The Square Root of *N* Plus One Sampling Rule

How Much Confidence Do We Have?

Hewa Saranadasa

STEVE SMITH

The accuracy of the 95% confidence probability statement for mean was compared for three distributions for sample size obtained from the square root of *N* plus one rule with the Edgeworth approximation derived sample size. Results showed that the sample size obtained from this rule is not even enough to declare less than 20% of defectives in a moderate size population with a high degree of confidence. Therefore, the author concludes this rule should not be used to select a sampling plan to infer a population defective rate.

Hewa Saranadasa, PhD, is a senior statistician at Pharmaceutical Sourcing Group Americas, a division of Ortho-MacNeil Pharmaceutical, Inc., 1000 Route 202, PO Box 300, Raritan, NJ 08869-0602, tel. 908.218.6321, fax 908.218.0230, hsaranad@psgus.jnj.com.

In the pharmaceutical industry, bulk product characteristics are routinely tested to determine whether the product meets in-house quality assurance specifications. Raw or processed materials are frequently stored in drums or containers. In these instances, one must know how to choose a representative sample that exhibits characteristics similar to those processed by the population as a whole. The common method in statistical practices is choosing a simple random sample from the population.

A simple method for choosing a sample size from a population is through what quality engineers refer to as the square root of *N* plus one sampling rule. This rule is apparently not statistically motivated nor is it mentioned by sampling theorists, practitioners, or reviewers of the field (1–3). Izenman provides a good summary of the origin of this rule (4). According to that article, this rule apparently originated in the 1920s from a need to provide agricultural regulatory inspectors with a convenient, memorizable rule for sample-size determination. In 1925 the Association of Official Agricultural Chemists (AOAC) set up a committee to study all aspects of the sampling problem for agricultural research as well as for regulatory activities. This rule was adopted formally and documented in an unpublished report by the AOAC committee in 1927 (5) for sampling of certain classes of foods. A companion article by Paul (6) also recommended the square-root rule for the sampling of bulk or powder drugs in small packages. However, Runkel (7) and Munch and Bidwell (8) warned that the rule had not been based on any theoretical significance. Their warning was basically ignored, and in the next 40 years, AOAC's square-root rule operated as the standard for sampling such agricultural lots as wheat flour, feeding stuffs, boxed dried fruit, and sacked cacao nibs. In 1967 and 1971 two articles, Quackenbush and Rund (9) and a December report by the Salmonella Committee of the National Research Council (10) reiterated that the rule did not have any statistical motivation and referred to it as the "commonly used rule-of-thumb sampling rule." Borland (11) demonstrated that the determination of sample size of the lot size as the square root of lot size might create a sense of false security. Therefore, he recommended not to use this rule for choosing a sampling plan.

Despite the lack of theoretical support, this rule has been adopted by many federal regulatory agencies. The recommen-

Table I: Percentage of times the 95% confidence statement would be wrong for a given population size and the sample size for normal-distribution data by simulation and the Edgeworth approximation (theory).

				Edgeworth approximation			$\sqrt{N}+1$ rule	
				Percentage of failing experiments			Percentage of failing experiments	
Distribution	Percentage to be failed[a]	N	n_1	Theory	Simulation	n_2	Theory	Simulation
Normal	10%	9	6	9.6	10.8	4	11.9	14.5
$\gamma_1 = 0$,		12	6	9.6	10.4	4	11.9	14.9
$\gamma_2 = 0$		15	6	9.6	10.9	5	10.5	12.1
		21	6	9.6	10.5	6	9.6	10.8
		27	6	9.6	10.8	6	9.6	10.8
		30	6	9.6	10.5	6	9.6	10.6
		33	6	9.6	10.9	7	9.0	9.8
		36	6	9.6	10.6	7	9.0	9.8
		50	6	9.6	10.7	8	8.5	9.1
		100	6	9.6	10.6	11	7.5	7.8
		200	6	9.6	10.7	15	6.8	7.0
		500	6	9.6	10.7	23	6.2	6.3
		1000	6	9.6	10.6	33	5.8	5.9
		1500	6	9.6	10.7	40	5.7	5.7
Normal	6%	9	9	0	0	4	11.9	14.4
$\gamma_1 = 0$,		12	12	0	0	4	11.9	14.7
$\gamma_2 = 0$		15	15	0	0	5	10.5	12.1
		21	21	0	0	6	9.6	10.9
		27	27	0	0	6	9.6	10.8
		30	28	5.9	5.9	6	9.6	10.9
		33	28	5.9	6.1	7	8.9	9.9
		36	28	5.9	6.0	7	8.9	9.8
		50	28	5.9	5.9	8	8.5	9.0
		100	28	5.9	6.1	11	7.5	7.8
		200	28	5.9	6.1	15	6.8	6.9
		500	28	5.9	6.1	23	6.2	6.3
		1000	28	5.9	6.0	33	5.8	5.9
		1500	28	5.9	5.9	40	5.7	5.6

