A Two-Component Model for Measurement Error in Analytical Chemistry

David M. ROCKE
Graduate School of Management
University of California
Davis, CA 95616

Stefan LORENZATO
Division of Water Quality
California State Water Resources
Control Board
Sacramento, CA 94244-2130

In this article, we propose and test a new model for measurement error in analytical chemistry. Often, the standard deviation of analytical errors is assumed to increase proportionally to the concentration of the analyte, a model that cannot be used for very low concentrations. For near-zero amounts, the standard deviation is often assumed constant, which does not apply to larger quantities. Neither model applies across the full range of concentrations of an analyte. By positing two error components, one additive and one multiplicative, we obtain a model that exhibits sensible behavior at both low and high concentration levels. We use maximum likelihood estimation and apply the technique to toluene by gas-chromatography/mass-spectrometry and cadmium by atomic absorption spectroscopy.

KEY WORDS: Atomic absorption spectroscopy (AAS); Coefficient of variation; Detection limit; Gas-chromatography/mass-spectrometry (GC/MS); Maximum likelihood; Quantitation level.

1. INTRODUCTION: MEASUREMENT NEAR THE DETECTION LIMIT

Traditionally, the description of the precision of an analytical method is accomplished by applying two separate models. One for describing zero and near-zero concentrations of the analyte (the compound that the analytical method is designed to measure) and another for quantifiable amounts. This traditional approach leaves a "gray area," of analytical responses in which the precision of the measurements cannot be determined. The model governing quantifiable amounts assumes that the likely size of the analytical error is proportional to the concentration. If this model is applied to analytical responses in the gray area, then there is an implicit assumption that the analytical error becomes vanishingly small as the measurements approach 0. From long experience, this assumption appears to be invalid. Similarly, if the zero-quantity model is applied, there is an implicit assumption that the absolute size of the analytical error is unrelated to the amount of material being measured. Based on similar empirical information, this assumption also cannot be supported.

The new model presented in this article resolves these difficulties by providing an estimate of analytical precision that varies between the two extremes described by the traditional models. The model provides a distinct advantage over existing methods by describing the precision of measurements across the entire usable range. Examples are given of two different analytical methods, an atomic absorption spectroscopy analysis for cadmium and a gas-chromatography/mass-spectrometry analysis for toluene, both of which support the validity of the new model. The new model is applicable to a wide variety of situations including nonlinear calibration and added-standards calibration. Discussion is provided on the application of the new model to some common issues such as determination of detection limits, characterization of single samples, and determination of sample size required for inference to given tolerances.

Many measurement technologies have errors whose size is roughly proportional to the concentration—this is often true over wide ranges of concentration (Caulcutt and Boddy 1983). One common way to describe this constant coefficient of variation (CV) model is that the measured concentration \( x \) is given by

\[
\log(x) = \log(\mu) + \eta \quad (1.1)
\]

or

\[
x = \mu e^{\eta}, \quad (1.2)
\]

where \( \mu \) is the true concentration and \( \eta \) is a normally distributed analytical error with mean 0 and standard deviation \( \sigma_\eta \). This model is widely used, but it fails to make sense for very low concentrations because it implies absolute errors of vanishingly small size.

On the other hand, one often considers the case of analyzing blanks (samples of zero concentration) and data near the detection limit by the model

\[
x = \mu + \epsilon \quad (1.3)
\]

with normally distributed analytical error \( \epsilon \). This is used for calibration with \( \mu = 0 \) and for the determination of
A MODEL FOR MEASUREMENT ERROR IN ANALYTICAL CHEMISTRY

177

detection limits (Massart, Vandeginste, Deming, Michotte, and Kaufman 1988). It is, however, a bad approximation over wider ranges of concentration because it implies that the absolute size of the error does not increase with the concentration, and this is rarely true of analytical methods.

