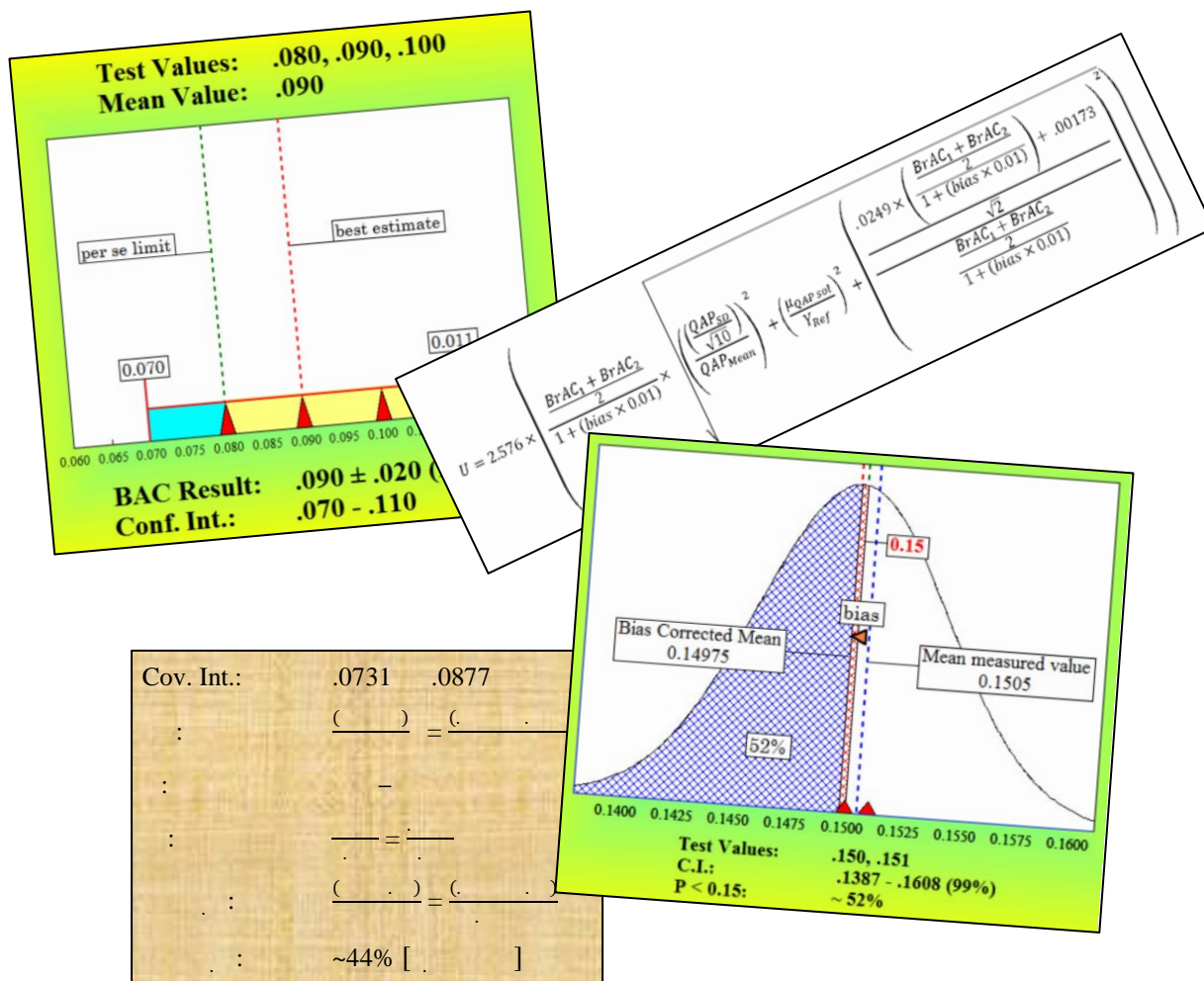


Introduction to Uncertainty



TED VOSK¹

Attorney at Law/Consultant
 Of Counsel, Cowan, Kirk, Gaston

One can never know how well he knows a thing, only how strongly he believes it.

¹ Correspondence should be directed to tvosk@comcast.net. See also, Vosk, *Measurement Uncertainty: Forensic Metrology and Result Interpretation Part I & II* in *Understanding DUI Scientific Evidence*, Chapters 7 & 8 (4th ed. Aspatore, 2011); Vosk & Emery, *Forensic Metrology: A Primer on Scientific Measurement for Lawyers, Judges and Forensic Scientists* (CRC Press – In Preparation).