[a]This percentage rate was used to calculate the sample size from the Edgeworth approximation.

dations for the drug testing also are published by the United Nations (12), which recommends this rule for composite sampling. Interestingly, forensic chemists also use this rule to select a random sample out of seized containers to estimate quantity of illicit drugs processed by defendants for sentencing purposes (4). Some citations are given, especially in the pharmaceutical industry, in the context of laboratory testing and are found in the DPSC standards manual (13) and the FDA *Investigations Operations Manual* (IOM) (14).

Several variations of the square root of N plus one rule have been used, including the following applications:

- The square root of the lot size plus one is taken to determine the sample size for the number of drums to inspect for creating a composite sample. One test is performed on the composite sample.
- The square root of the lot size plus one is taken to determine the number of units to inspect. Each unit is tested individually. The lot on zero defectives is accepted, or, in the case of continuous measurements, the lot is accepted if the average falls within given specifications.
- The square root of the number of cartons plus one is taken to determine the number of cartons to examine. The sample size is determined by other means. For example, the sample size (n) is chosen to determine whether at least a proportion of the population, say 90%, of the items or bags meet a particular quality standard with a degree of probability, say 95% or to obtain a sample size from ANSI/ASQC Z1.4 (15) military standard sampling plan tables. In this kind of application, the rule is used only to obtain a representative sample from the population.

The focus of this article is to examine the confidence of the square root of N plus one rule of the first two applications described. First, a comparison is made of the confidence of estimating the mean of population characteristic using the unfounded square root of N plus one rule with that of the sample size obtained from the Edgeworth expansion for various distributions. A simulation study was conducted to compare the precision of the classical confidence intervals on the mean for both methods. Then, the confidence of using this rule in the second application is compared with the sampling plan derived

Table II: Percentage of times the 95% confidence statement would be wrong for a given population size and the sample size for skewed distribution data (Chi squared) by simulation and the Edgeworth approximation (theory).

			Edgeworth approximation			$\sqrt{N}+1$ rule		
	Percentage			Percentage of failing experiments			Percentage of failing experiments	
Distribution	to be failed[a]	N	n_1	Theory	Simulation	n_2	Theory	Simulation
χ_6	10%	9	8	9.3	9.1	4	18.1	15.9
$\gamma_1 = 1.15$		12	9	10.1	9.1	4	18.4	15.5
$\gamma_2 = 2$		15	10	9.8	9.3	5	15.7	13.6
		21	10	10.2	9.8	6	14.0	12.9
		27	11	9.8	9.5	6	14.2	13.1
		30	11	9.8	9.2	6	14.2	12.6
		33	11	9.9	9.2	7	12.8	11.9
		36	11	9.9	9.6	7	12.9	12.4
		50	11	10.0	9.2	8	11.9	10.9
		100	11	10.1	9.7	11	10.1	9.6
		200	11	10.2	9.3	15	8.7	8.3
		500	11	10.2	9.3	23	9.4	7.3
		1000	11	10.2	9.5	33	6.7	6.5
		1500	11	10.2	9.4	40	6.4	6.4
χ_6	6%	9	9	0	0	4	18.1	16.1
$\gamma_1 = 1.15$		12	11	7.5	7.6	4	18.4	16.3
$\gamma_2 = 2$		15	14	6.6	6.7	5	15.7	14.2
		21	20	5.9	5.4	6	14.0	13.0
		27	25	5.9	6.0	6	14.2	12.4
		30	28	5.8	6.0	6	14.2	12.6
		33	30	6.0	6.0	7	12.8	11.7
		36	32	6.1	6.0	7	12.9	11.5
		50	41	6.0	6.1	8	11.9	10.8
		100	51	6.0	6.0	11	10.1	9.5
		200	55	6.0	6.0	15	8.7	8.3
		500	56	6.0	5.8	23	9.4	7.3
		1000	57	6.0	6.0	33	6.7	6.6
		1500	57	6.0	5.9	40	6.4	6.3

[a]This percentage rate was used to calculate the sample size from the Edgeworth approximation.

from hypergeometric distribution with no defective items in the sample as an acceptance criterion of the lot with certain confidence.