The solution proposed here is a combined model that reflects two types of errors. For example, in a gas-chromatography/mass-spectrometry (GC/MS) analysis one type of error is in the generation and measurement of peak area, which will generally have errors of size proportional to the concentration. The other type of error comes from the fact that, even when fed with a blank sample, the output is not a flat line but still retains some small variation. There are many sources of error in such an analysis; the idea here is merely to classify them into two types, additive and multiplicative. The model proposed is

\[ x = \mu e^\eta + \epsilon, \]  

where there are two analytical errors, \( \eta \sim N(0, \sigma_e^2) \) and \( \epsilon \sim N(0, \sigma_\epsilon^2) \), each normally distributed with mean 0. Here \( \eta \) represents the proportional error that is exhibited at (relatively) high concentrations and \( \epsilon \) represents the additive error that is shown primarily at small concentrations. Another way of writing this model is via a linear calibration curve as

\[ y = \alpha + \beta \mu e^\eta + \epsilon, \]  

where \( y \) is the observed measurement (such as peak area at a GC). Note that the new model approximates a constant standard deviation model for very low concentrations and approximates a constant CV model for high concentrations.

The approach proposed here should be contrasted with an alternative method, which is to model the standard deviation as a linear function of the mean concentration. The latter approach should work well over restricted ranges but suffers from some disadvantages when used over a wide range. First, the predicted standard deviation at zero concentration need bear little resemblance to the measured value when one regresses the standard deviation of replicates on the mean of the replicates. The proposed new model allows determination of the error structure of an analytical method near the detection limit and so provides an easy way to give precisions for such measurements, as well as the average of several such.

2. ESTIMATION

The approach we used is based on maximum likelihood estimation. An observed value \( x \) differs from the theoretical value \( \mu \) because of the two errors \( \eta \) and \( \epsilon \), which are not directly observed. Any combination of \( \eta \) and \( \epsilon \) that satisfies \( \epsilon = x - \mu e^\eta \) is possible. Consequently, the likelihood associated with a set \((\sigma_\eta, \sigma_\epsilon)\) of parameters given a set of \( n \) measurements \( x_i \) with known concentrations \( \mu_i \) is

\[ L = \prod_{i=1}^{n} \int_{-\infty}^{\infty} \frac{1}{2\pi \sigma_\eta \sigma_\epsilon} e^{-\eta^2/(2\sigma_\eta^2)} e^{-(x_i - \mu_i e^\eta)/(2\sigma_\epsilon^2)} d\eta. \]  

Maximizing this likelihood leads to estimates of the necessary parameters \( \sigma_\eta \) and \( \sigma_\epsilon \). More complex models, such as the estimation of a calibration curve, can be estimated in the same fashion, using maximum likelihood. For example, the calibration model (1.5) has likelihood

\[ L = \prod_{i=1}^{n} \int_{-\infty}^{\infty} \frac{1}{2\pi \sigma_\eta \sigma_\epsilon} e^{-\eta^2/(2\sigma_\eta^2)} e^{-(x_i - \mu_i e^\eta)/(2\sigma_\epsilon^2)} d\eta. \]  

Once estimates \( \sigma_\eta \) and \( \sigma_\epsilon \) have been derived, the precision of any measured value in the form (1.4) is (using the
Appendix B. From the usual maximum likelihood estimation asymptotic variances given by the negative inverse of the
covariance matrix (2.3).

If the covariance matrix \( \sigma^2 \) is V, then a straightforward
delta-method calculation shows that the variance of \( \mu \) is
\[
\text{var}(\mu) = \beta^2 + \mu^2 e^{\gamma^2} (e^{\gamma^2} - 1) \tag{2.4}
\]
where \( g = \beta^{-1} - \mu \beta^{-1} \). Note that, when sample sizes
are sufficiently large that \( \alpha \) and \( \beta \) can be regarded as
known, the \( g^2 V g \) term disappears, and this basically coincides with (2.3).

The computational approach used is outlined in
Appendix B. From the usual maximum likelihood estimation theory, the estimates of the parameters are normal with
asymptotic variances given by the negative inverse of the
information matrix, which we use for the matrix V mentioned previously when needed.

