Correct and Incorrect Estimation of Within-Day and Between-Day Variation

Michael J. Bookbinder and Kim J. Panosian

Between-day variance is an ambiguous term representing either total variance or pure between-day variance. In either case, it is often incorrectly calculated even though analysis of variance (ANOVA) and other excellent methods of estimation are available. We used statistical theory to predict the magnitude of error expected from using several intuitive approaches to estimation of variance components. We also evaluated the impact of estimating the total population variance instead of pure between-day variance and the impact of using biased estimators. We found that estimates of variance components could be systematically biased by several hundred percent. On the basis of these results, we make recommendations to remove these biases and to standardize precision estimates.

Additional Keyphrases: quality control - statistics - analysis of variance

Although analytical precision is usually evaluated (1) in terms of within-day and between-day variance, there is concern about how accurately these components are estimated (2-4). Should between-day imprecision include within-day imprecision, thus representing total imprecision, or should it represent only pure between-day imprecision (2, 3)? Between-day (or day-to-day) variance sometimes refers to pure between-day variance, especially when used with the term "component" (3, 5, 6); other times, it refers to total variance (7, 8). This ambiguity in definition affects the accuracy of the determination.

In addition, variance components are often estimated by intuitive (3) or biased methodologies instead of by accurate methods such as analysis of variance (ANOVA). When a variance is to be estimated, N aliquots are measured. The mean of these N readings is obtained, and the sum of the squared deviations from this mean is calculated. Commonly, the estimated variance is erroneously computed by dividing this sum by N, introducing bias into the estimate (9). The correct, unbiased estimate of the variance is obtained by dividing the sum by N - 1 (6, 7, 10-12).

We therefore decided to quantify the magnitude of the error inherent in using biased and intuitive approaches to variance component estimation. We also undertook to demonstrate the numerical difference between "total" and "pure between-day" variance.

Methods

The following questionable methods (3), numbered arbitrarily, were evaluated as to their ability to estimate variance components.

I. Methods Estimating $\sigma^2_{WD}$ and $\sigma^2_{TOTAL}$

A. Methods with $\sigma^2_{TOTAL}$ (SAMPLE) as the estimate
1. With bias. Method IA1: In this method the total biased sample variance (with a denominator of N instead of N - 1) is used to estimate the between-day variance ($N = S \times D$). We also examined the effect of using this method to estimate the total population variance ($\sigma^2_{WD} + \sigma^2_{BD}$).
2. Without bias. Method IA2: This method is similar to Method IA1, but the total unbiased sample variance is used to estimate the between-day variance. Division is by N - 1 instead of by N. We also examined the effect of using this method to estimate total population variance.

B. Other methods
Method IB: In this method, multiple aliquots are measured each day for D days. A daily average is calculated on each day's measurements. The between-day variance is estimated as the variance of these D averages. This variance estimate may be calculated in one of two ways—by dividing either by D - 1 (unbiased) or by D (biased).

II. Methods Estimating $\sigma^2_{WD}$

A. Method IIA: This method estimates $\sigma^2_{WD}$ by using S samples measured in one day. However, when the biased sample variance is used as an estimate of $\sigma^2_{WD}$, we have:

Estimated $\sigma^2_{WD} = \Sigma [(X_i - \overline{X})^2/S]$ (incorrect)

instead of:

Estimated $\sigma^2_{WD} = \Sigma [(X_i - \overline{X})^2/(S - 1)]$ (correct)
where $X_i$ is an individual measurement and $\bar{X}$ is the average of the $X_i$ values.

B. Method IIB: In this method, multiple aliquots are measured each day, and the variance (with bias) is calculated for each day. $\sigma_{BD}^2$ is estimated as the average of these $D$ variances.

For each of the above questionable methods of estimating $\sigma_{BD}^2$ and $\sigma_{ED}^2$, we calculated the expected error, using equations derived from the formulas for expected values of mean sum of squares (13). We calculated such expected or systematic errors for various values for $D, S, D, S, and R$. These equations and a method for performing ANOVA are given in the Appendix.

Results and Discussion

Of all the methodologies, only Method IA2 produced an unbiased estimate of $\sigma_{\text{TOTAL}}^2$ (see Appendix), but then only when there was one sample per day ($S = 1$). The expected inaccuracy or bias in using the other intuitive approaches is displayed in Figures 1 through 4 for selected combinations of $S, D, and R$. Other combinations of $S, D, and R$ yielded various other degrees of error.

We emphasize that when there is more than one sample per day, only ANOVA produces unbiased estimates of variance. However, other similar methods of variance-component estimation possess comparable accuracy (5).

