AIHA Journal

ISSN: 1542-8117 (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/uaih20

Statistical Methods for Establishing Equivalency of a Sampling Device to the OSHA Standard

K. Krishnamoorthy & Thomas Mathew

To cite this article: K. Krishnamoorthy & Thomas Mathew (2002) Statistical Methods for Establishing Equivalency of a Sampling Device to the OSHA Standard, AIHA Journal, 63:5, 567-571, DOI: 10.1080/15428110208984741

To link to this article: https://doi.org/10.1080/15428110208984741

Statistical Methods for Establishing Equivalency of a Sampling Device to the OSHA Standard

AUTHORS
K. Krishnamoorthy[a]
Thomas Mathew[b]

[a]Department of Mathematics, University of Louisiana at Lafayette, Lafayette, LA 70504;
[b]Department of Mathematics and Statistics, University of Maryland, 1000 Hilltop Circle, Baltimore, MD 21250

Occupational Safety and Health Administration (OSHA) regulations allow the use of an alternative sampling device for exposure monitoring provided the device has been demonstrated to be equivalent to the standard device. For example, the OSHA standard allows the use of an alternate cotton dust sampler that is equivalent to the Lumsden-Lynch vertical elutriator (VE); also, OSHA defines the accuracy of the monitoring device for measuring airborne chemicals such as benzene and sulfur dioxide. Typically, the OSHA criterion is that 90% of the readings of the sampling device should be within ±25% of the readings obtained by the standard device or within ±25% of the actual airborne chemical concentration. This article proposes two statistical tests for establishing that an alternative measuring device of airborne chemicals or dust is equivalent to the OSHA standard. The statistical tests are illustrated using an example.

Keywords: cotton dust sampler, lognormal distribution, nonparametric test, power, sample size

This research was supported by grant R01-0H03628-01A1 from the National Institute of Occupational Safety and Health.

Occupational Safety and Health Administration (OSHA) regulations provide, among other things, guidelines for exposure monitoring and measurement. OSHA also specifies the standard sampling device for specific exposure monitoring, and permits the use of an alternative device of equivalent accuracy and precision. Furthermore, it provides protocol to establish that a sampling device is equivalent to the standard or is acceptable for exposure monitoring. For example, the OSHA regulation for cotton dust[(1)] specifies the vertical elutriator (VE) as the standard for measuring cotton dust in the atmosphere of a cotton mill. It also states:

> if an alternative to the vertical elutriator cotton dust sampler is used, the employer shall establish equivalency by reference to an OSHA opinion or by documenting, based on data developed by the employer or supplied by the manufacturer, that the alternative sampling device meets the following criteria: (a) It collects respirable particulates in the same range as the vertical elutriator (approximately 15 microns); (b) Replicate exposure data used to establish equivalency are collected in side-by-side field and laboratory comparisons; and (c) A minimum of 100 samples over the range of 0.5 to 2 times the permissible exposure limit are collected, and 90% of these samples have an accuracy range of plus or minus 25 per cent of the vertical elutriator reading with a 95% confidence level as demonstrated by a statistically valid protocol.

As pointed out by Rockette and Wadsworth,[(2)] OSHA typically maintains the same criterion for equivalency among various standards. Specifically, similar criteria have been used in measuring airborne benzene[(3)] (www.osha-slc.gov/OshStd_data/1910_1028.html) and sulfur dioxide.[(4)] The OSHA requirements for equivalency can be regarded as a hypothesis testing problem about θ = proportion of the alternative device readings that are within ±25% of the standard readings. Specifically, the OSHA guidelines for establishing that an alternative device is equivalent to VE lead to the following hypothesis testing problem with

$$H_0{:}\theta \le 1 - p \text{ vs } H_a{:}\theta > 1 - p \quad (1)$$

where $p = 0.1$, and the level of significance $\alpha = 0.05$. Rockette and Wadsworth[2] proposed an approximate test for establishing the equivalency of an alternative device to the standard cotton dust sampler VE. Their numerical studies, however, indicate that their method lacks the required statistical properties. There are situations where the Type I error rate of their test can be much higher than the specified level of significance.