2.1 Elaborations of the Basic Model

The model we are developing will be of no practical use
unless it accurately describes the behavior of a wide range of
analytical data. If a test data set consists of replicated measurements at a variety of levels from 0 to (say) 500
ppb, then the precision at 0 will be a, and the CV at 500
ppb, then the precision at 0 will be a, and the CV at 500

A third potential elaboration is needed to address the
case in which a labeled standard is used to calibrate measured values, which adjusts for recovery efficiency. Consider
an analytical method to determine the concentration of a volatile organic such as toluene. One method of increasing accuracy is to spike the sample with a known concentration \( v \) of deuterated toluene and determine the estimated concentration \( \mu \) of toluene and \( \mu_d \) of deuterated toluene. One estimate of the true concentration of toluene is \( \mu_{\text{adj}} = \mu v/\mu_d \), which adjusts for recovery. If
\[
y = \alpha + \beta v + \epsilon \tag{2.7}
\]
and
\[
y_d = \alpha + \beta v + \epsilon_2 \tag{2.8}
\]
then the precision of \( \mu_{\text{adj}} \) depends on the precision of \( \mu \), the precision of \( \mu_d \), and their covariance. Using the delta
method we derive an approximate variance for \( \mu_{\text{adj}} \) as
\[
\text{var}(\mu_{\text{adj}}) \approx \sigma^2 \mu/v^2 + \sigma^2 v^2 \mu_d^2 /\mu_d^2 - 2 \sigma \mu v \mu_d /\mu_d^2.
\]

where \( \sigma^2, \sigma \), and \( \sigma \mu \), are derived from a multivariate
version of (2.5). In this case, it is necessary to have sufficient
data to estimate the covariances of the errors in (2.7) and (2.8).

3. APPLICATIONS

In this section we describe some ways that the new
model can be used. We concentrate especially on applications in
environmental monitoring, where detection and
measurement of low levels of toxic substances may be
quite important. Generally, these applications assume that
the parameters of the model have been determined, so we
will generally describe the ideas in terms of the simple
model (1.4). Extensions to the case in which calibration
error is also to be accounted for are straightforward.

3.1 Detection Limits

According to the preceding model, the observations at
true concentration \( \mu = 0 \) are normally distributed with
standard deviation \( \sigma_e \). If \( r \) replicates are used, then any
average of measured values greater than \( D - 3\sigma_e/\sqrt{r} \) is
extremely unlikely to have come from a zero concentration
sample. Use of the exact method of setting confidence
intervals described in Section 3.2 allows the precise
determination of the uncertainty. Of course, this assumes that the
replicates are true reruns of the entire process; other-
wise the error may not be reduced by a factor of \( \sqrt{r} \) but
by a much smaller amount.

An implication of this rule for environmental moni-
toring is that the accumulation of measurements at low
levels, even individually below the individual observation
detection limit can still provide quantitative evidence of the
concentration of a toxic substance. If the safe level is
near or below the detection limit, then it might make sense to
require replicate measurements (to reduce the effective

TECHNOMETRICS, MAY 1995, VOL. 37, NO. 2
3.2 Uncertainty of a Single Measurement

There are two primary approaches to this problem, an exact solution and a normal or lognormal approximation. In the exact solution, we construct a confidence interval for the true concentration by locating the points on either side of the measured value where a hypothesis test just rejects. If a measurement $x$ is obtained, and a 95% confidence interval is desired, then we need to find values $\mu_L$ and $\mu_U$ such that

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/(2\sigma^2)} (1 - \Phi(x - \mu_L e^u)/\sigma_x) dx = .025$$

and

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/(2\sigma^2)} (\Phi(x - \mu_U e^v)/\sigma_v) dx = .975,$$

where $\Phi()$ is the standard normal distribution function. This can be done by numerical solution of these equations, which yields a 95% confidence interval ($\mu_L$, $\mu_U$).

