DRUG DEVELOPMENT AND INDUSTRIAL PHARMACY, 17(10), 1381-1387 (1991)

DETERMINATION OF GEOMETRIC STANDARD DEVIATION FOR DISSOLUTION

N. R. Bohidar

Philadelphia College of Pharmacy and Science
and
Villanova University

ABSTRACT

The two important instances in which the scientist converts his/her experimental data to a logarithmic scale prior to computing the mean and standard deviation, are (i) when the distribution of the data is asymmetrical (e.g. percentage data) and (ii) when he/she intends to compare statistically the averages of two or more groups with unequal standard deviations. In either case, the mean is restored to its original scale by taking the anti-log of the log mean, which is the geometric mean. However, this procedure cannot be applied for computing the geometric standard deviation. The author of reference(1) erroneously claims that the anti-log of log standard deviation is the geometric standard deviation. This paper demonstrates the incorrectness of the procedure in reference (1), exhibits the exact statistical formula and introduces a novel method called "jackknife statistic" to confirm the results, based on the dissolution data associated with Product-C.

INTRODUCTION

The two important reasons for transforming the original scale of measurement of any pharmaceutical experimental data (e.g. dissolution, tablet strength) to a logarithmic scale prior to determining the mean and standard deviation are (i) when the frequency distribution of the response data is asymmetrical (skewed) (2) and (ii) when a statistical test for comparing the means of two or more groups (e.g. formulations) with distinctly unequal standard deviations, is undertaken (2,3). The logarithmic transformation has the unique property of achieving symmetricity of the data and of stabilizing the variances of the groups to be compared (2,3). Since the log mean (LM) and log standard deviation (LSD) are not meaningful for the purpose of interpretation, one computes the anti-log of LM to obtain the geometric mean (GM) [$GM = 10^{\bar{X}}$, where $\bar{X}$ = LM, X = log Y and Y = % dissolution] and computes the following formula to obtain the geometric standard deviation (GSD): $GSD = [(2.303)(GM)(S_X)]$, where S_X = LSD (see derivation next section). However, the author (TK) of reference (1) advocates that the geometric standard deviation should be computed by taking the anti-log of LSD (10^{Sx}) just as one obtains the GM by taking the anti-log of LM ($10^{\bar{X}}$) (1). (In the text this formula will be referred to as "ref(1) formula"). Because this formula is simpler and seems logical, several pharmaceutical scientists and statisticians associated with the pharmaceutical development, production and quality control departments brought the paper to the attention of the author (NRB) to examine the validity of the ref (1) formula. In the course of this paper, it will be shown conclusively that the ref(1) formula is absolutely incorrect. To begin with, consider a sample of 5

dissolution values, 90, 80, 90, 90 and 90, and their respective log values, 1.9542, 1.9031, 1.9542, 1.9542 and 1.9542. The GM and GSD are 87.90 and 4.6311, respectively. However, the ref(1) formula yields 1.054 since LSD = 0.0229. The mean (M) and standard deviation (SD) for the same data without the log transformation are 88.0 and 4.472, respectively. The GSD value is much closer to the SD than the value given by the ref(1) formula. Indeed, if the dissolution values were all 90, then the ref(1) formula would yield a value of 1.0, since LSD = 0.0. Therefore this is a contradiction, and the ref(1) formula is incorrect and should not be used in practice. In the following sections, the derivation of the GSD formula and the results of a dissolution experiment will be presented. In addition, a new and more accurate method of determining the GSD, called the jackknife statistic (JK) will be introduced (2,4,5). Note that, GM(JK) and GSD(JK) for the above data are 87.9 and 4.6716, respectively. This confirms the values of the GM and GSD given above.

