American Industrial Hygiene Association Journal

ISSN: 0002-8894 (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/aiha20

The Distribution of Student's t-Statistic for Small Samples from Lognormal Exposure Distributions

JOHN S. EVANS & NEIL C. HAWKINS

To cite this article: JOHN S. EVANS & NEIL C. HAWKINS (1988) The Distribution of Student's t-Statistic for Small Samples from Lognormal Exposure Distributions, American Industrial Hygiene Association Journal, 49:10, 512-515, DOI: 10.1080/15298668891380141

To link to this article: https://doi.org/10.1080/15298668891380141

Published online: 04 Jun 2010.

Am. Ind. Hyg. Assoc. J. 49(10):512-515 (1988)

The Distribution of Student's t-Statistic for Small Samples from Lognormal Exposure Distributions

JOHN S. EVANS and NEIL C. HAWKINS*

Department of Environmental Science and Physiology, Harvard University—School of Public Health, 665 Huntington Avenue, Boston, MA 02115

To assess compliance with industrial hygiene exposure criteria (*e.g.,* TLVs®), it may be necessary to perform statistical tests of hypotheses based on relatively small samples. For pollutants with long biological half-lives, the parameter most relevant for determining the risk faced by workers is the long-term arithmetic average concentration of the pollutant. In industrial environments it is common for pollutant concentrations to be approximately lognormal. Unfortunately, when based on small samples from lognormal distributions, the ordinary Student's t-statistic has some undesirable characteristics which are not recognized widely by practicing industrial hygienists. The difficulties in using the ordinary Student's t-statistic to evaluate the average exposure have been demonstrated. The properties of alternative test statistics have been explored. Some general observations on the implications of these findings have been made.

Introduction

To assess compliance with industrial hygiene exposure criteria (*e.g.,* TLVs®), it may be necessary to perform statistical tests of hypotheses based on relatively small samples. For pollutants with long biological half-lives, the parameter most relevant for determining the risk faced by workers is the long-term arithmetic average concentration of the pollutant. Although the Occupational Safety and Health Administration (OSHA) currently does not interpret exposure standards as limits on long-term arithmetic average concentrations, support for such an interpretation is growing.[1]

In industrial environments it is common for pollutant concentrations to be approximately lognormal. Unfortunately, when based on small samples from lognormal distributions, the ordinary Student's t-statistic has some undesirable characteristics which are not recognized widely by practicing industrial hygienists. In this paper, the authors demonstrate the difficulties in using the ordinary Student's t-statistic to evaluate the average exposure, explore the properties of alternative test statistics, and make some general observations on the implications of these findings.

The Student's t-Statistic

Two simple, one-sided hypothesis tests used in industrial hygiene are the compliance officer's test (designed to demonstrate noncompliance) and the plant industrial hygienist's hypothesis test (designed to demonstrate compliance). Both are based on the Student's t-statistic:

$$t = \frac{\bar{x} - x^*}{s_x/\sqrt{n}}$$

where $\bar{x}$ is the sample mean, x^* is the applicable standard, s_x is the sample standard deviation, and n is the sample size. In 1908, Gosset derived the distribution of the t-statistic for samples of random variables drawn from normal distributions, and he empirically verified the accuracy of his results using 750 random samples of size 4 drawn from data on the lengths of the middle fingers and the heights of 3000 criminals.[2]

The distribution of t depends on the sample size. As n approaches infinity, the t-distribution asymptotically approaches the normal distribution. Table I gives the values of t corresponding to several percentiles of the distribution for sample sizes of 5 and 30.[3] The results are symmetric about zero, and for n = 30, the values of the t-statistic are very close to the familiar 1.65 and 1.96 values associated with the 5th (95th) and 2.5th (97.5th) percentiles of the normal distribution. The tabulated values are strictly applicable only when samples are drawn from normal populations; however, they are approximately correct for many other cases in which the tails of the sampling distribution are not "too heavy."[4]

To use the t-statistic, one specifies an acceptable risk of Type I error and, on the basis of the number of samples collected, selects the value of t associated with this choice. For example, a compliance officer who was comfortable with a 5% risk of falsely declaring a violation and who had collected 5 air samples would use a critical t-value of +2.13 to identify violations of exposure standards. Sample data which resulted in t-statistics > 2.13 would be viewed as

TABLE I
Some Values of the t-Statistic

Percentile (%)	t-Statistic n = 5	n = 30
2.5	−2.78	−2.05
5.0	−2.12	−1.70
95.0	+2.12	+1.70
97.5	+2.78	+2.05

*Current address: The Dow Chemical Company, Health and Environmental Sciences, 1803 Building, Midland, MI 48674.