An approximate method is based on the estimated variance of $x$ given by

$$V(x) = \sigma_x^2 + x^2 e^{2\sigma^2} (e^{\sigma^2} - 1).$$

For low levels of $x$ (those in which the first term dominates), the distribution of $x$ is approximately normal. An approximate 95% confidence interval is formed as

$$x \pm 1.96\sqrt{V(x)}.$$

For high levels of $x$ (those in which the second term in $V(x)$ dominates) $\ln(x)$ is approximately normally distributed with variance $\sigma^2$ so that a 95% confidence interval for $\mu$ is

$$(\exp(\ln(x) - 1.96\sigma), \exp(\ln(x) + 1.96\sigma)).$$

Note that this interval, although symmetric on the logarithmic scale, is asymmetric on the original measurement scale.

3.3 Uncertainty of an Average of Several Measurements

Here, the exact method of confidence-interval determination would be burdensome because it would require the computation of a difficult convolution. So confidence intervals must be based on the approximate normality of $\bar{x}$ (for low levels) or of $\ln(x)$ (for high levels). For low levels, the average, $\bar{x}$, of $n$ measurements will be approximately normally distributed with standard deviation $\sqrt{V(\bar{x})/n}$. For larger values of $x$, it will be better to perform inference using the logarithms of the data, which will be more nearly normally distributed. The average

$$\sqrt{[\sigma^2 + \mu^2 e^{2\sigma^2} (e^{\sigma^2} - 1)]/r} = \sqrt{[.04 + (.09)e^{.01}(e^{.01} - 1)]/r}$$

and

$$= .0409/r.$$
For the chance of detection to be .95, we require \( \frac{r}{\sqrt{r}} > 1.645 \) or \( r > 2.77 \). In this case, the number of replicates should be at least 3.

4. EXAMPLES

In this section, we present two examples of analytical methods and examine the fit of the two-component model. The first example consists of graphite-furnace atomic absorption spectroscopy (AAS) measurements of cadmium, for known concentrations from 0 to 43 ppb, each quadruply replicated. The second example is GC/MS measurements of toluene from 4.6 picograms to 15 nanograms. For descriptions of these methods see Willard, Merritt, Dean, and Settle (1988). In both cases we use the calibration form (1.5) with likelihood (2.2).

4.1 Cadmium by Atomic Absorption Spectroscopy

The instrument used for these measurements was a Perkin-Elmer model 5500 graphite furnace atomic absorption spectrometer with a model HGA-500 furnace controller. The tubes were pyrolytically coated; injections were made onto L'vov platforms. A deuterium arc lamp was used for background correction. Table 1 shows the AAS measurements, and Table 2 gives the standard deviation of the replicate measurements of absorbance and the standard deviation of the logarithms of the replicates. Note that the measurements at the two lowest concentrations appear to have a constant standard deviation of the measured concentration and the measurements at the four highest concentrations appear to have a constant standard deviation of the measured log-concentration. Therefore, the data span the “gray area” of quantification. Starting values for the maximum likelihood estimator (MLE) procedure, which in this case are themselves quite good...
Table 4. Toluene Amounts by GC/MS

<table>
<thead>
<tr>
<th>Toluene amount (pg)</th>
<th>Peak area (pg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.6</td>
<td>29.80</td>
</tr>
<tr>
<td>23</td>
<td>44.60</td>
</tr>
<tr>
<td>116</td>
<td>207.70</td>
</tr>
<tr>
<td>580</td>
<td>894.67</td>
</tr>
<tr>
<td>3,000</td>
<td>5350.65</td>
</tr>
<tr>
<td>15,000</td>
<td>20718.14</td>
</tr>
</tbody>
</table>

Table 6. Initial Estimates From Pooled Data and Final Parameters by MLE for Toluene Data

