Why 30?

A Consideration for Standard Deviation

By Alex T.C. Lau

Q What is the minimum number of data points recommended to estimate standard deviation?

A This is one of the most common questions posed to statistical professionals. The common answer from most statistical professionals is "30." So, why 30?

To answer the "Why 30?" question, we need to have a brief discussion on the use of statistics to estimate population parameters.

What is a population parameter? Within the context of this discussion, let's assume you wish to calculate the reproducibility standard deviation of a test method for a specific material. The population of interest is then the population containing single test results from an infinite number of laboratories for the specific material using the same test method, as depicted in Fig. 1. The two main population parameters of interest are the mean (μ) and the standard deviation (σ).

If we had infinite resources, patience, and time, we would collect every result from the population and calculate the exact values for μ and σ using all the results. Since we do not have infinite resources, we can take a random sample of size *n* from the target population, do some math, and come up with statistics to estimate the desired population parameters as depicted in Fig. 2. In Fig. 2, the sample average $\bar{x}$ is an estimator for μ, denoted by the "^" above the character μ. Similarly, the sample standard deviation *s* is an estimator for σ. We will focus on the statistic *s*.

Estimators have variability themselves due to random sampling. In other words, if you repeat the random sampling process again, you will most likely get a numerically different value for the same statistic. This is referred to as the variability of the sample statistic *s*, which is a function of *σ* and sample size *n*.

The uncertainty (due to sampling) when using *s* as an estimator of true (population) σ can be visualized by way of a *confidence interval*. A confidence interval represents a range of plausible values within which the truth (σ) is thought to lie, with a specified level of confidence articulated in terms of probability (typically set at 95 percent).

A 95 percent confidence interval for σ of a *normally distributed* population can be constructed by multiplying the sample standard deviation *s* with lower and upper confidence limit factors as follows :

$$LCF_{95} \cdot s \le \sigma \le UCF_{95} \cdot s$$

where LCF_{95} is the 95 percent lower confidence limit factor, and UCF_{95} is the 95 percent upper confidence limit factor.

These confidence limit factors are based on a quantity known as the degrees of freedom (*df*) for *s* ($df = n - 1$), the specified level of confidence, and an associated Chi-square statistic. The confidence limit factors for s for different *df* values are given in Table 1.

For example, if the reproducibility standard deviation *s*, based on single results from seven laboratories is 1.5, a 95 percent confidence interval within

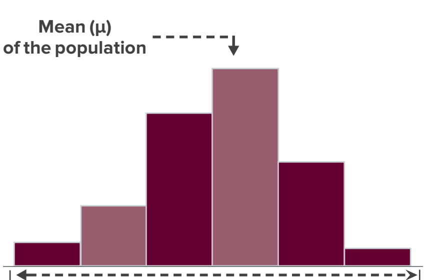
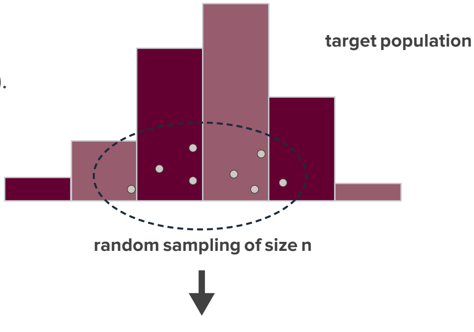


Fig. 1— Histogram of single results from the entire population of labs



$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \hat{\mu} \qquad s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}} = \hat{\sigma}$$