TABLE II
Empirical Values of the t-Statistic

Percentile (%)	Theory	t-Statistic (n = 5) Geometric Standard Deviation 1.01	1.5	2	3
2.5	-2.78	-2.75 ± .04	-4.18 ± .12	-5.54 ± .16	-9.86 ± .22
5.0	-2.12	-2.12 ± .03	-3.14 ± .06	-4.08 ± .10	-7.20 ± .15
95.0	+2.12	+2.20 ± .05	+1.54 ± .02	+1.38 ± .02	+1.14 ± .02
97.5	+2.78	+2.84 ± .08	+1.90 ± .03	+1.70 ± .02	+1.37 ± .02

providing adequate evidence of violation of exposure standards. Similarly, a plant industrial hygienist who was comfortable with a 5% risk of falsely declaring compliance and who had collected 5 air samples would use a critical t-value of −2.13 to demonstrate compliance with exposure standards. Sample data which yielded t-statistics < −2.13 would be viewed as providing adequate evidence of compliance.

As illustrated below, these hypothesis tests do not behave as expected when based on small samples drawn from lognormal distributions.

Methods

Random samples of size 5 were drawn repeatedly from lognormal distributions. For each sample of size 5, the t-statistic was computed. The empirical distributions of the t-statistic were generated by constructing cumulative frequency distributions of the t-statistics resulting from 1000 samples of size 5. To assess the precision of the results, 20 independent realizations of 1000 samples of size 5 were generated for each case of interest. The variability of estimates of the fractiles of the t-distributions was assessed by calculating the standard error of each calculated percentile.

Using this approach, the empirical distribution of the t-statistic was assessed for random samples of size 5 from lognormal distributions with standard geometric deviations of 1.01, 1.5, 2 and 3.

Results

The results are summarized in Table II. For the essentially normal case (GSD = 1.01), the empirical results are virtually identical to the theoretical distribution of the t-statistic. For the other cases more relevant for industrial hygiene, however, the empirical results do not conform to the t-distribution. The degree of divergence increases with increasing variability in the environment of interest.

These results indicate that the actual risk of Type I error faced by a compliance officer is less than expected, but the risk of Type I error faced by the plant industrial hygienist is greater than expected. Table III gives the actual risks of Type I error that would be faced by industrial hygienists evaluating samples of size 5 from lognormal distributions. The results are based on the assumption that critical t-values of −2.13 and +2.13, which nominally result in 5% Type I errors for this sample size, are used in evaluation of the data.

The obvious problem with the ordinary t-statistic is that the plant industrial hygienist using it, in many cases, would be left with a false sense of security. Data, analyzed using the ordinary t-statistic, might suggest that a work environment was safe when in fact it was not.

A second problem is related to the asymmetry in the power of the test. In the design of a hypothesis test, one must be concerned with not only the risk of Type I error but also with the power of the test. A compliance officer interested in evaluating the power of a test based on the t-statistic might rely on the theoretical distribution of the t-statistic in making power calculations. For a sample size of 5 and a specified risk of Type I error of 5%, the power of the test would be estimated using the following equation:

$$\text{Power} = P(t > 2.13)$$

Figures 1 and 2 compare the theoretical estimates of power of the hypothesis tests with the empirical estimates