For Kris: Who believed in me when no one else did and made all things possible.

INTRODUCTION

Revealed Truth, absolute and known, is not the domain of science. Rather, it is relative inference. From observation and data, to the relationships alive therein, to varying degrees of certitude never complete. That's the promise of science... and the best it can do.

Exactly what constitutes “science” has long been a matter of debate. The word “science” comes from the Latin *scientia*, meaning knowledge. This is natural enough as science is characterized by the goal of acquiring knowledge. There are many activities where the goal is the acquisition of knowledge, however, that would not be seen as science. For example, both philosophy and religion are frequently relied upon as a source of, or tool for obtaining, knowledge, but neither constitutes science. So, what is it that separates these disciplines from what would be characterized as science?

Some consider falsifiability and testability to be the hallmarks of science. The object here is the contention being made. A contention is scientific if it is open to being empirically tested and falsified. Others see science not so much as centering on the status of particular propositions, but instead as more of a puzzle solving activity. Here, the focus is on the actions engaged in. Actions are scientific if they are aimed at empirically solving unanswered questions. Still others claim that science is characterized by the ability to successfully predict future phenomenon. This looks at forecasting. Forecasts are scientific if they are empirically based and able to predict unknown outcomes with a high degree of certainty.

There is nothing necessarily incompatible about these different conceptions. In their less dogmatic forms, each seems to capture some aspect of the scientific enterprise. Like the three blind men who, upon touching different parts of an elephant described very different beasts, these descriptions of science seem to convey distinct characteristics of a greater whole. Which

aspect one focuses on may simply be a function of whether we are considering a body of accumulated knowledge or instead the practices engaged in to obtain such knowledge.

One thing common to each of these frameworks is that each has an empirical basis. Whatever science is, all agree that it requires collection and analysis of information *from* the physical world in order to gain knowledge *about* the physical world. It “is based on the principle that...observation is the ultimate and final judge of the truth.”¹ For any claim to be considered scientific, it must either be derived from and/or withstand the test of, systematic observation. Thus, unlike religion or philosophy which may or may not choose to incorporate rigorously obtained empirical information, science requires it.

Another fundamental aspect of science is that in addressing “truth”, science does not deal in absolutes.

Ontologically, although science must be empirically based, its true subject matter, as a human enterprise, is simply information and the relationships inherent within it. Although we ascribe an external reality to revealed physical principles and/or conclusions, in the end we are limited by our evolved senses and manner of reason. And while these tools may serve us well in an evolutionary sense, only faith connects their products to fundamental reality. In the end, science deals strictly with information, the manner in which it is processed and the relationships found or created within it by inference. Our scientific “truths” consist of little more than models of information networks that rigorously correspond to experience, which may or may not correspond to deeper reality.

Epistemologically, when science is relied upon to guide our understanding of the physical world, the absolute is still denied to us by the fact that *uncertainty is inherent to all scientific knowledge*. The relationships perceived and the conclusions drawn there-from are all soft edged to

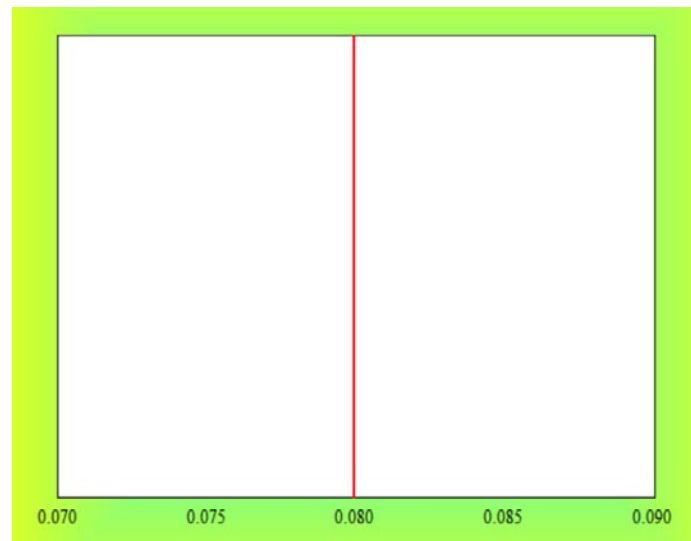
some extent. Although the degree of belief associated with a given proposition may be high, science cannot prove it. “It is scientific only to say what is more likely and what is less likely.”² Thus, even when the connection between experience and physical reality is accepted, the best we can do is speak of the relative strength of our inferences.