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Starting value</th>
<th>Final value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>-1.6</td>
<td>11.51</td>
</tr>
<tr>
<td>$\beta$</td>
<td>1.546</td>
<td>1.524</td>
</tr>
<tr>
<td>$\sigma_a$</td>
<td>0.10</td>
<td>0.1032</td>
</tr>
<tr>
<td>$\sigma_r$</td>
<td>6.0</td>
<td>5.698</td>
</tr>
</tbody>
</table>

estimates of the parameters, are shown in Table 3. Starting values for the regression parameters $\alpha$ and $\beta$ were derived by ordinary linear regression of the absorbance measurements on the concentration. The starting value for the standard deviation $\sigma_a$ of zero measurements was estimated from the two lowest concentrations because there appears not to be an upward trend until the third highest concentration. A starting value for the standard deviation of the log absorbances for high concentrations $\sigma_r$, which is approximately the coefficient of variation, was derived by pooling the highest four levels, where the CV appears constant. These initial estimates are of such high quality that they are scarcely changed by the MLE iterations. Final values are also shown in Table 3. Note that the calculation is not extremely sensitive to the starting value because the same optimum was arrived at from $\alpha = 0$, $\beta = 2$, $\sigma_a = .03$, and $\sigma_r = .4$, for example. Somewhat plausible values must be used, however, to avoid numerical instability.

The results are shown graphically in Figure 1, which plots the data, the calibration curve, and an estimated envelope; Figure 2, which shows the actual and predicted standard deviation; and Figure 3, which shows the residuals from the model fit.

If we now treat these estimated parameters as known, we can use the model for further analysis. For example, if the limit of detection is defined as the absorbance (and associated implied concentration) that falls three standard deviations above 0, we find that it lies at about .4 ppb.

Confidence intervals for concentration can be derived using the methods of Section 3.2. For example, with the parameters estimated for the cadmium data, an absorbance of 6 implies an estimated concentration of 2.75 ppb, with 95% confidence interval (2.47, 3.04). This is almost symmetric, with half widths of .28 and .29. An absorbance of 50 implies an estimated concentration of 21.76 ppb, with 95% confidence interval (20.69, 22.88) and half-widths 1.07 and 1.12. A normal approximation using Equation (2.3) yields confidence intervals of (2.49, 3.01) for an absorbance of 6 and (21.23, 22.29) for an absorbance of 50. The approximation is quite acceptable for the low absorbance but not for the higher one. For the absorbance of 50, an acceptable approximation is gained by using the approximate lognormality of the measurements at high levels, leading to a 95% confidence interval of (20.72, 22.85). Note, however, that there is no satisfactory alternative to the exact confidence interval provided by the new model that is accurate over the entire range of measurements.

4.2 Toluene by GC/MS

The instrument used here was a Trio-2 GC/MS (VG Masslab) with electron ionization at 70 electron volts with a 30-meter DB-5 GC column. Table 4 shows an analysis of amount of toluene by GC/MS for known amounts of from 4.6 picograms to 15 nanograms in 100 PL of extract. (1 picogram in 100 PL corresponds to a concentration of .01 ppb.) The quantitation is done by peak area.
DAVID M. ROCKE AND STEFAN LORENZATO

Figure 5. Actual and Predicted Standard Deviation for Toluene by GC/MS. The dots represent the standard deviation of the replicates and the solid line is the predicted standard deviation using (3.4) and (3.5) as appropriate. Note the log-log scale.

The relationship between amount and peak area is satisfactorily linear and the behavior of the errors is generally consistent with the new model as shown in Table 5. Starting values were derived from ordinary linear regression and from an examination of the standard deviations of the raw and logged data. The MLE estimation gives apparently acceptable results as shown in Table 6.

The results are shown graphically in Figure 4, which plots the data, the calibration curve, and an estimated envelope; Figure 5, which shows the actual and predicted standard deviation; and Figure 6, which shows the residuals from the model fit. These plots are all on a log-log scale.