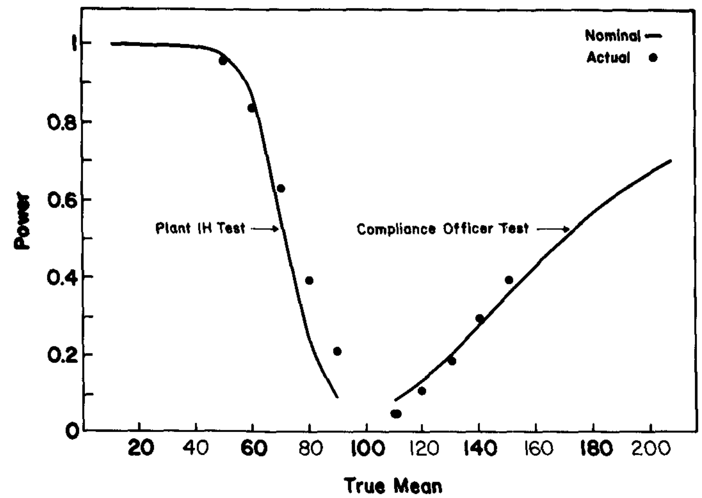

Figure 1—Nominal and actual powers of t-tests (GSD = 1.5, n = 5, and 5% Type I error).

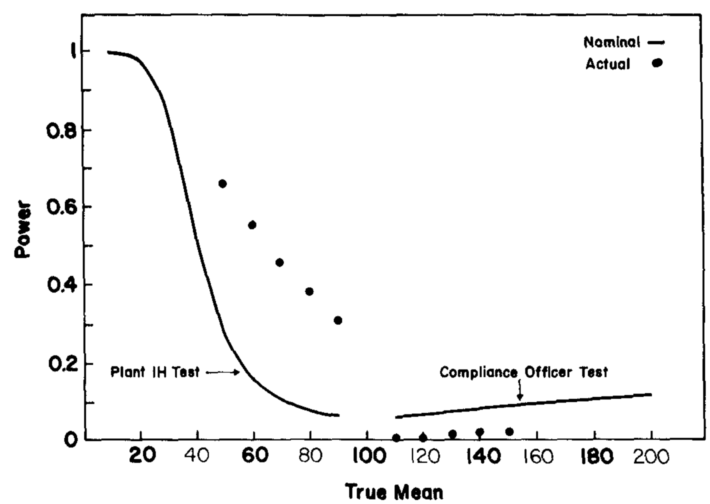

Figure 2—Nominal and actual powers of t-tests (GSD = 3.0, n = 5, and 5% Type I error).

obtained by Monte Carlo simulation. For environments with relatively little variability (*e.g.*, GSD = 1.5), the empirical estimates of power compare well with the theoretical estimates. The only region of appreciable divergence is near the standard (*e.g.*, 90 < true mean < 110), where there is a decided asymmetry in the performance of the test. For environments with greater variability (*e.g.*, GSD = 3.0), this asymmetry is more pronounced. The power to demonstrate compliance when the true mean is 50% of the standard is greater than 65%. In contrast, the power to demonstrate noncompliance when the true mean is 150% of the standard is less than 5%. In addition to the actual asymmetry in the power of the t-test, there are systematic differences between the theoretical and empirical estimates of power. The theoretical estimates of the power of the t-test underestimate its true power to demonstrate compliance and overestimate its power to demonstrate noncompliance.

Alternatives to the Ordinary t-Statistic

In view of these difficulties, it is important to consider alternative approaches that might be used in assessing compliance with industrial exposure standards. In the standard calculation of the t-statistic, the sample mean and standard deviation are estimated directly from the data:

$$\bar{x} = (1/n) \Sigma x_i$$

$$s_x^2 = [1/(n-1)] \Sigma (x_i - \bar{x})^2$$

One alternative would involve calculation of $\bar{y}$ and s_y, the sample mean and standard deviation of the natural logarithms of the data, and subsequent estimation of $\bar{x}$ and s_x, using the relationships given by Aitchison and Brown[5]:

$$\bar{y} = (1/n) \Sigma \ln (x_i)$$

$$s_y^2 = [1/(n-1)] \Sigma [\ln (x_i) - \bar{y}]^2$$

$$\bar{x} = \exp (\bar{y} + 0.5\, s_y^2)$$

$$s_x^2 = \bar{x}^2 [\exp (s_y^2) - 1]$$

A second alternative, recently proposed by Rappaport and Selvin,[6] involves a slight variation on this basic concept. In their test, the standard, x*, is substituted for the sample mean, $\bar{x}$, in the estimation of the sample variance, s_x^2, and $\sqrt{n-2}$ is used rather than $\sqrt{n}$ as the denominator in the calculation of the t-statistic.