A few comments about the intuitive methodologies follow:

Method IA1 is a common way of estimating $\sigma_{BD}^2$. The total biased sample variance is used as the estimate. The degree of error in this method (Figure 1) is variable, making the estimates from this method difficult to interpret. The degree of error depends on the values of $S, D, and R$. Method IA1 is inaccurate because there is division by $N$ instead of by $N - 1$ and because the total sample variance is used to estimate the pure between-day variance.

Method IA1 may also be used to estimate the total population variance. However, the method remains biased (Figure 2, lower curve, and Figure 3), even in the case of one sample per day.

Method IA2 overcomes this problem by dividing the total sum of squares by $N - 1$ instead of by $N$. However, as an estimate of $\sigma_{BD}^2$ Method IA2 produced large errors (Table 1) and did so variably. Pure between-day precision of assays calculated by Method IA2 would not be comparable between researchers except under special circumstances.

Table 1. Amount of Overstatement of Actual Between-Day Variance When Estimated by Method IA2

<table>
<thead>
<tr>
<th>$R$</th>
<th>$S$</th>
<th>$D = 3$</th>
<th>$D = 15$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>1</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>-7</td>
<td>19</td>
</tr>
<tr>
<td>1.00</td>
<td>1</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>75</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>68</td>
<td>94</td>
</tr>
<tr>
<td>4.00</td>
<td>1</td>
<td>400</td>
<td>400</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>375</td>
<td>395</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>368</td>
<td>394</td>
</tr>
</tbody>
</table>

*Systematic error, see text.
$R = \frac{\sigma_{BD}^2}{\sigma_{BD}^2}, S = \text{no. of samples per day, } D = \text{no. of days.}

Fig. 1. Systematic error from using Method IA1 (total unbiased sample variance) to estimate $\sigma_{BD}^2$ when $S = 3$

Fig. 2. Systematic error from using Method IA1 (total biased sample variance) to estimate variance when $S = 1$ and $R = 0.25$

The upper curve is the systematic error in estimating $\sigma_{BD}^2$ and the lower curve is the error in estimating $\sigma_{\text{TOTAL}}^2$.

Fig. 3. Systematic error from using Method IA1 (total unbiased sample variance) to estimate $\sigma_{\text{TOTAL}}^2$ when $S = 3$

The error curves for Method IA2 (total unbiased sample variance) in estimating $\sigma_{\text{TOTAL}}^2$ fall a little higher. The error curves for Method IA2 in estimating $\sigma_{BD}^2$ are similar to the lower curve, letting the abscissae represent $S$.
The correct approach is to use the total unbiased sample variance (Method IA2) to estimate total variance. In the special case of $S = 1$, we found that Method IA2 is accurate in estimating total population variance. However, this method will not generally be accurate if $S > 1$. The error will be greater if $s^2_{BD}$ is relatively large (Figure 3).

Method IB estimates the between-day variance as the variance among the daily means. This approach is more sophisticated than that of the Method IA group for estimating $s^2_{BD}$ when there are multiple aliquots per day. When the variance (of the daily means) is obtained by dividing by D instead of by $D - 1$, this method produces biased estimates that will be too small or too large, depending on the particular values of D, S, and R (Figure 4). Dividing by $D - 1$ instead of by D will produce estimates with better accuracy, particularly when either S is large or R is small. Apparently, averaging removes enough of the within-day component if such a component is small or if there are many samples per day.

Method IIA is the method of estimating $s^2_{BD}$ by measuring one sample per day, and then using the sample variance (dividing by D) as the estimate. This method yields the usual $D/(D - 1)$ bias, which may be acceptable if D is large. If all researchers used the same number of specimens and all used this biased method, precision estimates would still be comparable, but such a situation is unlikely.

Method IIB estimates $s^2_{BD}$ as the average of multiple within-day variances, where each day’s variance is biased. The end-result is thus biased. Note that there may be extreme bias despite the presence of many data. If $S = 3$, then the systematic error is about $-33\%$, no matter how many days are involved.

Whether corrected for bias or not, Method IIB estimates $s^2_{BD}$ with greater precision than does Method IIA because Method IIB is based on a multi-day average rather than data from only one day.

Two precautions are in order. First, there is concern (3) that total variance, when reported as being less than within-day variance, may represent some other quantity. Sometimes between-day variance, instead of referring to total variance, apparently refers to pure between-day variance. Although "between-day variance" usually refers to total variance ($s^2_{BD} + s^2_{DB}$) in assay evaluations (7, 8), the same phrase has also been used to refer to pure between-day variance in control chart applications (14-17). Ambiguity in the meaning of between-day variance creates difficulty in comparing estimates of precision.