The Edgeworth expansion

It is well known that the distribution of sample mean approximately follows a normal distribution when the sample size is large by the central limit theorem (CLT). One must then determine the sample size for the asymptotic to work. The general approach is to use Cochran's rule (16); that is, the minimum sample size is $25\gamma_1^2$ for the 95% confidence probability statement to be correct more than 94% of the time, in which γ_1 is the skewness of the underline distribution of what is being sampled. Sugden et al. (17) improved Cochran's rule by applying Edgeworth approximations to the distribution of the sample mean under simple random sampling using the results published in their 1997 article (18). Their results show that 28 extra sampling units are needed over Cochran's rule to compensate for not knowing the variance of the distribution. For example, if the kurtosis of the distribution is zero and the skewness is γ_1, then the smallest sample size is $28 + 25\gamma_1^2$ to have the 95% confidence probability statement to be correct more than 94% of the time.

Let $X_1, X_2, \ldots, X_N$ be the values of a variate X in a finite population of N units, and let $\bar{X}$ and

$$S^2 = \sum_{i=1}^{N} \frac{(X_i - \bar{X})^2}{N-1}$$

be the mean and finite population variance, respectively. In a simple random sample of n units drawn without replacement from this population, let $\bar{x}$ and s^2 be the corresponding sample quantities. Using the well-known results for the mean and variance of the sampling distribution of $\bar{x}$ (16), one defines

$$U_n = \sqrt{\frac{n}{1-f}}\,\frac{(\bar{x} - \bar{X})}{s}$$

Table III: Percentage of times the 95% confidence statement would be wrong for a given population size and the sample size for skewed distribution data (exponential) by simulation and the Edgeworth approximation (theory).

				Edgeworth approximation			$\sqrt{N}+1$ rule	
				Percentage of failing experiments			Percentage of failing experiments	
Distribution	Percentage to be failed[a]	N	n_1	Theory	Simulation	n_2	Theory	Simulation
Exp(0.5)	10%	9	8	10.9	9.3	4	30.5	19.1
$\gamma_1 = 2$		12	11	9.3	8.3	4	31.6	20.2
$\gamma_2 = 6$		15	13	9.4	9.4	5	26.2	18.2
		21	16	9.9	8.6	6	22.9	16.6
		27	18	9.9	8.6	6	23.4	16.4
		30	19	9.8	8.6	6	23.4	16.2
		33	19	10.0	8.7	7	20.8	15.1
		36	19	10.1	9.0	7	20.9	15.6
		50	21	9.9	9.1	8	19.0	14.2
		100	22	10.0	9.2	11	15.3	12.4
		200	23	9.9	8.9	15	12.7	10.5
		500	23	10.0	8.9	23	10.0	8.9
		1000	23	10.0	9.1	33	8.5	7.8
		1500	23	10.0	9.0	40	7.9	7.5
Exp(0.5)	6%	9	8	10.9	9.9	4	30.5	19.4
$\gamma_1 = 2$		12	11	7.5	8.1	4	31.6	20.5
$\gamma_2 = 6$		15	14	5.6	8.1	5	26.2	17.7
		21	19	6.9	7.1	6	22.9	16.3
		27	25	5.7	6.6	6	23.4	16.2
		30	28	5.3	6.3	6	23.4	16.0
		33	30	6.1	6.7	7	20.8	14.8
		36	33	5.8	5.9	7	20.9	14.4
		50	45	5.9	6.1	8	19.0	14.1
		100	78	5.9	5.8	11	15.3	12.0
		200	99	6.0	6.0	15	12.7	10.6
		500	110	6.0	6.0	23	10.0	9.2
		1000	114	6.0	5.9	33	8.5	8.0
		1500	115	6.0	6.1	40	7.9	7.6

[a]This percentage rate was used to calculate the sample size from the Edgeworth approximation.

as the studentized mean, in which

$$f = \frac{n}{N}$$

is the sampling fraction.