It appears that a conclusion stronger than $\theta > 1 - p$ will be of practical interest. The stronger hypothesis requires that the proportion of alternative device readings greater than 1.25 times of the standard readings, and less that 0.75 times of the standard readings, must both be less that $p/2$. In other words, the authors would like to see that "both tails" are small enough, because if the devices are truly equivalent, the proportion of alternative device readings that are too large or too small compared with the standard readings should both be small. This stronger hypothesis implies H_a in Equation 1. The parametric test procedure that the authors have developed uses such a stronger alternative hypothesis. Also, the problem is somewhat easier to solve with the stronger alternative hypothesis.

It is well established now that exposure data usually follow lognormal distribution (see Esmen and Hammad,[5] Rappaport and Selvin,[6] Lyles and Kupper[7]). Therefore, in this article, a parametric test is first developed (assuming that the readings are lognormally distributed) for establishing the equivalency of a sampling device to the standard. The proposed test is exact and simple to use. Necessary table values are given for all practical sample sizes and requirements for establishing equivalency of a measuring device to the OSHA standard. Power and sample size calculations are also given. A nonparametric procedure is outlined for situations when the exposure data do not follow a lognormal distribution. The parametric and nonparametric test procedures are illustrated using a practical example.

A TEST WHEN THE DATA ARE LOGNORMAL

Let (X_i, Y_i), $i = 1, \ldots, n$ denote n paired readings taken by the standard device (X) and the alternative device (Y). Because the data consist of paired measurements, it is not assumed that X_i and Y_i are independent; however, independence is assumed across pairs, and this is clearly realistic. A bivariate lognormal distribution for (X_i, Y_i) also is assumed. The test procedure that is developed in this section requires only the normality of $\ln(Y) - \ln(X)$. Let θ denote the proportion of Y measurements that fall between $(1 - \delta)X$ and $(1 + \delta)X$. Because X is a positive random variable, the following can be written:

$$\theta = P\left(1 - \delta \le \frac{Y}{X} \le 1 + \delta\right) \quad (2)$$

To establish that an alternative device is equivalent to the standard one, it should be tested whether the collected data provide sufficient evidence to conclude that $\theta > 1 - p$, where p is a small specified value, say, 0.05 or 0.10. Note that in the context of the earlier example dealing with establishing an alternative cotton dust sampler to be equivalent to the vertical elutriator, $\delta = 0.25$ and $p = 0.10$.

Let $D = \ln(Y) - \ln(X)$. Further, let μ and σ, respectively, denote the mean and standard deviation of D. Under the assumption of lognormal distribution, D is normally distributed with mean μ and standard deviation σ. Let $a = \ln(1 - \delta)$ and $b = \ln(1 + \delta)$. Then Equation 2 can be written as

$$\theta = P(a \le D \le b) \quad (3)$$

Instead of the testing problem (1), the stronger hypothesis mentioned in the introduction is now formulated. This hypothesis requires that the two tail probabilities $P(D < a)$ and $P(b < D)$ are both less than $p/2$. Thus, the null and alternative hypotheses become

$$\begin{aligned} &H_0: P(D < a) \ge p/2 \quad \text{or} \quad P(b < D) \ge p/2 \quad \text{vs} \\ &H_a: P(D < a) < p/2 \quad \text{and} \quad P(b < D) < p/2 \end{aligned} \quad (4)$$

It turns out that the problem is somewhat easier to handle if the hypothesis in Equation 4 is used as opposed to that in Equation 1. It is also easy to see that H_0 in Equation 4 holds if H_0 in Equation 1 holds, and the H_a Equation 1 holds if H_a in Equation 4 holds. Thus, whenever the H_0 in Equation 4 is rejected at the level α, the H_0 in Equation 1 also will be rejected at a level not exceeding α. Therefore, in the following is developed a test procedure for the hypotheses in Equation 4. It should however be noted that Equations 1 and 4 are not equivalent. Notice that the null and alternative hypotheses in Equations 1 and 4 are set up so that rejection of the null hypothesis indicates that the devices are equivalent.