TABLE III
Risk of Type I Error

Geometric Standard Deviation	Actual Risk of Type I Error	
	Plant IH	Compliance Officer
1.5	10.5 ± .4	1.7 ±.1
2	15.1 ± .4	1.2 ±.1
3	25.5 ± .4	0.3 ±.1

TABLE IV
Comparative Power Analysis (GSD = 1.5, n = 5)

	True Mean						
Test Statistic	50	70	90	100	110	130	150
Ordinary t-test	96	64	22	10/2	5	19	40
Lognormal t-test	93	57	19	9/2	5	18	37
Rappaport's t-test	54	22	6	3/3	6	25	52

TABLE V
Comparative Power Analysis (GSD = 3.0, n = 5)

	True Mean						
Test Statistic	50	70	90	100	110	130	150
Ordinary t-test	66	46	31	26/0	0.5	1.3	2.3
Lognormal t-test	47	31	21	16/0	0.3	0.6	1.1
Rappaport's t-test	8	4	3	2/4	5	10	15

As shown in Tables IV and V, there are no appreciable differences between the results obtained by working with the raw data and those obtained by first transforming the data. The modifications introduced by Rappaport and Selvin,[6] do have a beneficial effect. Of most importance, with these modifications, the empirical estimates of the Type I error of the test are maintained below the nominal 5% level. In addition, these modifications have the effect of greatly reducing the asymmetry in the power of the test.

It is worth noting, however, that the sample size calculations given by Rappaport and Selvin[6] are somewhat optimistic. For example, they suggest that in an environment with a GSD of 3, a sample size of 5 is adequate to achieve 90% power (with 5% Type I error) for a comparison of a true mean of 300 with a standard of 100. These simulations indicate that the power of Rappaport and Selvin's test in this circumstance is only about 60%.

Conclusions

This analysis clearly demonstrates the problems that arise when the ordinary t-statistic is used to evaluate compliance with occupational health standards. Nominal estimates of the rates of Type I error and of the power of the test are misleading. These problems are most pronounced when small samples are drawn from extremely skewed distributions.

The problems are not eliminated simply by working with logarithmically transformed data. Fortunately, the modifications recently proposed by Rappaport and Selvin[6] are beneficial and represent the best approach available to practicing industrial hygienists. Industrial hygienists using their approach, however, should be aware that, even with these modifications, the t-statistic is not distributed symmetrically and that the actual power of the test is somewhat less than Rappaport and Selvin[6] suggest.

Acknowledgment

The authors thank Elizabeth Miesner for assisting with many of the preliminary calculations involved in this project. Support for this project was provided, in part, by the Occupational Safety and Health Educational Resource Center, NIOSH Grant Award 5 T15 OH07096.

References

1. "Occupational Exposure to Benzene: Final Rule" *Federal Register 52* (11 Sept. 1987). pp. 34460–34578.
2. **Kruskal, W.H. and J.M. Tanur:** *International Encyclopedia of Statistics.* New York: The Free Press, 1978. pp. 409–413.
3. **Burington, R.S. and D.C. May:** *Handbook of Probability and Statistics with Tables.* New York: McGraw-Hill, 1970.
4. **Ott, L.:** *An Introduction to Statistical Methods and Data Analysis.* North Scituate, Mass.: Duxbury Press, 1977. p. 102.
5. **Aitchison, J. and J.A.C. Brown:** *The Lognormal Distribution—With Special Reference to Its Use in Economics.* Cambridge, United Kingdom: Cambridge University Press, 1957. p. 38.
6. **Rappaport, S.M. and S. Selvin:** A Method for Evaluating the Mean Exposure from a Lognormal Distribution. *Am. Ind. Hyg. Assoc. J. 48*:374–387 (1987).

10 October 1986; Revised 6 May 1988