The Edgeworth expansion of the probability distribution of U_n can be written as

$$\Pr(U_n \leqslant u) = \Phi(u) + \frac{q_1(u)\phi(u)}{\sqrt{n}} + \frac{q_2(u)\phi(u)}{n} + O\left(n^{-3/2}\right)$$

in which

$$q_1(u) = \gamma_1\{\frac{1}{2}\sqrt{(1-f)} + \frac{1}{3}\frac{1-f/2}{\sqrt{(1-f)}}(u^2-1)\}$$

$$\phi(u) = \frac{1}{\sqrt{2\pi}}\exp(-\frac{u^2}{2})$$

$$q_2(u) = u[\gamma_2\{\frac{2-6f+3f^2}{24(1-f)}(u^2-3) - \frac{f}{2}\} - \{1 + \frac{1}{4}(u^2-3)\} - \gamma_1^2\{1-f+\frac{2-f}{3}(u^2-3)\}] - u\gamma_1^2[\frac{(1-f/2)^2(u^4-10u^2+15)}{18(1-f)}]$$

$$\gamma_1 = \frac{\frac{1}{N}\sum_{i=1}^{N}(X_i-\bar{X})^3}{\sigma^3}$$

$$\gamma_2 = \frac{1}{N}\frac{\sum_{i=1}^{N}(X_i - \bar{X})^4}{\sigma^4} - 3$$

$$\sigma^2 = \frac{(N-1)S^2}{N}$$

γ_1 and γ_2 are the skewness and kurtosis of the distribution, respectively.

Minimum sample size

To derive the minimum sample size, the author will first consider Cochran's original condition and then generalize this condition to $(1 - \alpha)100\%$ confidence probability statement on the average would be wrong not more than $(\alpha + \epsilon)100\%$ and $\epsilon > 0$. That is,

$$\Pr(U_n \leqslant u) - \Pr(U_n \leqslant -u) > 1 - \alpha - \epsilon$$

When $u = 1.96$, $\alpha = 0.05$, and $\epsilon = 0.01$ the above statement becomes Cochran's original condition. Using the Edgeworth expansion above, one obtains the following condition

$$\frac{2q_2(u)\phi(u)}{n} > -\epsilon$$

After some algebra, one obtains the quadratic inequality for n (17)

$$An^2 + Bn + C < 0$$

in which

$$A = \frac{36\epsilon}{N\phi(u)} - \frac{\gamma_2}{N^2}\{9H_3(u) + 36u\} + \frac{\gamma_1^2}{N^2}\{72u + 24H_3(u) + H_5(u)\}$$

$$B = -[\frac{36\epsilon}{\phi(u)} - \frac{\gamma_2}{N}\{18H_3(u) + 36u\} + \frac{\gamma_1^2}{N}\{144u + 72H_3(u) + 4H_5(u)\} + \frac{1}{N}\{72u + 18H_3(u)\}]$$

$$C = 72u + 18H_3(u) - 6\gamma_2 H_3(u) + \gamma_1^2\{72u + 48H_3(u) + 4H_5(u)\}$$

$H_3(u) = u^3 - 3u$ and $H_5(u) = u^5 - 10u^3 + 15u$ are the Hermite polynomials used in the derivation of the Edgeworth expansion.

The smallest root of the inequality is

$$n = \frac{-B - \sqrt{B^2 - 4AC}}{2A}$$

If the distribution is normal ($\gamma_1 = \gamma_2 = 0$), then $f = 0$ (i.e., n is very small compared with N), $u = 1.96$, and $\epsilon = 0.01$, and the above equation gives $n = 28$. This is the minimum sample size needed to satisfy Cochran's condition for normally distributed data.

A simulation study

The author simulated 100 samples of size N (N was selected to represent each of the following sample-size categories: small [9–30], moderate [30–100], and large [>100]) from each of three distributions: a normal distribution and two skewed distributions (a Chi-squared distribution with 6 degrees of freedom and an exponential distribution with parameter 0.5). One thousand bootstrap samples of size n_1 and n_2 were generated from the simulated random samples of size N without replacement. The sample sizes n_1 and n_2 correspond to the Edgeworth approximation and square root of N plus one rule. The numbers of simulated experiments in which the population averaged outside of ±1.96 standard deviation units from the sample average were counted. The experiment was repeated for the 95% confidence probability statement to be wrong not more than 6% and 10% of the time. The average percentage of failing experiments was calculated for each method. Tables I–III show the results.