A comparison is given in Table 7 between the predicted standard deviation of the peak area from the model developed in this article and the predicted standard deviation from a variance function in which the standard deviation is a linear function of the mean. Both fit well for high concentrations, but the linear function is quite inaccurate for lower concentrations. For these large ranges, it appears that the new model fits the behavior of the data better.

5. CONCLUSION

In this article we have provided a new model for analytical error that behaves like a constant standard deviation model at low concentrations and like a constant CV model at high concentrations. The importance of this new model is that it provides a reliable way to estimate the precision of measurements that are near the detection limit so that they can be used in inference and regulation. Although the illustrations used in the article are for linear calibration analyses, the model itself is flexible enough to be used in a wide variety of situations, including nonlinear calibration, censored data, and added-standards methods.

ACKNOWLEDGMENTS

We are grateful to A. Russell Flegal, Institute of Marine Sciences, University of California, Santa Cruz, and A. Daniel Jones, Facility for Advanced Instrumentation, University of California, Davis, for providing the data used as examples in this article, and to the editor, an associate editor, and two referees for suggestions that substantially improved the article. The research reported in this article was supported in part by the California State Water Resources Control Board, by National Institute of Environmental Health Sciences, National Institutes of Health Grant P42 ES04699, and by National Science Foundation Grant DMS-93.01344.

APPENDIX A: THE HORWITZ TRUMPET

Horwitz (1982; Horwitz et al. 1980) examined over 150 independent Association of Official Analytical Chemists interlaboratory collaborative studies covering numerous
topics from drug preparations and pesticide formulations (at the high concentration end) to aflatoxin (at the low end). Over these studies, there was a pattern to the inter-laboratory CV that Hall and Selinger (1989) described as the Horwitz trumpet. The average CV of the interlaboratory errors increased from a few percent for concentrations in the double-digit percentage range, to about 50% for concentrations in the 1 ppb range. Horwitz suggested an empirical formula relating the concentration \( C \) to the coefficient of variation \( \text{CV}(C) \) [and therefore the precision \( \sigma(C) \)] as follows:

\[
\text{CV}(C) = 0.006C^{-0.5} \quad (A.1)
\]

and

\[
\sigma(C) = 0.006C^{0.5}. \quad (A.2)
\]

After an analysis of Horwitz's data using a binomial model with variable apparent sample size, Hall and Selinger suggested a similar formula:

\[
\text{CV}(C) = 0.02C^{-1.15} \quad (A.3)
\]

and

\[
\sigma(C) = 0.02C^{0.85}. \quad (A.4)
\]

Both the Horwitz trumpet and the present article suggest an increase in the CV at low concentrations; however, there are several important differences. First, Horwitz was interested in the variation from laboratory to laboratory, whereas the model in this article is intended for intralaboratory precision (there is no reason why it should not apply also across laboratories, however). Second, and more important, Horwitz was describing how the precision changes when one moves from one analytical method intended for a certain range of concentrations to another method intended for a different range. The two-component error model is for a single analytical method across its useful range. There is no reason why one model cannot be used within methods and a different one between methods. In the language of this article, Horwitz had a model for how \( \sigma(C) \) changes with the analytical method, but the two-component model describes how the precision changes within a given method as the concentration approaches the detection limit of that method.

It is of particular note that the Horwitz trumpet cannot be used to serve the purpose of this article in describing the transition of the error structure in a given analytical method as the concentration changes from high levels to those near the detection limit. This is because the Horwitz model in both its original form and in Hall and Selinger's emendation imply that the standard deviation of a measurement at low levels approaches 0. In particular, if used inappropriately to describe the interlaboratory error under zero concentration, this model would imply zero error. In fact, most laboratories might report the compound as "not detected," but this is a far cry from zero error.

In summary, both the Horwitz trumpet and the two-component error model have a proper place in understanding the errors of analytical method. The former is useful to describe the way in which the size of the errors from analytical methods change as the method itself and the intended range of concentration changes. The latter, the focus of this article, allows better estimation of quantitation error for a given analytical method for measurements within a few multiples of the detection limit.

