Modern Measurement, Probability, and Statistics: Some Generalities and Multivariate Illustrations

Stu Hunter Conference 3/9/15

Stephen B. Vardeman

IMSE and Statistics Departments
Iowa State University
vardeman@iastate.edu
Motivation

- Physical measurement is the lens through which society builds its understanding of its world, and as the complexity of that understanding develops, so also must the sophistication of measurement.

- Effective use of new measurement technologies is inextricably tied to the advance of appropriate modeling and data analysis technology.

- This offers fascinating opportunities for the development of interesting and important new statistical methods for measurement.
Modern Metrology and Statistics

• New probability modeling is required for increasingly complex measurements ... but
  – “Differences” from “center” still function as “bias”
  – Hierarchical compounding still describes “sources of variation” like R&R, batch-to-batch variation, etc.

• Mixture features allow for “outliers” (and their rational down-weighting)

• Bayes methods are natural and often almost essential
  – High dimensionality of parameters
  – Rational representation of “Type B” uncertainty
  – A ready-made path to inference (MCMC)
The Plan Here

• State a formal generalization of standard univariate modeling of measurement

• Provide several multivariate examples of the basic message
  – determination of *cubic crystal orientation* via electron backscatter diffraction
  – determination of *particle size distribution* through sieving
  – analysis of theoretically monotone functional responses from *thermogravimetric analysis* in a materials study
  – measurement error analysis for *laser CMM outputs*
  – location and characterization of a *manufactured surface* based on data from a portable laser CMM
Multivariate Measurement

Suppose measurand $X$ is inherently multivariate. This could be as complicated as

- a set of points in 3-D on a 2-D surface of an object
- a bivariate time series of points representing the weight of a specimen and temperature as it is heated and loses mass through evaporation
- an "orientation" of an object in 3-D (represented in some appropriate way)
- a 3-D "field" of real values representing density of a specimen at a grid of points internal to the specimen

Advance in science and technology now typically requires treating a complex measurand and a measurement of it, say $Y$, as entities *rather than in terms of single elements or summaries*.
Measurement Error

Standard univariate modeling is

\[ y = x + \varepsilon \quad \text{or} \quad y - x = \varepsilon \]

Sometimes applying this component-wise to \( X \) and \( Y \) makes sense, and \( Y = X + E \)

But this may not be general enough ... e.g. where \( X \) and \( Y \) are rotation matrices specifying a 3-D orientation, one wants \( Y = EX \) for an “error” rotation matrix \( E \)
General Measurement Error

One might assume that both measurements and measurands belong to some group $\mathcal{G}$ with operation $\oplus$ (generalizing ordinary addition) and that a measurement error $\mathbf{E}$ is an element of $\mathcal{G}$ for which $\mathbf{Y} = \mathbf{E} \oplus \mathbf{X}$

Probability modeling concerns distributions on $\mathcal{G}$

“Good” distributions for $\mathbf{E}$ are ones “concentrating” around the group identity
General Measurement Error

A “central value” $\Delta$ for a distribution of $E$ then serves as “bias”

Compounding of errors from multiple sources can be handled naturally as, e.g.,

$$Y = E \oplus E_C \oplus E_B \oplus E_A \oplus X$$

Measurements from a single level of a “random effect” then share a single instance of a corresponding error
Modeling Outliers with Mixtures

Surprisingly big measurement errors need modeling—

For $F_\theta$ a parametric distribution on $\mathcal{G}$ centered at the group identity element and for which the distribution spread increases with $\theta$ in some well-defined sense, for some $p < .5$ and some $\theta < \theta^*$ a possible (no bias) model is

$$E \sim (1 - p)F_\theta + pF_{\theta^*}$$

With appropriate kinds of replication, parameters and measurands can be estimated in a principled way, e.g.,

$$Y_i = E_i X \quad \text{for} \quad i = 1, 2, \ldots, n \quad \text{for iid} \ E_i$$
Bayes Analysis

Realistic multivariate models have numbers of parameters outstripping standard sample sizes—for an obvious example, the general \( p \)-variate normal has parameter dimension \( 2p + \binom{p}{2} \).

The Bayes framework provides a coherent way to incorporate “Type B” knowledge.

MCMC provides a convenient and general “hammer” for inference.
Example 1-EBSD and Measurement of Cubic Crystal Orientations (Materials Science)
Example 1-Orientations

“Orientations” are rotation matrices that would rotate the “world” coordinate system to a coordinate system consistent with the crystalline structure

These form a group (called $SO(3)$)

Bingham, Qui, Du, Nordman and Vardeman have developed “Uniform Axis Random Spin” models for perturbations of a central orientation $S$
Example 1-UARS Models for Random Orientations

UARS rotation with central value $\mathbf{I}$ can be generated by choosing a uniformly distributed direction from the origin to define a positive axis of rotation, and then spinning the coordinate system in a clockwise direction about that axis according to some circular distribution.