Hypergeometric distribution

Define N, p, and n to be the lot size, proportion of defective items in a lot, and sample size, respectively. The number of defective items in the sample is a random number, and each item has equal probability, p, of being a defective item. The probability of the number of defective items, d, in the sample follows a hypergeometric distribution given by the following function:

$$\Pr(D = d) = \frac{\binom{Np}{d}\binom{N(1-p)}{n-d}}{\binom{N}{n}}$$

$$\max\{0, n - N + Np\} \leqslant d \leqslant \min\{n, Np\}$$

in which Np is the number of defective/nonconforming items and $N(1 - p)$ is the number of nondefective/conforming items in the population. The notation

$$\binom{N}{n}$$

indicates the number of different samples of size n drawn from a population or lot size N and is given by

$$\binom{N}{n} = \frac{N!}{(N-n)!n!}$$

in which $n! = n(n - 1)(n - 2) \dots 1$ (e.g., $4! = 4 \times 3 \times 2 \times 1 = 24$).

Table IV: Sample size from the hypergeometric distribution and the Type I error rate for the square root of *N* plus 1 rule for testing various acceptable defective levels for a lot.

Lot size	Sample size from hypergeometric distribution with 5% Type I error rate[a]				Type I error rate for $\sqrt{N}+1$ rule[a]			
N	10%	5%	1%	$\sqrt{N}+1$	20%	10%	5%	1%
9	7	8	9	4	14.3	30.0	41.3	52.4
12	9	10	12	4	20.3	38.5	51.2	63.3
15	10	12	14	5	15.4	33.8	48.0	62.5
21	13	16	20	6	13.5	32.9	49.0	66.4
27	15	19	25	6	16.1	37.2	54.4	72.5
30	15	21	27	6	17.0	38.8	56.3	74.7
33	16	22	30	7	12.8	33.7	52.2	72.7
36	17	24	32	7	13.4	34.8	53.6	74.4
50	19	29	43	8	11.5	33.0	53.4	76.9
100	23	39	78	11	6.4	25.8	48.7	79.1
200	26	47	126	15	2.8	17.7	41.4	79.1
500	28	54	196	23	0.4	7.9	28.5	75.3
1000	28	56	238	33	0	2.8	17.3	69.0
1500	28	57	255	40	0	1.8	12.1	64.8
2000	29	57	265	46	0	0.7	9.0	61.2

[a]The maximum acceptable proportions of defective columns are 20, 10, 5, and 1%. The Type I error rates reported in the table are listed as percentages.

Suppose one wants to derive a sample size by the following testing hypothesis: H_0:$p>\theta_0$ versus H_1:$p\leqslant\theta_0$ in which θ_0 is the maximum proportion of defectives that could still be considered as acceptable. Let α be the Type I error rate (i.e., accepting a lot when it has more than the maximum acceptable proportion of defectives). The required minimum sample size (n) to test the above null hypothesis with α100% significant level when no defective items found in the sample can be obtained solving the following inequality for n:

$$\frac{(N-N\theta_0-1)!(N-n)!}{N!(N-N\theta_0-n-1)!}\leqslant\alpha$$

Table IV shows the sample size from the hypergeometric distribution and the acceptance probability of a lot when it is not at the acceptable quality level (i.e., $p>\theta_0$), Type I error rate, from the square root of *N* plus one rule. The acceptance criterion is no defectives found in the sample.

The results shown in Table IV demonstrate that for this sampling plan (accept the lot if no defectives are found) the sample size is largely unaffected when the population size is large enough (>2000) for testing the given defective levels in the lot with a given confidence. For an infinite population, sample sizes 29, 58, and 297 are needed to achieve 95% confidence to declare not more than 10, 5, and 1% defective rates in the population, respectively. Obviously, more samples are needed to test smaller defective rates to achieve the same confidence. The results also showed a frequent occurrence of Type I error rates (accepting a lot when it has more than the given maximum rate of defectives) with the sample size obtained from the square root of *N* plus one rule as a sampling plan. This sample size is not even sufficient to declare less than 20% of defectives in a moderate population with a high degree of confidence. Therefore, this rule should not be used to select a sampling plan to infer a population defective rate.

Concluding comments

Using the results of the simulation study described in this article, one can make the following conclusions: If the underline distribution is normal and the population size is greater than 30, then the sample size obtained from the square root of *N* plus one rule guarantees that at least 90% of the time the 95% confidence interval would cover the population mean. If the underline distribution deviates from normal distribution to a highly skewed distribution, then the accuracy of the 95% confidence statement with sample size obtained from the square root of *N* plus one rule decreases. For example, for exponentially distributed data, the accuracy would be not more than 88% for a moderate-size population (30–100). The simulation results also showed that the sample size obtained from the Edgeworth approximation does not provide good accuracy for small samples (see the exponential distribution in Table III).