Second, even though it may be recognized that the total unbiased sample variance when $S = 1$ is a valid estimator of total population variance, the concept of "total variance" must be used with caution. The researcher must not conclude that total sample variance is always an unbiased estimator of total population variance ($s^2_{BD} + s^2_{DB}$). If $S > 1$, then the total unbiased sample variance is not, in general, an accurate estimator of the total population variance.

In summary, all variance estimates should be computed without bias by dividing by $N - 1$ unless N is large. To estimate components, ANOVA or the methods listed below should be used. The divisor and method should be clearly indicated, to avoid confusion and ensure comparability of results. In light of recent considerations (2, 3), inclusion of methodological detail is critical if researchers use methods different from those used previously. An accurate method to estimate $s^2_{BD}$ would be Method IIA modified for division by $S - 1$. After $s^2_{BD}$ is so obtained, Method IA2, for one sample per day, may be used to estimate total population variance. $s^2_{BD}$ may be estimated by subtracting $s^2_{DB}$ from $s^2_{TOTAL}$. Other accurate methods are also available (6, 18).

The appropriate phrase, "pure between-day" or "total" variance, should be used. If total variance is used, the writer should make clear whether this phrase refers to the population estimate or the total sample variance, because the latter may be quite different from the former. In addition, even when the estimates are unbiased, the researcher must avoid using too few specimens (II, 19) and must handle outliers adequately (II).

References

Fig. 4. Systematic error from using the variance of the daily averages to estimate $s^2_{BD}$ when $S = 3$

The solid lines are the error from using the biased variance of the daily averages as the estimate (Method IB). The dotted lines represent the error from using the unbiased variance of the daily averages to estimate $s^2_{BD}$ (Method IB modified for division by $D - 1$ instead of by D). R = 4.0; O, R = 1.0; A, R = 0.25

1730 CLINICAL CHEMISTRY, Vol. 32, No. 9, 1986

Appendix

The expected values of the methods were derived from the equations for the expected value of mean squares (13):

<table>
<thead>
<tr>
<th>Method used</th>
<th>Actual value estimated</th>
<th>Expected value of the method</th>
</tr>
</thead>
<tbody>
<tr>
<td>IA1</td>
<td>σ_BD or σ_TOT</td>
<td>(D - 1) × σ_BD/D + (N - 1) × σ_WD/N</td>
</tr>
<tr>
<td>IA2</td>
<td>σ_BD or σ_TOT</td>
<td>[(D - 1) × S × σ_BD^2/(N - 1)] + σ_WD</td>
</tr>
<tr>
<td>IB (unbiased)</td>
<td>σ_BD</td>
<td>σ_BD + σ_WD/S</td>
</tr>
</tbody>
</table>

IB (biased) \[ \sigma_B = (D - 1) \times (\sigma_B + \sigma_WD) \]

IA, B \[ \sigma_W = (S - 1) \times \sigma_WD \]

where \( S \) = samples per day, \( D \) = number of days, and \( N = S \times D \). Note that when \( S = 1 \), then \( N = D \), and the expected value of using Method IA2 becomes \( \sigma_B + \sigma_W = \sigma_{TOTAL} \).

The systematic error (SE) for each of the methods was calculated by using the expected value (above) as follows:

\[
SE = \frac{\text{(expected value) - (actual value)}}{\text{(actual value)}} \times 100\%
\]

where "actual value" is the quantity being estimated as indicated above.

Estimation of variance components by using analysis of variance. Let the measurement of sample number "i" on day "j" be called \( X_{ij} \). First calculate the "sum of squares within-days" (SSWD) and the "sum of squares between-days" (SSBD):

\[
SSWD = \sum \sum (X_{ij})^2 - \sum (X_{ij})^2/S
\]

\[
SSBD = \sum (X_{ij})^2/S - (\sum X_{ij})^2/(S \times D)
\]

where:

\[ X_{ij} = \sum X_{ij} \text{ summed over all } j \text{, and } X_{..} = \sum X_{ij} \]

Next, calculate the mean squares, both within-day (MSQWD) and between-day (MSQBD):

\[
MSQWD = SSWD/(D \times (S - 1))
\]

\[
MSQBD = SSBD/(D - 1)
\]

Finally, calculate the correctly estimated variances components:

\[
\text{Estimated } \sigma_W^D = MSQWD
\]

\[
\text{Estimated } \sigma_B^D = (MSQBD - MSQWD)/S
\]