A parameter (say, $\kappa$) controlling the spread of the circular distribution then controls the size of random rotations produced by the $\text{UARS}(\mathbf{I}, \kappa)$ model.
Example 1-UARS Inference

There are tractable densities for UARS variables wrt the uniform distribution on SO(3), leading to likelihoods

*Products of uniform distributions on “central values” and Jeffreys priors on concentration parameters (for the circular distributions) give Bayes models and extremely effective inference*

MCMC in (even hierarchical) versions is “easy”
Example 1-Modeling for EBSD

Multiple scans are made of the surface of a metal specimen, involving a number of fixed locations inside a number of grains in the material

\[ X_{ij} = \text{the orientation at the } j\text{th location inside grain } i \]

and

\[ Y_{ijk} = \text{the } k\text{th measurement of } X_{ij} \]

Modeling can proceed as follows
Example 1-EBSD Modeling Elements

Model as

\[ X_{ij} = L_{ij} G_i O \text{ and } Y_{ijk} = E_{ijk} S_k X_{ij} \]

For some central orientation \( O \) and independent

\( G_i \) iid \( \text{UARS}(I, \kappa_G) \) grain effects

\( L_{ij} \) iid \( \text{UARS}(I, \kappa_L) \) location effects

\( S_k \) iid \( \text{UARS}(I, \kappa_S) \) scan effects

\( E_{ijk} \) iid \( \text{UARS}(I, \kappa) \) repeatability errors
Example 1-Bayes Inferences

Parameters are \( \mathbf{O}, \kappa_G, \kappa_L, \kappa_S, \) and \( \kappa \)

Well-calibrated credible sets for these are available (and allow comparison of “variance components”)

Further, the MCMC basis of inference makes \textit{interpretable} credible sets available for every orientation measurand \( X_{ij} \)
Example 1-Set of Cones Credible Set
The REAL problem actually isn’t so simple ... though EBSD data are usually treated as orientations, they are really representatives of equivalence classes (of size 24 in the cubic crystal case) of indistinguishable orientations.

The real problem motivates even more interesting methodology development based on the foregoing.
Example 1-Set of Cones Credible Set for the Real Problem
Example 2-Sieving and “Particle Size Distributions”

From environmental soil studies, to analysis of cement composition, to mastication studies, to preparation of plastic bonded explosives powders or drug delivery systems, it is often important to know what size particles make up a granular material or powder.

“Particle size distributions” are of interest.
Example 2-Sieving

A standard method of characterizing a sample from such a particle system is to run it through a set of progressively finer sieves, weighing the amount of material caught on each sieve.
Example 2-Notation

For sieve sizes $C_1, \ldots, C_{k-1}$ with

$0 \leq C_0 < C_1 < \cdots < C_{k-1} < C_k \leq \infty$ suppose that actual weight fractions of particles in a specimen of material of total weight $m$ in intervals $[C_{i-1}, C_i)$ for $i = 1, 2, \ldots, k$ are respectively $p_1, p_2, \ldots, p_k$

Analysis of these weight fractions of material is what is typically meant by analysis of a particle size distribution based on a sieving study
Example 2-Modeling

Under a modeling assumption of successive random sampling of particles, renewal theory implies there is a function $CW(s, \theta)$ of features $\theta$ of the joint distribution of the size-weight pairs in the "population" of particles, giving the limit as $m \to \infty$ of the weight fraction of particles with size less than or equal to $s$

Further, multivariate normal limits apply
Example 2-Modeling

For $q_i = \ln p_i - \ln p_1$ for $i = 2, \ldots, k$ and $q = (q_2, \ldots, q_k)'$

$q \sim \text{MVN}_{k-1} \left( \Delta(\theta), \Sigma(\theta) \right)$

for $\Delta(\theta) = (\Delta_2(\theta), \ldots, \Delta_k(\theta))'$ where

$$\Delta_i(\theta) = \ln \frac{(CW(C_i, \theta) - CW(C_{i-1}, \theta))}{(CW(C_1, \theta) - CW(C_0, \theta))}$$

and $\Sigma(\theta)$ also depends upon the features of the particle size-weight distribution
Example 2-Modeling Caveat

The modeling just thus far covers only variation in real weight fractions

It is possible to add parameters allowing for and describing *precision of weight measurement* and produce more flexible forms for $\Sigma$ depending upon them as well as upon $\theta$

Limited testing suggests that not too much is lost ignoring this issue
Example 2-Data Analysis

10-20 sieves are common and sample sizes (numbers of specimens sieved) are often small

Bayes analyses with MVN likelihoods provide inference for $\theta$, values of $CW$ and $CW^{-1}$ and predictions of new vectors of fractions $p_1, \ldots, p_k$ based on iid $q_i$. 
Example 2-One Sample Data
Example 2-iid Model (4 Parameters)
Example 2-iid Model (7 Parameters)
Example 2-Hierarchical Modeling

Leyva, Page, Vardeman, and Wendelberger analyzed 2 results from each of 6 different batches of a material.