Note that D is normally distributed with mean μ and standard deviation σ. The hypotheses in Equation 4 is now stated in terms of μ and σ. It can be easily verified that $P(b < D) \le p/2$ and $P(D < a) \le p/2$ are equivalent to $\mu + z_{1-p/2}\sigma < b$ and $a < \mu - z_{1-p/2}\sigma$, respectively. Thus, Equation 4 is equivalent to the hypotheses

$$\begin{aligned} &H_0: \mu - z_{1-p/2}\sigma \le a \quad \text{or} \quad \mu + z_{1-p/2}\sigma \ge b \quad \text{vs} \\ &H_a: \mu - z_{1-p/2}\sigma > a \quad \text{and} \quad \mu + z_{1-p/2}\sigma < b \end{aligned} \quad (5)$$

The hypotheses in Equation 5 can be tested as follows. For a given data set $\{d_i = \ln(X_i) - \ln(Y_i), i = 1, \ldots, n\}$, compute

$$\bar{d} = \frac{1}{n}\sum_{i=1}^{n} d_i, \quad \text{and} \quad s_d^2 = \frac{1}{n-1}\sum_{i=1}^{n}(d_i - \bar{d})^2$$

The null hypothesis will be rejected at the level of significance α if

$$\bar{d} - ks_d > a \quad \text{and} \quad \bar{d} + ks_d < b \quad (6)$$

where k is to be determined so that

$$P(\bar{d} - ks_d > a \quad \text{and} \quad \bar{d} + ks_d < b) = \alpha \quad (7)$$

After some algebraic manipulation (see Appendix I), it can be shown that the critical value k in Equation 7 satisfies

$$\frac{1}{2^{(n-1)/2}\Gamma\left(\frac{n-1}{2}\right)} \times \int_0^{\eta^2/k^2r^2} (2\Phi(\eta - kr\sqrt{w}) - 1)e^{-w/2}w^{[(n-1)/2]-1}\,dw = \alpha \quad (8)$$

where Φ is the standard normal distribution function, $\eta = \sqrt{n}z_{1-p/2}$ and $r = \sqrt{n/(n-1)}$. We used a numerical integration procedure and a root finding method to compute the values of k that satisfy Equation 8. These values are presented in Table I for various values of n, p, and α.

POWER CALCULATION

To understand the nature of the power of the proposed test, the powers were computed using Monte Carlo simulation. Notice that the power (that is, probability of rejecting H_0 when H_a is true) is given by