THEORY

Let Y_1, Y_2, ---, Y_n be a random sample of n dissolution values and let $X_i = \log_{10} Y_i$ $(i = 1,2---n)$. Then, by definition, $e^{2.303\bar{X}}$ is the geometric mean of the Y_i's, where, $\bar{X} = \Sigma X/n$. The interest here is to derive the variance (V) of $e^{2.303\bar{X}}$, a function of $\bar{X}$. (Note that the square root of V provides the GSD). In general terms (3), let the function be $F(\Theta_1,\Theta_2,--- \Theta_k)$ and $V[F(\Theta)] = (\delta F/\delta\Theta_1)^2 V(\Theta_1) + (\delta F/\delta\Theta_2)^2 V(\Theta_2) + ---- + (\delta F/\delta\Theta_k)2V(\Theta_k)$ + covariance terms,
$(\delta F/\delta\Theta_i)(\delta F/\delta\Theta_j)COV(\Theta_i,\Theta_j)$ $i,j = 1,2,---,k$ $(i \neq j)$.
Note that, covariance terms do not appear in the case being considered.

For the present derivation, $F(\bar{X}) = e^{2.303\bar{x}}$ and

$$V[F(\bar{x})] = [\delta F/\delta \bar{x}]^2 V(\bar{X}) = (2.303)^2[e^{2.303\bar{x}}]^2 V(\bar{X})$$
$$= (2.303)^2[e^{2.303\bar{x}}]^2 S^2_X/n$$
$$= (2.303)^2(GM)^2 S^2_X/n$$

From the above expression, the formula for the GSD is obtained as,

Geometric Standard Deviation = $(2.303)(GM)(S_X)$

Geometric Standard Error = $(2.303)(GM)(S_{\bar{X}})$

where $(S_{\bar{X}}) = S_X/\sqrt{n}$.

This method of derivation is generally known as the Taylor series expansion method, error propagation method or statistical differential method (3,5).

Jackknife Statistic (2,4,5):

The derivation of the variance of a complicated function of the estimate, such as the one for the GM, results in a complex expression. The jackknife statistic not only provides a simpler means to determine the variance but also it has many desirable statistical properties. Because it requires a large number of repetitive computations, one would need to use either a programmable calculator or a desk-top PC. Let X_1, X_2, ---- X_n be a sample with n observations. Let Q(all) denote the estimator of the parameter GM based on all the available data. Let Q(i) denote the estimator based on the sample size of (n-1), where the i^{th} observation has been deleted (skipped). The pseudo values are computed as follows:

$Q^*(i) = n[Q(all)] - (n-1)[Q(i)]$, $i = 1, 2, ---, n$. The jackknife estimate of $GM(JK) = \Sigma Q^*(i)/n$

and the jackknife estimate of

$GSD(JK) = [\Sigma[Q^*(i) - GM(JK)]^2/n-1]^{\frac{1}{2}}$. The added accuracy of the estimate emanates from the fact that one has to compute as many pseudo-values as there are observations in the sample.

TABLE I

GEOMETRIC STANDARD DEVIATION AND GEOMETRIC CONFIDENCE LIMITS FOR PRODUCT-C DISSOLUTION

NO.	STATISTIC	BATCH-P	BATCH-Q	BATCH-R
1	n	10	10	10
2	range	95-99	86-100	83-102
3	GM	97.3868	96.4051	96.6496
4	GSD	1.4354	4.3738	5.9651
5	95% GCL	96.36-98.41	93.28-99.53	92.38-100.92
6	GM(JK)	97.3859	96.3353	96.7114
7	GSD(JK)	1.4430	4.4391	5.9842
8	95% GCL(JK)	96.35-98.42	93.22-99.45	92.43-100.99
9	LSD	0.0064	0.0197	0.0268
10	Ref(1) formula for GSD	1.0150	1.0464	1.0637

GM= Geometric Mean, GSD = Geometric Standard ion, GCL = Geometric Confidence , JK = Jackknife, LSD = Log Standard Deviation

DISSOLUTION STUDY: RESULTS AND DISCUSSION

The dissolution testing was performed on the three selected batches, P, Q and R, of Product-C, using 100 RPM basket method with 900 ml. HCl (1:100) medium at 30 minutes. The statistical analysis was conducted on the spectrophotometric assay values of the 30-minute samples expressed as a percent of label claim. The statistical results are presented in Table I.

