1,000 loans are originated in the first month of the year and your estimated incidence rate at the beginning of the year is 5%.

Population Count	1,000
Expected Incidence Rate (%)	5
Desired Precision Level (%)	2
Confidence Level	95 (one-sided)
Sample Size	**242**

You must sample 242 loans each month. Because your desired 2% precision and 95% confidence is required for a single month rather than over 12 months, you must sample many more loans in the current month.

Observations

If you experiment with the Calculator, you will notice the following:

Population has the least impact on sample size. After volume of about 12,000 units, the sample size required remains quite flat. This means that sample size as a percentage of population decreases as population grows. This often seems counter-intuitive: how do you get the same

[3] The statistical sample size calculations in Cogent QC Systems adjust automatically to the variations in defect rate across audit periods.

precision and confidence by sampling 320 loans whether your population is 10,000 or 100,000 or 1,000,000? It's the same phenomenon as in election polling. With a small sample, you can make far-reaching inferences.

Incidence Rate is the factor over which you have most control. The lower the incidence (or defect) rate, the fewer the loans to be sampled and audited. Therefore, improved quality reduces the number of audits required, in addition to producing superior products. Note that your definition of 'defective' has an impact on incidence rate also.

Precision Level has the greatest impact on sample size. The narrower the precision level (e.g., 1% instead of 2%), the more precisely you can infer to the population. But the price you pay is that you must sample many more loans. Similarly, as you shorten the period over which you wish to achieve those precision and confidence levels (e.g., one month vs. one year), the sample size increases significantly.

Confidence Level has the second greatest impact on sample size. In quality control, we are most interested in the upper bounds of the inferred defect rate (the worst case) so a one-sided confidence level (or interval) is often sufficient.