TABLE I. Critical Values k Satisfying Equation 8

	$\alpha = 0.10$				$\alpha = 0.05$				$\alpha = 0.01$			
n/p	0.10	0.05	0.01	0.001	0.10	0.05	0.01	0.001	0.10	0.05	0.01	0.001
3	3.77921	4.72193	6.58854	8.77168	5.43349	6.78223	9.45417	12.58061	12.30000	15.34604	21.37613	28.43584
4	2.93653	3.62909	4.99924	6.60214	3.80066	4.69121	6.45405	8.51717	6.68139	8.23521	11.31662	14.92467
5	2.59725	3.19067	4.36356	5.73539	3.19209	3.91659	5.34873	7.02473	4.96304	6.08130	8.29000	10.88000
6	2.41329	2.95341	4.01985	5.26648	2.87519	3.51406	4.77634	6.25241	4.15711	5.07305	6.88449	9.00244
7	2.29719	2.80368	3.80295	4.97054	2.67979	3.26640	4.42439	5.77829	3.69178	4.49283	6.07527	7.92657
8	2.21673	2.69992	3.65261	4.76539	2.54652	3.09769	4.18498	5.45565	3.38886	4.11607	5.55000	7.22860
9	2.15737	2.62336	3.54166	4.61392	2.44937	2.97474	4.01062	5.22079	3.17545	3.85000	5.18153	6.73823
10	2.11158	2.56430	3.45604	4.49697	2.37500	2.88072	3.87747	5.04140	3.01615	3.65206	4.90737	6.37392
11	2.07502	2.51716	3.38765	4.40358	2.31610	2.80633	3.77197	4.89957	2.89186	3.49903	4.69430	6.09180
12	2.04512	2.47852	3.33169	4.32704	2.26813	2.74568	3.68620	4.78399	2.79309	3.37617	4.52491	5.86652
13	2.02013	2.44626	3.28484	4.26300	2.22819	2.69528	3.61480	4.68788	2.71185	3.27551	4.38566	5.68188
14	1.99887	2.41882	3.24496	4.20851	2.19435	2.65250	3.55428	4.60650	2.64373	3.19132	4.26918	5.52752
15	1.98053	2.39514	3.21058	4.16144	2.16522	2.61583	3.50236	4.53653	2.58597	3.12000	4.17024	5.39638
16	1.96452	2.37446	3.18055	4.12038	2.13994	2.58386	3.45708	4.47568	2.53596	3.05771	4.08454	5.28335
17	1.95040	2.35622	3.15403	4.08413	2.11765	2.55575	3.41735	4.42216	2.49251	3.00407	4.01053	5.18448
18	1.93783	2.33998	3.13047	4.05184	2.09788	2.53076	3.38206	4.37467	2.45412	2.95660	3.94486	5.09844
19	1.92658	2.32545	3.10934	4.02292	2.08018	2.50846	3.35048	4.33218	2.41994	2.91439	3.88680	5.02099
20	1.91644	2.31233	3.09026	3.99680	2.06424	2.48834	3.32201	4.29391	2.38939	2.87656	3.83456	4.95183
21	1.90722	2.30041	3.07293	3.97306	2.04978	2.47012	3.29625	4.25922	2.36184	2.84256	3.78768	4.88985
22	1.89881	2.28953	3.05710	3.95140	2.03661	2.45350	3.27271	4.22757	2.33682	2.81165	3.74510	4.83339
23	1.89109	2.27954	3.04259	3.93150	2.02454	2.43827	3.25123	4.19865	2.31397	2.78345	3.70626	4.78209
24	1.88398	2.27036	3.02919	3.91320	2.01342	2.42423	3.23143	4.17197	2.29310	2.75768	3.67057	4.73493
25	1.87741	2.26186	3.01683	3.89625	2.00317	2.41132	3.21313	4.14741	2.27370	2.73393	3.63775	4.69178
30	1.85071	2.22730	2.96654	3.82730	1.96158	2.35889	3.13901	4.04770	2.19657	2.63856	3.50662	4.51807
35	1.83108	2.20188	2.92950	3.77652	1.93108	2.32044	3.08469	3.97462	2.14063	2.56956	3.41164	4.39263
40	1.81590	2.18222	2.90085	3.73726	1.90755	2.29079	3.04281	3.91826	2.09790	2.51689	3.33925	4.29695
45	1.80375	2.16648	2.87790	3.70577	1.88876	2.26711	3.00932	3.87328	2.06406	2.47513	3.28168	4.22095
50	1.79375	2.15352	2.85899	3.67983	1.87331	2.24764	2.98184	3.83630	2.03633	2.44103	3.23499	4.15924
60	1.77816	2.13330	2.82950	3.63934	1.84926	2.21733	2.93903	3.77873	1.99350	2.38832	3.16258	4.06370
70	1.76646	2.11812	2.80734	3.60892	1.83122	2.19460	2.90695	3.73559	1.96168	2.34914	3.10879	3.99268
80	1.75728	2.10621	2.78995	3.58502	1.81709	2.17681	2.88181	3.70180	1.93687	2.31863	3.06690	3.93743
90	1.74984	2.09655	2.77583	3.56564	1.80564	2.16240	2.86146	3.67443	1.91688	2.29401	3.03316	3.89292
100	1.74366	2.08852	2.76410	3.54953	1.79614	2.15043	2.84455	3.65171	1.90036	2.27368	3.00524	3.85612
120	1.73391	2.07586	2.74559	3.52409	1.78117	2.13157	2.81793	3.61591	1.87442	2.24180	2.96153	3.79846
145	1.72492	2.06417	2.72850	3.50061	1.76737	2.11418	2.79340	3.58293	1.85063	2.21257	2.92147	3.74571
300	1.69859	2.02995	2.67841	3.43173	1.72702	2.06337	2.72168	3.48653	1.78181	2.12800	2.80562	3.59294
500	1.68575	2.01325	2.65396	3.39809	1.70737	2.03864	2.68680	3.43964	1.74866	2.08732	2.74999	3.51969
∞	1.64485	1.95996	2.57583	3.29051	1.64485	1.95996	2.57583	3.29051	1.64485	1.95996	2.57583	3.29051