In this context, however, one of the remarkable aspects of science is its ability to *rigorously characterize* the relative strength of our inferences and communicate the *level of uncertainty* associated with certain types of claims. That is, we can grade the degree of fuzziness surrounding certain scientific claims so that their strengths can be widely understood and compared. This is critical to our understanding of science or any claim made in its name. In this manner, science brings order to our world through the creation and utilization of models which enable us to make predictions and/or inferences with *relative degrees of certainty* concerning the physical phenomena of our experience.

From this brief discussion, a simple picture of what constitutes science can be painted. Science focuses on the quest for, and acquisition of, knowledge of the physical world as it is perceived and processed by our senses. It requires the systematic collection of empirical information followed by an assessment of the strengths and weaknesses of that information. Where the information permits, relationships are discovered and causal inferences are made creating knowledge in the form of an explanation for what has been observed and the ability to determine what will happen in the future when certain criteria are satisfied. Finally, based on our information, the significance of that knowledge is evaluated through a rigorous determination of its limits and, in particular, the degree of certainty associated with it. Absent a rigorous understanding of these limits, however, the weight accorded scientific knowledge/conclusions is a matter of faith. Science requires more.

The uncertainty associated with a piece of scientific knowledge provides a measure of its practical epistemological robustness. It represents the degree to which our knowledge concerning the relevant physical phenomenon is imperfect. Although uncertainty may be minimized, it cannot be eliminated: it is inherent to all scientific knowledge. Fortunately, science has developed methods for rigorously characterizing and communicating the level of uncertainty associated with certain types of claims. These methods permit us to assign a “degree of belief” that can be placed in such knowledge expressed as a likelihood or level of confidence.

Consider the case of scientific measurement. To most individuals, the value obtained by a scientific measurement means exactly what it says. Thus, if the value reported by a breath test instrument is 0.08, most individuals will assume that the measurement represents a true and accurate value for an individual’s BrAC of 0.08.

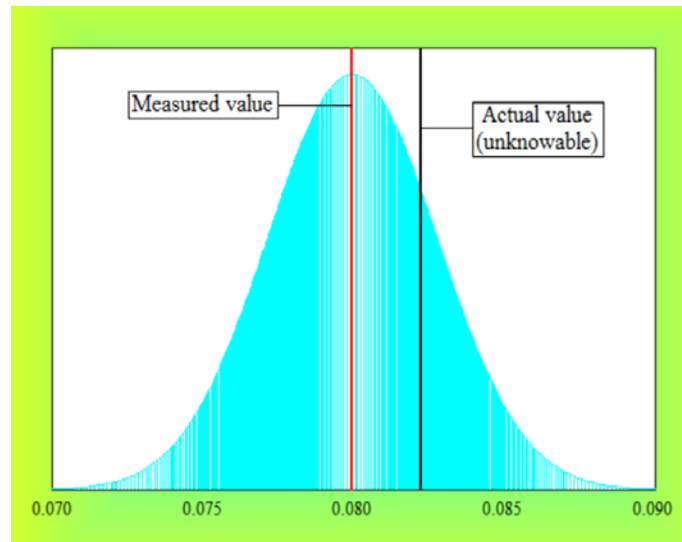


For even the most carefully performed measurement, however, the value of a thing being measured (the “measurand”) can never be known exactly; all that can ever be given is an estimated value.³ In fact:⁴

...for a given measurand and a given result of measurement of it, there is not one value but an infinite number of values dispersed about the result that are consistent with all of the observations and data and one’s knowledge of the

physical world, and that with varying degrees of credibility can be attributed to the measurand.

Thus, in the context of a breath test yielding a value of 0.08, the truth is more accurately represented as a packet of values, any of which may represent an individual's actual BrAC.



In essence, “[t]he result of a measurement is a probability distribution that provides an unambiguous encoding of one’s state of knowledge about the measured quantity.”⁵ The uncertainty associated with a measurement supplies a quantitative statement characterizing the dispersion of values that can actually and “reasonably be attributed to the measurand.”⁶ “Requirements for measurement accuracy translate into a need to know not only the results of measurements but the uncertainties associated with the results.”⁷ “The result of a measurement cannot be correctly evaluated without knowing its uncertainty.”⁸

Knowledge of the uncertainty associated with measurement results is essential to the interpretation of the results. Without quantitative assessments of uncertainty, it is impossible to decide whether observed differences between results reflect more than experimental variability, whether test items comply with specifications, or whether laws based on limits have been broken. Without information on uncertainty, there is a risk of misinterpretation of results. Incorrect decisions taken on such a basis may result in unnecessary expenditure in industry, incorrect prosecution in law, or adverse health or social consequences.⁹

Accordingly, “[a] result is complete only when accompanied by a quantitative statement of its uncertainty.”¹⁰ The most common way of expressing measurement uncertainty is as a coverage interval. It consists of a best estimate of the “true” value of the measurand accompanied by a range of values that can also be attributed to the “true” value with a given level of confidence (probability).