Fig. 2— Visualizing Use of Statistics to Estimate Population Parameters

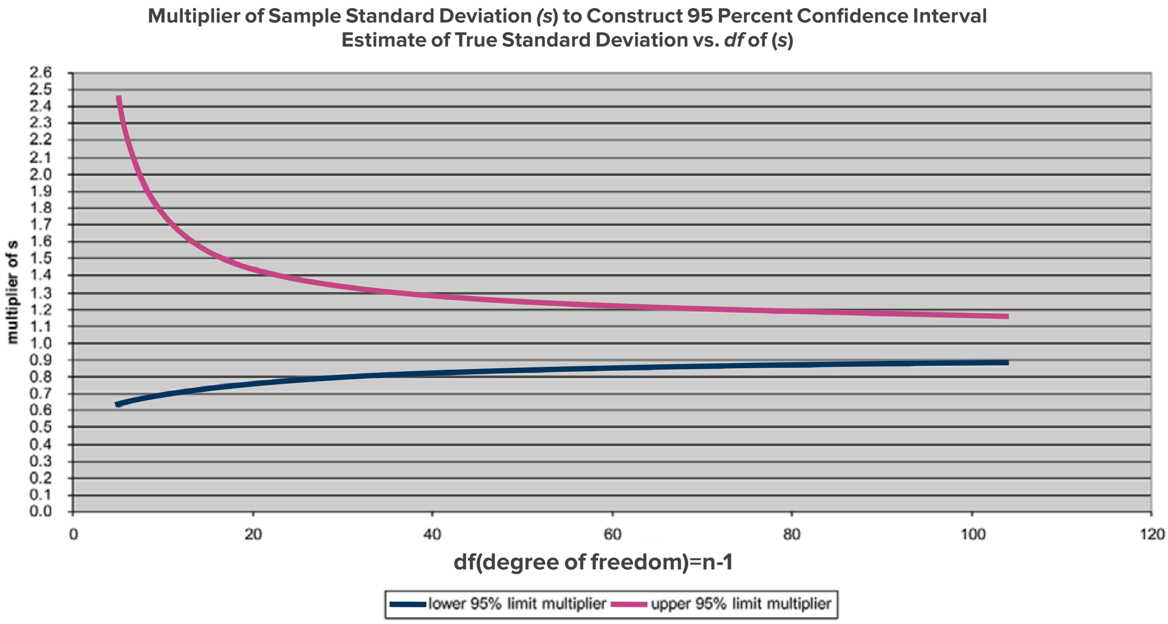


Fig.3 — Visualizing 95 Percent Confidence Interval as a Function of *df.*

Table 1 — Multipliers of *s* to Construct 95 Percent Confidence Interval for σ

df = *n* - 1	Lower 95% Limit Multiplier	Upper 95% Limit Multiplier	df = *n* - 1	Lower 95% Limit Multiplier	Upper 95% Limit Multiplier
5	0.624	2.453	31	0.802	1.329
6	0.644	2.202	32	0.804	1.323
7	0.661	2.035	33	0.807	1.316
8	0.675	1.916	34	0.809	1.310
9	0.688	1.826	35	0.811	1.304
10	0.699	1.755	36	0.813	1.299
11	0.708	1.698	37	0.815	1.294
12	0.717	1.651	38	0.817	1.289
13	0.725	1.611	39	0.819	1.284
14	0.732	1.577	40	0.821	1.280
15	0.739	1.548	41	0.823	1.275
16	0.745	1.522	42	0.825	1.271
17	0.750	1.499	43	0.826	1.267
18	0.756	1.479	44	0.828	1.263
19	0.760	1.461	45	0.829	1.260
20	0.765	1.444	46	0.831	1.256
21	0.769	1.429	47	0.832	1.253
22	0.773	1.415	48	0.834	1.249
23	0.777	1.403	49	0.835	1.246
24	0.781	1.391	50	0.837	1.243
25	0.784	1.380	60	0.849	1.217
26	0.788	1.370	70	0.858	1.198
27	0.791	1.361	80	0.866	1.183
28	0.794	1.352	90	0.873	1.171
29	0.796	1.344	100	0.879	1.161
30	0.799	1.337			

which the true (population) reproducibility standard deviation σ may lie, using Table 1 for *df* = (7 − 1) = 6, is between (0.644) x 1.5 = 0.966 and (2.202) x 1.5 = 3.303. If *s* is calculated using single results from 31 laboratories, the 95 percent confidence interval is between (0.799) x 1.5 = 1.198 and (1.337) x 1.5 = 2.00. Note that the width of the confidence interval for *n* = 31 is 0.802 units, as compared to 2.337 units for *n* = 7, representing a 66 percent reduction in uncertainty by increasing the sample size from 7 to 31.

Although increasing the sample size reduces uncertainty, there is a point of diminishing return. The statistics community has generally accepted 30 degrees of freedom as providing the minimum acceptable uncertainty for estimating σ. As shown in Fig. 3, the rate of improvement (reduction in confidence interval width) begins to approach the point of diminishing return beyond 30.

The takeaway is that *while* a standard deviation estimate can be calculated from a small dataset, decisions based on such estimates (from small datasets) can be highly unreliable. More importantly, due to the distribution of the statistic *s*, the lower and upper 95 percent limit multipliers move rapidly and asymmetrically away from 1, as can be seen in Fig. 3, as the degrees of freedom approach 1. Hence, for small sample size *n* (low degrees of freedom), the risk of underestimating population σ is substantial.

Alex T.C. Lau, TCL Consulting, Whitby, Ontario, Canada, is chairman of the coordinating subcommittee on quality assurance and statistics (D02.94) and of subcommittee on precision (D02.01.0B), which are part of the committee on petroleum products, liquid fuels, and lubricants (D02). An ASTM International fellow, Lau is also a member of the committees on quality and statistics (E11), accreditation and certification (E36), and sports equipment and facilities (F08).

John Carson, Ph.D., of P&J Carson Consulting LLC, is the Data Points column coordinator. He is also chairman of the subcommittee on statistical quality control (E11.30), part of the committee on quality and statistics (E11), and a member of the committee on environmental assessment, risk management, and corrective action (E50).

Send your statistics questions to Maryann Gorman at mgorman@astm.org.