$$P(\bar{d} - ks_d > a \quad \text{and} \quad \bar{d} + ks_d < b \quad \text{under } H_a)$$

where the alternative hypothesis H_a is given in Equation 5. The size of the test, that is, probability of rejecting the H_0 when it is actually true, is given by

$$P(\bar{d} - ks_d > a \quad \text{and} \quad \bar{d} + ks_d < b \quad \text{under } H_0)$$

where the null hypothesis H_0 is given in Equation 5. The size of test should be less than or equal to the nominal level α. For a given a and b, the size of the test should be equal to α when $\mu - z_{1-p/2}\sigma = a$ and $\mu + z_{1-p/2}\sigma = b$, or equivalently, $\mu = (a + b)/2$ and $\sigma = (b - a)/(2z_{1-p/2})$. In view of the OSHA criterion given in the introduction, the powers were computed when $\delta = 0.25$, $p = 0.1$, and $\alpha = 0.05$, and for sample sizes 10, 20, 25, and 100. The powers and sizes are computed using the Monte Carlo method with 100,000 runs. The powers are reported in Table II.

The sizes of the test are given in the first five rows of Table II. It is clear that the size of the test is always less than or equal to α when $\mu - z_{1-p/2}\sigma \leq a$ or when $\mu + z_{1-p/2}\sigma \geq b$, and is equal to the nominal level 0.05 when $\mu - z_{1-p/2}\delta = a$ and $\mu + z_{1-p/2}\sigma = b$. It can also be observed from Table II that for a given n the power is an increasing function of $(\mu - z_{1-p/2}\sigma) - a$ and $b - (\mu + z_{1-p/2}\sigma)$. Furthermore, for a given $(\mu - z_{1-p/2}\sigma) - a > 0$ and $b - (\mu + z_{1-p/2}\sigma) < 0$, the power increases with respect to the sample size. Thus, Table II indicates that the power function of the proposed test satisfies all the natural requirements.

In Table III is given sample sizes required to have a power of 0.85 at the nominal level $\alpha = 0.05$, $\delta = 0.25$, and $p = 0.10$. For example, if an experimenter believes that the 5th percentile of $\ln(Y/X)$ is around -0.21, and the 95th percentile of $\ln(X/Y)$ is about 0.15, then the required sample size to attain a power of 0.85 is 35.

A NONPARAMETRIC APPROACH

If the data do not satisfy any parametric distributional assumption, then the following distribution-free approach can be used to test

