What Do We Mean by "Zero Defects"?

Part 1 of 3: Sampling from a Process

BY STEPHEN N. LUKO AND DEAN V. NEUBAUER

Q: What does the phrase "zero defects" mean, and what can we say about a process for which this phrase has been applied?

A. The phrase "zero defects" denotes a philosophical position with respect to quality whereby zero non-conformances are the standard in quality excellence. This also gives it great psychological appeal. The philosophy took shape in the aerospace industry and in U.S. Department of Defense standards in the 1960s and 1970s. Later, the philosophy migrated to the automotive industry and to its supply base.

Today, this standard is widely used. Unfortunately, many professionals in industry have a mistaken notion about zero defects. If literally zero defects are found in a sample, what can we say about the process that produced the sample? Many people think that a process producing a sample having zero defects must also have a quality level of zero defects and thus can be expected never to produce a defect. This perception is the subject of this and two future articles concerning common variations on this theme.

> **Many people think that a process producing a sample having zero defects must also have a quality level of zero defects and thus can be expected never to produce a defect.**

There are three commonly found cases where zero defects might apply: 1) sampling from a process, or a practically unlimited supply or lot; 2) sampling from a continuum, as when looking for events that can occur in any number in time, space, area, volume, etc.; and 3) sampling a finite lot. These cases are more thoroughly discussed in the ASTM standard E2334, Practice for Setting an Upper Confidence Bound for a Fraction or Number of Non-Conforming Items, or a Rate of Occurrence for Non-Conformities, Using Attribute Data, When There Is a Zero Response in the Sample.

In this article, we examine Case 1, probably the most common form. We will not make a distinction between variable or attribute type data. For the purposes of this discussion, when sampling a process, we have a sample size, n, and compare each object sampled to a requirement. The requirement might be a pure go/no-go (attribute), or it might be comparing the measurement result to a limit (variable). In any event, when all n objects meet the requirement, we have a sample with $x = 0$ non-conformances. We further make the assumption that the process being sampled is in a state of statistical control and that the sample adequately represents the process. The former requirement means that the process is stable so that statistical inference may be used for the prediction of future process output. The latter requirement simply means that the sample is a simple random sample.

The point of sampling a process is typically to gain information about the true fraction non-conforming (or some other attribute fraction), p. We can never infer that $p = 0$ based on a limited sample that happens to turn up $x = 0$ non-conformances. We expect that if $x = 0$ has been observed, p must be small, but not zero. Further, if $x = 0$ has been observed, the larger the sample size, the smaller we would expect the true p to be. We can think of the fraction p

as the probability that a single observation from the process will be non-conforming, where each object has this same probability.

In sampling theory, the mathematical model that governs this type of sampling is the binomial distribution. In the binomial distribution, we require two numbers to know everything that can be known statistically about a sampling result. These are the sample size, *n*, and the true probability that any object will be non-conforming, *p*. We always know our sample size, but one typically does not know *p*, and so we must estimate it. Generally, if $x = r > 0$ non-conformances were observed in a sample of size *n*, one might take the estimate of *p* as the fraction *r*/*n*. This is called a point estimate of *p*. In reality, the true *p* is likely to be somewhat different than this estimate. We do not know this difference, except that when $x = 0$ occurs, the point estimate would then be 0, which, as discussed above, we can never accept.

Still one requires an estimate of *p*. So, what is the largest value of *p* that could have given rise to $x = 0$ in a sample of *n* with some reasonable probability under the assumed conditions of this sampling model? By "reasonable probability" we usually mean something on the order of an error of 5 or 10 percent (0.05 to 0.1). This can be expressed as follows:

$$(1-p)^n \geq 1-C \qquad (1)$$

In Equation 1, *p* is the probability that a single object is non-conforming, and (1 - *p*) is the probability that it is conforming. For *n* independent conforming objects, the joint probability of this result is the product of (1 - *p*) taken *n* times, which is the left-hand side of Equation 1. We require this to be a small but reasonable probability. That is, the right side is shown as "1 - *C*", where *C* is the confidence in our final result. Note than when 1 - *C* = 0.1, *C* = 0.9 or 90 percent confidence. Solving the inequality (Equation 1) for *p* gives:

$$p \leq 1 - \sqrt[n]{1-C} \qquad (2)$$

Thus, Equation 2 is the upper confidence bound for *p* at confidence level *C* when $x = 0$ is observed in a sample of size *n*. For example, if $n = 22$ and $C = 0.9$, the upper bound for *p* is approximately 0.1, or 10 percent, using Equation 2. We may claim 90 percent confidence in stating this result, meaning that there is still a 10 percent chance (risk) that the true *p* could be larger than 10 percent. Should we want to claim 95 percent confidence, the upper bound for *p* would increase slightly to 12.73 percent, but the risk of underestimating *p* shrinks to 5 percent.

Three types of questions may be answered using Equation 1. Knowing any two quantities of *n*, *p* or *C*, we can solve for the other. For example, we might want to know the sample size required to demonstrate that *p* is not more than 2 percent with confidence 90 percent, or what confidence has been demonstrated that *p* is not more than 2 percent using a sample of size 100. Solving Equation 1 for *n* and *C* gives the following formulas:

$$n \geq \frac{\ln(1-C)}{\ln(1-p)} \qquad (3)$$

$$C \geq 1-(1-p)^n \qquad (4)$$

Using Equation 3 to demonstrate that *p* is not more than 2 percent using 90 percent confidence (*C*), requires a sample size of $n = 114$. With $n = 100$, one can claim a confidence of 86.74 percent that *p* is not more than 2 percent, using Equation 4. These useful formulas represent one of the simplest cases of statistical inference. We conclude this article with a table that shows the required sample size (*n*) having 0 non-conforming to demonstrate that the true process fraction non-conforming is not more than p with confidence C.

STEPHEN N. LUKO, *Pratt & Whitney Aircraft, is the immediate past chairman of Committee E11 on Quality and Statistics, and a fellow of ASTM.*

DEAN V. NEUBAUER, *Corning Inc., is vice-chairman of Committee E11 on Quality and Statistics, chairman of Subcommittee E11.30 on Statistical Quality Control and chairman of E11.90.03 on Publications; he also coordinates the DataPoints column.*

	CONFIDENCE, *C*					
p	0.60	0.75	0.80	0.90	0.95	0.99
0.001	916	1,386	1,609	2,302	2,995	4,603
0.005	183	277	322	460	598	919
0.010	92	138	161	230	299	459
0.020	46	69	80	114	149	228
0.030	31	46	53	76	99	152
0.040	23	34	40	57	74	113
0.050	18	28	32	45	59	90
0.060	15	23	27	38	49	75
0.070	13	20	23	32	42	64
0.080	11	17	20	28	36	56
0.090	10	15	18	25	32	49
0.100	9	14	16	22	29	44

Table 1: Relationship Among Sample Size, n, Confidence, and Demonstrated Process Fraction Non-Conforming When Sampling a Process and x = 0 Non-Conformances Appear

Statistics play an important role in the ASTM International standards you write, and a panel of experts is ready to answer your questions about how to use statistical principles in ASTM standards. Please send your questions to SN Editor in Chief Maryann Gorman at mgorman@astm.org or ASTM International, 100 Barr Harbor Drive, P.O. Box C700, West Conshohocken, PA 19428-2959.