**APPENDIX B: COMPUTATIONAL ISSUES IN ESTIMATION**

Actually performing the maximization of the likelihood requires the numerical optimization of a numerically integrated likelihood. For the numerical integration, we used a Gauss–Hermite rule with centering depending on a second-order Taylor approximation to the logarithm of the integrand. This led to highly accurate results. Specifically, we need to approximate

\[
\int_{-\infty}^{\infty} \frac{1}{2\pi \sigma_\eta \sigma_\xi} e^{-\eta^2/(2\sigma_\eta^2)} e^{-(x-\mu)^2/(2\sigma_\xi^2)} d\eta, \quad (A.5)
\]

Gauss–Hermite integration is a form of Gaussian integration for use with a kernel of the form \( w(x) = \exp(-x^2/2) \) (Davis and Rabinowitz 1984). One approximates an integral

\[
\int_{-\infty}^{\infty} f(x)w(x)dx \quad (A.6)
\]

by

\[
\sum_{i=1}^{m} w_i f(x_i), \quad (A.7)
\]

where the Gaussian points \( x_i \) and weights \( w_i \) can be obtained from a variety of computer programs. (Davis and Rabinowitz 1984, IMSL Library 1989, NAG Library 1987). This is exact when \( f(x) \) is a polynomial of degree \( 2m + 1 \) or less.

The Equation (A.5) can be transformed to resemble (A 6) by a linear transformation. In fact, we transform (A.5) so that it resembles \( w(x) = \exp(-x^2/2) \) by matching the first two derivatives of the logarithm at the mode. A function like

\[
g(x) = e^{-(x-x_0)^2/2} \quad (A.8)
\]

satisfies

\[
\frac{d \log(g(x))}{dx} = -(x-x_0)/k^2 \quad (A.9)
\]

and

\[
\frac{d^2 \log(g(x))}{dx^2} = -1/k^2. \quad (A.10)
\]

Then \( x_0 \) is a 0 of \( d \log(g(x))/dx \) and

\[
k = \left( \frac{d^2 \log(g(x))}{dx^2} \right)^{-1/2} \bigg|_{x=x_0} \quad (A.11)
\]

Thus, if we match the integrand of (A.5) in this fashion and transform using \( \eta^* = (\eta - \eta_0)/k \), we approximate the form of the kernel.
Now the log of the integrand of (A.5) is

$$h(\eta) = -\log(2\pi \sigma, \sigma,\nu) - \eta^2/(2\sigma^2) - (x - \mu\nu)^2/(2\sigma^2).$$  
(A.12)

Then

$$h'(\eta) = -\eta/\sigma^2 + \mu e^\eta(x - \mu e^\eta)/\sigma^2$$  
(A.13)

and

$$h''(\eta) = -1/\sigma^2 + \mu e^\eta(x - \mu e^\eta)/\sigma^2 - \mu^2 e^{2\eta}/\sigma^2.$$  
(A.14)

We find a root of Equation (A.13) numerically and use the associated transformation to approximate the integral.

This procedure results in fairly quick calculation of an accurate approximation of the likelihood for use in maximum likelihood estimation. In the calculations in this article, we use a 12-point approximation. The numerical optimization was performed with a quasi-Newton method using BFGS rank-two updates and a trust region approach (Dennis and Schnabel 1983). The implementation used was the IMSL routine DUMINF (1989). Starting values for the iterations when the model is in the simple form (1.4) can be derived from two simple calculations. First, \(\sigma_c\) can be estimated by the standard deviation of zero or near-zero replicates. Then, \(\sigma_a\) can be estimated by the standard deviation of the logarithms of some high concentration. For the calibration form (1.5), initial estimates of \(\alpha\) and \(\beta\) can be derived from linear regression of the average response on the log concentration.

REFERENCES


IMSL Library (1989), Houston, TX: IMSL Inc.