A cursory examination of the results indicates that the two methods of determination, statistical differential and jackknife statistic, yielded essentially identical results. The unanimity of the results of the two methods unequivocally demonstrates that the ref(1) formula is erroneous and the values given in the last row of Table I are absolutely meaningless. Table II depicts the detailed computation of the GSD(JK) for Batch-R. The contents are self-explanatory.

TABLE-II

COMPUTATION OF JACKKNIFE STATISTICS FOR BATCH-R

Y	LOG Y	Skip 1	Skip 2	Skip 3	Skip 4	Skip 5	Skip 6	Skip 7	Skip 8	Skip 9	Skip 10
100	2.000	1.9956	2.0000	2.0000	2.0000	2.0000	2.0000	2.0000	2.0000	2.0000	2.0000
99	1.9956	1.9868	1.9868	1.9956	1.9956	1.9956	1.9956	1.9956	1.9956	1.9956	1.9956
97	1.9868	1.9638	1.9638	1.9638	1.9868	1.9868	1.9868	1.9868	1.9868	1.9868	1.9868
92	1.9638	2.0043	2.0043	2.0043	2.0043	1.9638	1.9638	1.9638	1.9638	1.9638	1.9638
101	2.0043	1.9823	1.9823	1.9823	1.9823	1.9823	2.0043	2.0043	2.0043	2.0043	2.0043
96	1.9823	2.0086	2.0086	2.0086	2.0086	2.0086	2.0086	1.9823	1.9823	1.9823	1.9823
102	2.0086	1.9191	1.9191	1.9191	1.9191	1.9191	1.9191	1.9191	2.0086	2.0086	2.0086
83	1.9191	1.9868	1.9868	1.9868	1.9868	1.9868	1.9868	1.9868	1.9868	1.9191	1.9191
97	1.9868	2.0043	2.0043	2.0043	2.0043	2.0043	2.0043	2.0043	2.0043	2.0043	1.9868
101	2.0043										
LM:	1.9852	1.9835	1.9840	1.9850	1.9875	1.9830	1.9855	1.9826	1.9925	1.9850	1.9830
GM:	96.650	96.272	96.383	96.605	97.163	96.161	96.716	96.073	98.288	96.605	96.161
LSD:	0.0268	0.0279	0.0281	0.0284	0.0273	0.0275	0.0284	0.0270	0.0141	0.0284	0.0275
Pseudo Values:*		100.05	99.050	97.050	92.031	101.045	96.048	101.84	81.905	97.050	101.05
GM(JK):		96.711									
GSD(JK)		5.9842									

* Pseudo Value = 10 x G.M. (all) - 9 x (G.M. (i), where i = 1,2,....10

ACKNOWLEDGEMENTS

Thanks are due to Messrs. Frederick J. Mark and Stephen B. Cowdery of Merck and Co. for drawing the author's attention to the Biometrics paper (Ref.1) which prompted the preparation of the present paper for publication. Grateful thanks are due to Mrs. B. Mukerji for carrying out the jackknife statistic computations, promptly and efficiently, on a PC "EVEREX" with Lotus-1-2-3. Special thanks are due to Mrs. B.J. Tomlinson for exhibiting her scientific acumen in accomplishing the task of typing the manuscript in the word processor.

REFERENCES

1. T.B. Kirkwood, Geometric Mean And Measures of Dispersion. Biometrics. Correspondence Section, Vol. 35, No. 4, 908-909, 1979.
2. J. Neter, W. Wasserman and G. Whitmore, "Applied Statistics." Allyn and Bacon, Inc. Chapter 13, pg. 338 and Chapter 15, pg. 399, 1982.
3. S.L. Meyer, "Data Analysis For Scientists and Engineers." John Wiley and Sons, N.Y. Chapter 10, pg. 39, 1975.
4. F. Mosteller and J.W. Tukey, "Data analysis and Regression." Addison-Wesley, Inc. Chapter 8, 1977.
5. N.R. Bohidar, Jackknife Approach to Estimation of Apparent Half-life and its Confidence Limits. Proceedings American Statistical Association, Biopharmaceutical Section, pg. 94-95, 1986.