$$H_0: \theta \leq 1 - p \text{ vs } H_a: \theta > 1 - p \qquad \textbf{(9)}$$

where p as defined in the previous section, and θ is the proportion

TABLE II. Powers and Sizes of the Test Equation 5

μ	σ	$\mu - z_{1-p/2}\sigma$	$\mu + z_{1-p/2}\sigma$	Sample Size 10	20	25	100
0	0.15528	−0.25541	0.25541	0.043	0.037	0.033	0.011
−0.1	0.15528	−0.35541	0.15541	0.027	0.013	0.001	0.000
−3.22693E−2	0.17	−0.31189	0.24736	0.026	0.019	0.015	0.003
−3.22693E−2	0.16	−0.29545	0.23091	0.039	0.037	0.034	0.022
−3.22693E−2	0.15528	−0.28768	0.22314	0.050	0.050	0.050	0.050
−3.22693E−2	0.14	−0.26255	0.19801	0.097	0.139	0.156	0.377
−3.22693E−2	0.13	−0.24610	0.18156	0.150	0.250	0.294	0.752
−3.22693E−2	0.11	−0.21320	0.14866	0.359	0.632	0.726	0.998
−3.22693E−2	0.10	−0.19675	0.13222	0.509	0.830	0.904	0.999
−3.22693E−2	0.09	−0.18031	0.11577	0.689	0.955	0.984	1.000

Note: $\delta = 0.25$; $\alpha = 0.05$; $p = 0.10$; $a = \ln(1-\delta) = -0.28768$; and $b = \ln(1+\delta) = 0.22314$.

of the alternative device readings Y that fall between $(1-\delta)X$ and $(1+\delta)X$. That is,

$$\theta = P((1-\delta)X \leq Y \leq (1+\delta)X) \tag{10}$$

Note that the number of Y measurements that fall in $((1-\delta)X, (1+\delta)X)$ follows a binomial distribution with the number of trials n and success probability θ. Thus, the problem of testing equivalency of sampling devices is reduced to hypothesis testing about a binomial success probability. Let k be the number of Y measurements that actually fall in $((1-\delta)X, (1+\delta)X)$. Using the well-known normal approximation to a binomial distribution, one gets a $100(1-\alpha)\%$ one-sided lower confidence limit for θ as

$$\hat{\theta} - z_{1-\alpha}\sqrt{\frac{\hat{\theta}(1-\hat{\theta})}{n}}$$

where $\hat{\theta} = k/n$. The null hypothesis in Equation 9 will be rejected if the above lower limit is greater than $(1-p)$. This approximate lower limit is satisfactory provided $n\theta > 5$ and $n(1-\theta) > 5$. An exact method due to Clopper and Pearson[8] gives the lower confidence limit for θ as

$$\theta_l = \text{Beta}(n, n-k+1, \alpha) = \frac{2nF_{2n,2(n-k+1),\alpha}}{2(n-k+1) + 2nF_{2n,2(n-k+1),\alpha}}$$

where Beta(a, b, c) denotes the 100cth percentile of the beta distribution with shape parameters a and b and $F_{a,b,c}$ denotes the 100cth percentile of the F distribution with numerator degrees of freedom a, and the denominator degrees of freedom b. The null hypothesis in Equation 9 will be rejected if $\theta_l > 1 - p$. In other words, it is concluded that the sampling devices are equivalent if the lower bound for θ is larger than $1 - p$.

EXAMPLE

The testing methods presented in the preceding sections are now illustrated using the data set given in Rockette and Wadsworth.[2] The data represent 60 pairs of readings taken at four different work sites. The readings VE are from vertical elutriator and the readings AD are from an alternative device. For easy reference, the data are given in Table IV.

First, the method given for lognormally distributed data is applied to test whether the AD device is equivalent to the VE. Probability plots using the software MINITAB indicated that the lognormality holds for the difference ln(VE) − ln(AD). Therefore, the parametric method is applicable for this example. Let d = ln(VE) − ln(AD). Then, the computed values are $\bar{d} = 0.0020422$ and $s_d = 0.0550842$. When $n = 60$, $p = 0.10$, and $\alpha = 0.05$, the critical value $k = 1.84296$. Further, when $\delta = 0.25$, $a = \ln(1-\delta) = -0.2877$, and $b = \ln(1+\delta) = 0.2231$. Recall that the null hypothesis $H_0{:}\theta < 0.90$ will be rejected in favor of $H_a{:}\theta > 0.90$ if $\bar{d} - ks_d > a$ and $\bar{d} + ks_d < b$ (see Equation 6). For this example,

$$\bar{d} - ks_d = -0.0998 > a \text{ and } \bar{d} + ks_d = 0.1039 < b.$$

Thus, it can be concluded that the data provide enough evidence to show that the alternative sampling device satisfies all the requirements of OSHA regulations to be equivalent to the VE.

