JOURNAL OF FORENSIC SCIENCES

J Forensic Sci, January 2010, Vol. 55, No. 1
doi: 10.1111/j.1556-4029.2009.01256.x
Available online at: interscience.wiley.com

TECHNICAL NOTE
CRIMINALISTICS

Reinoud D. Stoel,[1] *Ph.D. and Annabel Bolck,*[1] *Ph.D.*

Correction to Tzidony and Ravreby (1992): A Statistical Approach to Drug Sampling: A Case Study

ABSTRACT: In 1992, Tzidony and Ravreby presented a confidence interval for the total weight of a seizure of illicit drugs present in a population. Their approach has subsequently been applied by several researchers in the field. The formula on which their approach is based does, however, not fully take into account the proportion of drug units found in the sample. In this paper, a modification is presented that consistently uses the correct sample size in all terms of the confidence interval, based on the proportion of drug units found in the sample. The effective sample size is smaller than the original sample size, and this should consequently be accounted for in the estimation of the standard error and in the corresponding *t*-distribution. The new confidence interval is again based on the assumption that the proportion of drug units in the population is known after sampling.

KEYWORDS: forensic science, drugs, statistics, quantity estimation, sampling, toxicology

Quite some literature has been devoted to the determination of a lower limit to the proportion of units in a population which contains drugs, at a given confidence level. One of the first references is the paper by Frank et al. (1), followed by a paper by Tzidony and Ravreby (2) in which they extended the approach with a method for calculating the total weight of a drug present in a population within a given confidence interval.

The method presented by Tzidony and Ravreby (2) has subsequently been used and applied by several scholars in the field of the forensic sciences (3–6). Their method is based on the assumption that the weight of a single unit is a normally distributed random variable. Given the common practice of relatively small sample size it is appropriate to use the *t*-distribution to provide a (1-α)100% confidence interval for the average weight of a drug unit in the population, by their formula 7

$$\bar{X} - \frac{S}{\sqrt{n}} t_{n-1,1-\alpha/2} \le \mu \le \bar{X} + \frac{S}{\sqrt{n}} t_{n-1,1-\alpha/2} \qquad (1)$$

where μ and $\bar{X}$ are the average weight of a drug unit in, respectively, the population and the sample, S is the sample standard deviation, n is the sample size, and $t_{n-1,1-\alpha/2}$ is the $1-\alpha/2$ percentile of the *t*-distribution, with $n-1$ degrees of freedom.

When the total population is finite of size N, and n/N is >0.1, a finite population correction factor, $\sqrt{\frac{N-n}{N}}$, should be used. This results in Tzidony and Ravreby's formula 8

$$\bar{X} - \frac{S}{\sqrt{n}} t_{n-1,1-\alpha/2} \sqrt{\frac{N-n}{N}} \le \mu \le \bar{X} + \frac{S}{\sqrt{n}} t_{n-1,1-\alpha/2} \sqrt{\frac{N-n}{N}} \qquad (2)$$

[1]Netherlands Forensic Institute, The Hague, The Netherlands.

Received 29 May 2008; and in revised form 14 Oct. 2008; accepted 31 Dec. 2008

The confidence interval for the total weight (W) of the drug exhibit is subsequently obtained by their formula 9, which adds the amount of drug units in the population as estimated by the product of proportion of drug units found in the sample, and the total population size (i.e., PN).

$$\widehat{P}N\left\{\bar{X} - \frac{S}{\sqrt{n}} t_{n-1,1-\alpha/2}\sqrt{\frac{N-n}{N}}\right\} \le W \le \widehat{P}N\left\{\bar{X} + \frac{S}{\sqrt{n}} t_{n-1,1-\alpha/2}\sqrt{\frac{N-n}{N}}\right\} \qquad (3a)$$

Please note that the uncertainty in $\widehat{P}$ is not taken into account (4). In the following we continue in this line of reasoning by ignoring the uncertainty in $\widehat{P}$. The confidence interval is, as a consequence, interpreted conditionally on $\widehat{P}$. Explicitly including the uncertainty in $\widehat{P}$ will result in several problems, as will be discussed below.