A 7-dimensional parameter was broken into a 4-dimensional part common across lots and a 3-dimensional part that was allowed to vary lot-to-lot.

“Flat” priors were used to make inferences for parameters, lot-specific cumulative weight fraction functions and their inverses, and predictions for new vectors of (the 20) weight fractions.
Example 2 - Data from 6 Batches
Example 2-Inferences from 6 Batches
Example 2-Outliers in Sieving Cement (Round Robin of 21 Labs)
Example 2-Modeling Outliers

Page and Vardeman used a partially specified contamination model (PCM) for analysis.

Observed \( q_i \) for \( i = 1, 2, \ldots, 21 \) were modeled as iid

\[
(1 - p) \text{MVN}_{k-1}(\Delta(\theta), \Sigma(\theta)) + p \text{MVU}_{k-1}(G, H)
\]

for specified \( G \leq H \)

A beta prior for the mixing parameter and a “flat” prior for \( \theta \) enabled inference for the reference \( CW(s, \theta) \) and posterior probabilities that each lab is “outlying”
Example 2-Bayes Mixture Analysis
Example 3-Thermogravimetric Analysis

It is often of interest how the mass of a specimen changes with temperature (e.g. as volatiles evaporate or oxidation occurs)

A "measurement" in a TGA study is actually a time series of data pairs \((\text{temp}(t), \text{weight}(t))\) taken at a very high time \((t)\) frequency

Without measurement error (in both temperature and weight), both time series would typically be monotone and the observed points would fall on a curve representing a monotone (weight) function of temperature
Example 3-An Example
Example 3-Analysis

There seems to be no choice but to begin by (ignoring time order and) smoothing \((temp, weight)\) data pairs to create a \(weight(temp)\) function. Subject matter interest then concerns the location (in temp) and level (in weight) of the plateaus ... ? and perhaps some comparison of the “shapes” of the shoulders ?

After choosing a fine temperature grid, first differences are essential.
Example 3-Smooth and 1\textsuperscript{st} Differences on a Temperature Grid
Example 3-Definition and Modeling of Features

Upon adoption of some definition of (temp) intervals where weight is nearly constant (derivative is essentially 0) over which one averages weight to get the *elevation* of a plateau, *and* a definition of *location* of a plateau (such as the first temperature after a large negative derivative with an essentially 0 derivative) one has *numerical features* to be modeled.

The most obvious place to begin is then with a (7-dimensional) MVN distribution for (4) elevations and (3) locations.
Example 3-MVN Modeling of Features

Frequentist joint inference will already typically be problematic, as numbers of TGA runs for a single set of conditions will not approach the dimensionality of the unconstrained parameter space

(Simple addition) Compounding of MVNs can be useful for random effects modeling

Vaca, Hamada, Moore, Burr and Vardeman have made some Bayes efforts here (labs and repeats)
Example 3-Detailed Analysis of Shapes

Physical differences between lab set-ups, exact temperature increase schedules, batches of a material, etc. can vary shapes of the "shoulders" of the weight-versus-temperature curve as it drops from one plateau to the next.

One might try to establish a “standard” shape for each section of a type of TGA curve and model deviations from it.
Example 3-Standard Shapes

To account for small differences in

\[ \Delta \text{temp} \text{ and } \Delta \text{weight} \]

for a fixed section of TGA plots that “should be the same” it seems best to rescale both temp and weight to \([0,1]\) before analysis

(Estimated) Standard shapes can then be made by smoothing points from several rescaled weight versus temp functions
Example 3-Deviations From and Between (Estimated) Standard Shapes

Mean 0 Gaussian processes $d(t)$ on $[0,1]$ conditioned to have $d(0)=d(1)=0$ might model deviations from and between standard shapes (parameters control smoothness and variance) (Some modification of basic stationary models is probably needed to handle variance increasing with $t$ on $[0,1]$)

Vaca, Hamada, Moore, Burr and Vardeman have made some Bayes efforts here (labs and repeats)
Example 4-Bayes Error Analysis for a Laser CMM Observation
Example 4-A Few Applications

Portable laser CMMs are e.g. used in machine calibration, **CAD-based inspection**, and tool-building and setup in manufacturing, in accident reconstructions, and for as-built documentation and 3-D modeling of existing buildings.