Now the nonparametric procedure for this example will be applied. For the data given in Table IV, observe that the AD reading falls between 0.75 × VE and 1.25 × VE for each pair (VE, AD). So $n = 60$, and the observed number of successes is 60. Notice that $\bar{\theta} = k/n = 60/60 = 1$. Therefore, the formula for the standard error of $\bar{\theta}$, that is, $\sqrt{\bar{\theta}(1-\bar{\theta})/n}$ is not valid, and hence the approximate lower limit based on the normal approximation (see the previous section) cannot be constructed. On the other hand, using the exact method due to Clopper and Pearson,[8] the 95% one-sided lower limit for θ can be estimated as

$$\frac{120F_{120,2,0.05}}{2 + 120F_{120,2,0.05}} = \frac{120 \times 0.325544}{2 + 120 \times 0.325544} = 0.951297$$

TABLE III. Sample Sizes to Attain a Power of 0.85

μ	σ	$\mu - z_{1-p/2}\sigma$	$\mu + z_{1-p/2}\sigma$	k	Sample Size	Power
−3.22693E−2	0.14	−0.26255	0.19801	1.77347	133	0.8502
−3.22693E−2	0.13	−0.24610	0.18156	1.84333	63	0.8535
−3.22693E−2	0.11	−0.21320	0.14866	1.93108	35	0.8546
−3.22693E−2	0.10	−0.19675	0.13222	2.04978	21	0.8467
−3.22693E−2	0.09	−0.18031	0.11577	2.19435	14	0.8501

Note: $\delta = 0.25$; $\alpha = 0.05$; $p = 0.10$; $a = \ln(1-\delta) = -0.28768$, and $b = \ln(1+\delta) = 0.22314$.

TABLE IV. VE and AD Readings in Micrograms per Cubic Meter

Site 1		Site 2		Site 3		Site 4	
VE	AD	VE	AD	VE	AD	VE	AD
75	72	120	118	250	230	500	521
80	65	122	127	258	262	522	500
82	84	118	110	300	280	480	495
74	79	132	126	375	402	602	610
85	85	114	120	314	322	620	640
89	93	125	129	270	305	525	536
90	86	140	132	285	287	490	508
84	80	117	125	320	329	544	520
77	85	113	113	305	280	610	630
78	72	125	130	275	285	485	495
74	77	126	120	320	305	560	533
80	80	118	115	318	300	525	527
81	86	125	120	292	275	532	540
84	83	129	134	340	320	547	560
82	84	132	140	360	350	590	589

Because this one-sided lower limit for the proportion of agreement is greater than 0.90, this test also supports the authors' earlier conclusion (based on the parametric test) that the AD is equivalent to the VE.

DISCUSSION

It should be clear now that the tests proposed in this article are more general and simple to use. If there is a situation that demands that 95% of the readings should be within $\pm10\%$ of the standard with 99% confidence, then the tests can be used by taking $p = 0.05$, $\alpha = 0.01$, and $\delta = 0.10$. Regarding the choice between the parametric and nonparametric tests, note the following: use of a binomial distribution to test the proportion of agreement of the readings can be justified if the readings do not follow a parametric distribution. However, also note that the use of a binomial distribution ignores the degree of agreement of the readings (i.e., the magnitude of the difference between the readings) of the samplers. It is intuitive that dichotomizing the numerical data leads to loss of information, and as a result any testing procedure based on dichotimized data are less efficient than those based on the actual numerical data. Therefore, if the readings follow a lognormal distribution (even if approximately), then the parametric test developed for lognormal data should be used.

ACKNOWLEDGMENTS

The authors thank a referee and the editorial board for their valuable comments and suggestions.