Confidence interval 3*a*, however, seems to contain an error. The effect of $\widehat{P}$ is not consistently taken into account at all arguments. As there are $(1-\widehat{P})n$ units without drugs in the sample, the effective sample size on which the estimation of the total weight is based should equal $\widehat{P}n$ instead of n. A similar argument applies to N in the population correction factor which should equal $\widehat{p}N$. Formula 3*a* should, therefore, be rewritten as

$$\widehat{P}N\left\{\bar{X} - \frac{S}{\sqrt{\widehat{P}n}} t_{\widehat{P}n-1,1-\alpha/2}\sqrt{\frac{\widehat{P}N-\widehat{P}n}{\widehat{P}N}}\right\} \le W \le \widehat{P}N\left\{\bar{X} + \frac{S}{\sqrt{\widehat{P}n}} t_{\widehat{P}n-1,1-\alpha/2}\sqrt{\frac{\widehat{P}N-\widehat{P}n}{\widehat{P}N}}\right\}, \qquad (3b)$$

which is equivalent to

$$\widehat{P}N\left\{\bar{X}-\frac{S}{\sqrt{\widehat{P}n}}t_{\widehat{P}n-1,1-\alpha/2}\sqrt{\frac{N-n}{N}}\right\}\leq W \leq \widehat{P}N\left\{\bar{X}+\frac{S}{\sqrt{\widehat{P}n}}t_{\widehat{P}n-1,1-\alpha/2}\sqrt{\frac{N-n}{N}}\right\}. \quad (3c)$$

So, the decrease of the effective sample size $\widehat{P}n$ if $\widehat{P}<1$ should also be taken into account in the standard error, as well as in the number of degrees of freedom. The consequence of this in practice is that both the standard error and the 1-(1/2α) percentile of the *t*-distribution will increase, resulting in a broader, more conservative, confidence interval.

Discussion

In this short paper, we presented an alternative specification for the confidence interval of the total weight of a drug exhibit corrected for the proportion of drug units found in the sample. If one wants to correct, the correction should be applied consistently. In other words, the uncertainty introduced by the fact that the effective sample size is smaller in the presence of units not containing drugs should also be accounted for in the estimation of the standard error, $\frac{S}{\sqrt{\widehat{P}n}}$, and the 1-(1/2$\alpha$) percentile of the *t*-distribution, $t_{\widehat{P}n-1,1-\alpha/2}$.

The fact remains that $\widehat{P}$ is an estimator and therefore stochastic. It is convenient to assume the number of positives known by the time the sampling is perfomed. The interval in formula 3 can then be seen as an approximation of the true interval.

Incorporating the uncertainty in $\widehat{P}$ will result in a complex confidence interval with the degrees of freedom of the *t*-distribution being a random variable. An alternative specification of 3*a* could be to ignore $\widehat{P}$ completely in the interval and base the total weight estimation on all units instead of only the units containing drugs by writing

$$N\left\{\bar{X}-\frac{S}{\sqrt{n}}t_{n-1,1-\alpha/2}\sqrt{\frac{N-n}{N}}\right\}\leq W^* \leq N\left\{\bar{X}+\frac{S}{\sqrt{n}}t_{n-1,1-\alpha/2}\sqrt{\frac{N-n}{N}}\right\}. \quad (4)$$

This can be seen as a confidence interval for W^* not affected by the uncertainty in $\widehat{P}$. Comparison of this confidence interval to the abovementioned intervals, and inclusion of the uncertainty in $\widehat{P}$ will be the topic of future research. The purpose of this paper, however, is only to bring the error in the formula by Tzidony and Ravreby (2) to the attention.

Acknowledgments

We gratefully acknowledge the contribution of Ivo Alberink and Klaas Slooten for their helpful comments.

References

1. Frank RS, Hinkley SW, Hoffman CG. Representative sampling of drug seizures in multiple containers. J Forensic Sci 1991;36:350–7.
2. Tzidony D, Ravreby M. Statistical approach to drug sampling: a case study. J Forensic Sci 1992;37:1541–9.
3. Azoury M, Grader-Sageev D, Avraham S. Evaluation of a sampling procedure for heroin street doses. J Forensic Sci 1998;43:1203–7.
4. Aitken CGG, Lucy D. Estimation of the quantity of a drug in a consignment from measurements on a sample. J Forensic Sci 2002;47:968–75.
5. Aitken CGG, Taroni F. Statistics and the evaluation of evidence for forensic scientists. 2nd ed. Chichester: John Wiley and Sons Ltd., 2004.
6. ENFSI. Guidelines on representative drug sampling, http://www.enfsi.eu/get_doc.php?uid=81 (accessed September 30, 2008).

Additional information and reprint requests:
Reinoud D. Stoel, Ph.D.
Department of Digital Technology and Biometry
Netherlands Forensic Institute
The Hague
The Netherlands
E-mail: r.stoel@nfi.minjus.nl