How precise is an \((x, y, z)\) (relative to a coordinate system on the CMM) position measured by such a device?
Example 4-Geometry (From a Classic 1945 QST Radar Article)

QST
Example 4-Simple Geometry and Type B Uncertainty

Relationships between what is actually observed and the Euclidean measurand are simple:

\[
x = r \cos \theta \cos \phi \\
y = r \cos \theta \sin \phi \\
z = r \sin \theta
\]

In many applications (e.g. scanning of a building exterior) no real repeating of a particular measurement is possible
Example 4-Bayes/Type B Modeling

Treat observed \((\phi, \theta, r)\) as measurements from distributions with central values (measurands) \((\mu_\phi, \mu_\theta, \rho)\) and use these plugged into the previous equations as Euclidean measurands.

Natural choices of models for angles are distributions on the group \(SO(2)\), for example the vonMises distributions with pdfs

\[
f(\alpha | \mu, \kappa) = \exp \left( \kappa \cos(\alpha - \mu) \right) \frac{1}{2\pi I_0(\kappa)} \quad \text{for} \quad \alpha \in [-\pi, \pi]
\]
Example 4-vonMises Densities

Here are polar plots of \((1+\text{density})\) for several vonMises distributions:

- **red**: \(\mu = 0, \kappa = 3\)
- **dark blue**: \(\mu = 0, \kappa = 15\)
- **violet**: \(\mu = \frac{\pi}{4}, \kappa = 15\)
- **light blue**: \(\mu = \frac{\pi}{2}, \kappa = 40\)
Example 4-Likelihood and Prior

With (independent) vonMises distributions for bearing and elevation centered at the actuals and with Type B information about precision encoded in kappas independent of an observed range uniformly distributed in some (Type B specified) range about the actual and uniform priors on angles and (improper) uniform prior on range, the posterior is simple
Example 4-Posterior and Uncertainty

A posterior density is proportional to
\[
\exp\left(\kappa_\phi \cos(\phi - \mu_\phi) + \kappa_\theta \cos(\theta - \mu_\theta)\right)I[r - \Delta < \rho < r + \Delta]
\]

The actual angles are (independent) vonMises also independent of the uniform actual range

Simulating and answering any probability question about actual x-y-z coordinates is then trivial (e.g. making a 95% probability ball for the actual location)
Example 5-Portable Laser CMM (Location and) Characterization of a Manufactured Surface
Example 5-CMM Inspection

Portable laser CMMs have applications in checking conformance of very large manufactured parts to design specifications.

Colleagues at ISU have used them in studies of form variation seen in molded fiberglass wind turbine blades, objects that can reach 75m in length.

Many measured points on a surface of such a blade can be recorded and need to be somehow compared to what is "ideal"
Example 5-Some (?Enlightened?) Speculation About Modeling

A CAD model for the designed surface might give rise to an "average part surface for a particular mold" through the addition of a realization from some mean 0 Gaussian spatial process (in the "z" direction as a function of $(x, y)$)

Then particular part surfaces for a given mold might be modelled as perturbations of the mold average through addition of an independent realization of some other mean 0 Gaussian spatial process

That is, the ideas of hierarchical probability modeling can be employed to describe part-to-part variation in form
Example 5-Connecting a Model for Part Surfaces to Data

For a given set-up of the CMM, there is the location and orientation of the part relative to its coordinate system.

For a part coordinate system set up on a corner of the part, bounds could easily be set (and then uniform distributions produced for describing) the location of the origin of that system and the size of the rotation of part axes from the CMM axes.

One then has a model for a realized part surface in a coordinate system corresponding to the CMM that allows for part-to-part and mold-to-mold variation.
Example 5-Model for Data Given Surface

One model for the set of observed points on the realized part surface (giving rise to the set of measurements taken) is that they are \textit{a priori} iid and uniform in the "x-y" coordinates defining the untranslated and unrotated ideal surface with (untranslated and unrotated) "z" coming from the addition of mold and part perturbations of the ideal design.

Then each of these individual locations on the part surface has its own \((\mu_\phi, \mu_\theta, \rho)\) and given these parameters observation of \((\phi, \theta, r)\) is subject to same kind of randomness just discussed.
Example 5-MCMC

This gives a probability model for many observed and actual locations on the surface, a part coordinate system origin, and an orientation.

It has parameters associated with the Gaussian processes and the conditional distributions of observed angles and range given their actual values.

Upon putting priors on the part coordinate system origin and location, and Gaussian process parameters, and specifying type B figures, MCMC sampling of the posterior of the parameters of the Gaussian process (that quantify variation in molds and parts from a given mold) and finding predictions for mold average surfaces and part actual surfaces seems possible.
Summary

• Modern metrology motivates interesting highly multivariate modeling and inference problems
• Standard concerns of bias, precision, variance sources remain but must be technically rethought
• Mixture modeling effectively treats outliers
• Bayes methods are effective and almost essential