REFERENCES

1. "Occupational Exposures to Cotton Dust: Final Mandatory Occupational Safety and Health Standards, Part III" *Federal Register* (23 June 1978).
2. **Rockette, H.E., and H.M. Wadsworth:** Equivalency of alternative cotton dust samplers to the vertical elutriator. *Am. Ind. Hyg. Assoc. J. 46*:716–719 (1985).
3. "Occupational Exposure to Benzene." *Federal Register* (3 May 1977).
4. "Sulfur Dioxide Occupational Exposure." *Federal Register* (24 November 1975).
5. **Esmen, N., and Y. Hammad:** Lognormality of environmental sampling data. *Environ. Sci. Health A12*:29–41 (1977).
6. **Rappaport, S.M., and S. Selvin:** A method for evaluating the mean exposure from a lognormal distribution. *Am. Ind. Hyg. Assoc. J. 48*: 374–379 (1987).
7. **Lyles, R.H., and L.L. Kupper:** On strategies for comparing occupational exposure data to limits. *Am. Ind. Hyg. Assoc. J. 57*:6–15 (1996).
8. **Clopper, C.J., and E.S. Pearson:** The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika 26*:404–413 (1934).

APPENDIX I

For a given a and b, let μ_0 and σ_0 be respectively the values of μ and σ such that $a = \mu_0 - z_{1-p/2}\sigma_0$ and $b = \mu_0 + z_{1-p/2}\sigma_0$. Note that in order to control Type I error rate, determine k such that

$$P(\bar{d} - ks_d > \mu_0 - z_{1-p/2}\sigma_0 \text{ and } \bar{d} + ks_d < \mu_0 + z_{1-p/2}\sigma_0) = \alpha \tag{I.1}$$

After standardizing $\bar{d}$ in (I.1), it can be seen that (I.1) is equivalent to

$$P\left(\frac{Z/\sqrt{n} + z_{1-p/2}}{s_d/\sigma_0} > k \quad \text{and} \quad \frac{-Z/\sqrt{n} + z_{1-p/2}}{s_d/\sigma_0} > k\right) = \alpha \tag{I.2}$$

where Z is the standard normal random variable. let $r = \sqrt{n/(n-1)}$, $\delta = \sqrt{n}z_{1-p/2}$, and $W = [(n-1)s_d^2]/\sigma_0^2$. In terms of these notations, the two inequalities $(Z/\sqrt{n}+z_{1-p/2})/(s_d/\sigma_0) > k$ and $(-Z/\sqrt{n}+z_{1-p/2})/(s_d/\sigma_0) > k$ are together equivalent to $-\delta + kr\sqrt{W} < Z < \delta - kr\sqrt{W}$, which cannot hold unless $-\delta + kr\sqrt{W} < \delta - kr\sqrt{W}$, i.e., unless $W < \delta^2/k^2r^2$. Thus, (I.2) can be equivalently expressed as

$$\begin{aligned}
\alpha &= E_W[P(-\delta + kr\sqrt{W} < Z < \delta - kr\sqrt{W} \mid W)] \\
&= E_W\left[P\left(-\delta + kr\sqrt{W} < Z < \delta - kr\sqrt{W} \mid W < \frac{\delta^2}{k^2r^2}\right)\right] \\
&\quad + E_W\left[P\left(-\delta + kr\sqrt{W} < Z < \delta - kr\sqrt{W} \mid W \geq \frac{\delta^2}{k^2r^2}\right)\right] \\
&= E_W\left[P\left(-\delta + kr\sqrt{W} < Z < \delta - kr\sqrt{W} \mid W < \frac{\delta^2}{k^2r^2}\right)\right]
\end{aligned} \tag{I.3}$$

where E_W denotes the expectation with respect to the distribution of W. Noticing that W is a chi-square random variable with degrees of freedom $n - 1$, (I.3) can be written as

$$\frac{1}{2^{(n-1)/2}\Gamma\left(\frac{n-1}{2}\right)} \int_0^{\delta^2/k^2r^2} (2\Phi(\delta - kr\sqrt{w}) - 1)e^{-w/2}w^{[(n-1)/2]-1}\,dw = \alpha$$