It also is very interesting to determine whether the population is large enough and whether 28 observations would give 94% accuracy of the 95% confidence statement for the mean to be true for normally distributed data. If the data distribution was moderately skewed, then a sample size of 57 would give the same accuracy as normally distributed data. This sample size is independent of the population size when the population size is large enough. Even with 12 data points, the confidence statement would be accurate >90% if the data are normal or have a moderately skewed distribution.

The results summarized in the last section showed a frequent occurrence of Type I errors (accepting a lot when it has more than the given maximum rate of defectives) with the sample size derived from the square root of *N* plus one rule as a sampling plan. This sample size is not even sufficient to declare less than 20% of defective in a moderate-size population with a high degree of confidence. Therefore, this rule should not be used to select a sampling plan to infer a population defective rate.

Acknowledgments

The author thanks the referees and editors for their helpful suggestions and careful reading of this article.

References

1. F.F. Stephan, "History of the Uses of Modern Sampling Procedures," *J. Amer. Statist. Assoc.* **43,** 12–39 (1948).
2. M.H. Hansen, "Some History and Reminiscences on Survey Sampling," *Statist. Sci.* **2,** 180–190 (1987).
3. D. Bellhouse, "A Brief History of Survey Sampling," in *Handbook Of Statistics*, P.R. Krishnaiah and C.R. Rao, Ed. (North-Holland, Amesterdam, 1988), pp. 1–15.
4. A.J. Izenman, "Statistical and Legal Aspect of the Forensic Study of Illicit Drugs," *Satist. Sci.* **16,** 1, 35–57 (2001).
5. F.C. Blanck, "Report of the Committee on Sampling," *J. Assoc. Official Agricultural Chemists* **10,** 92–98 (1927).
6. A.E. Paul, "Sampling Drugs," *J. Assoc. Official Agricultural Chemists* **10,** 99–102 (1927).
7. H. Runkel, "Report on Methods for Sampling Flour," *J. Assoc. Official Agricultural Chemists* **9,** 423–426 (1926).
8. J.C. Munch and G.L. Bidwell, "What Constitutes an Adequate Sample?" *J. Assoc. Official Agricultural Chemists,* **11,** 220-222 (1928).
9. F.W. Quackenbush and R.C. Rund, "The Continuing Problem of Sampling," *J. Assoc. Official Analytical Chemists* **50,** 997–1006 (1967).
10. E.M. Foster, "The Control of *Salmonellae* in Processed Foods: Classification Systems and Sampling Plan," *J. Assoc. Official Analytical Chemists* **54,** 259–266 (1971).
11. K. Borland, "The Fallacy of Square Root Sampling Rule," *J. Amer. Pharm. Assoc.* **39** (7), 373-377 (1950).
12. United Nations, *Recommended Method for Testing Cannabis: Manual for Use by National Narcotics Laboratories* (United Nations, Division of Narcotic Drugs,Vienna, 1987).
13. *DPSC Standards for the Manufacture and Packaging of Drugs, Pharmaceuticals and Biological Products (*Defense Personnel Support Center, Philadelphia, PA, 1968).
14. Office of Regulatory Affairs, *FDA Investigation Operations Manual*, Chapter 4, Subsection 427.2, www.fda.gov/ora/inspect_ref/iom/ChapterText/420_part2.html (2002).
15. ANSI/ASQC Z1.4, *Sampling Procedures and Tables for Inspection by Attributes* (American Society for Quality Control, Milwaukee, WI, 1993).
16. W.G. Cochran, *Sampling Techniques (*Wiley Interscience, New York, NY, 3d ed., 1977).
17. R.A. Sugden, T.M.F. Smith, and R.P. Jones, "Cochran's Rule for Simple Random Sampling," *J.R. Statist. Soc. B.* **62** (4), 787–793 (2000).
18. R.A. Sugden and T.M.F. Smith, "Edgeworth Approximations to the Distribution of the Sample Mean Under Simple Random Sampling," *Statistics & Probability Letters* **34,** 293–299 (1997). **PT**

FYI

New regulatory department

The Biotechnology Industry Organization (BIO) has established a Science and Regulatory Affairs Department with Gillian R. Woollett, MA, DPhil, as its first vice-president. Woollett was most recently associate vice-president for Biologics and Biotechnology at Pharmaceutical Research and Manufacturers of America (PhRMA).

The new department will act as a liaison between the biotechnology industry and the federal government to help c reate a responsive regulatory system.

BIO represents more than 1000 biotechnology companies, academic institutions, and state biotechnology centers worldwide. For more information, visit www.bio.org.

Circle/eINFO 44