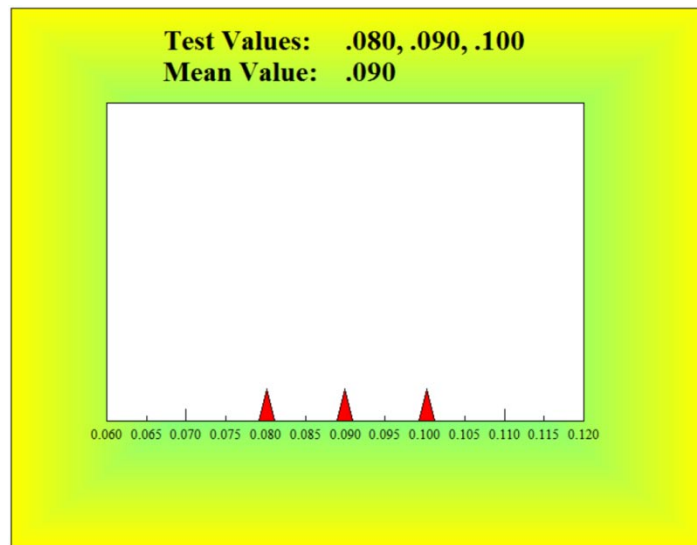
<u>Measurement Result</u>	
Measured mean value:	22 mg/dL
Uncertainty ($k=2$; 95%):	± 2.5 mg/dL
<u>Result:</u>	22 ± 2.5 mg/dL (95%)

When the result of a forensic measurement is reported simply as “a number,” it does not reflect the accuracy of the measurement and cannot be properly interpreted.”¹¹ “For example, methods for measuring the level of blood alcohol in an individual...can do so only within a confidence interval of possible values.”¹² “Estimating and reporting measurement uncertainty with the number completes the picture and allows us to properly use the result to make reliable and defensible decisions.”¹³ The importance of uncertainty as part of a complete result can be illustrated using a simple example provided by the National Academy of Sciences.¹⁴

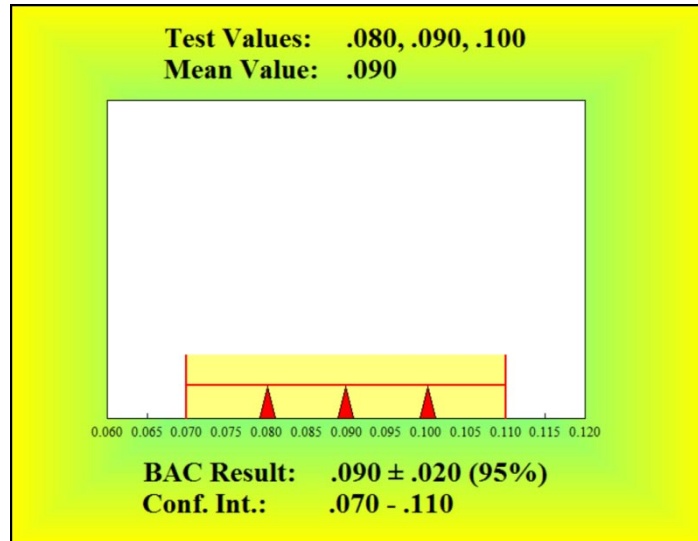
Consider...a case in which an instrument (e.g., a breathalyzer such as Intoxilyzer) is used to measure the blood-alcohol level of an individual three times, and the three measurements are 0.08 percent, 0.09 percent, and 0.10 percent. The variability in the three measurements may arise from the internal components of the instrument, the different times and ways in which the measurements were taken, or a variety of other factors. These measured results need to be reported, along with a confidence interval that has a high probability of containing the true blood-alcohol level (e.g., the mean plus or minus two standard deviations). For this illustration, the average is 0.09 percent and the standard deviation is 0.01 percent; therefore, a two-standard-deviation confidence interval (0.07 percent, 0.11 percent) has a high probability of containing the person’s true blood-alcohol

level. (Statistical models dictate the methods for generating such intervals in other circumstances so that they have a high probability of containing the true result.)

In this example, each of the measured values equals or exceeds a BrAC of 0.08, the level at which most states have defined the crime of *per se* DUI. If a jury is supplied with only the values reported by the instrument, the picture created overwhelmingly suggests, and is likely to result in, guilt.



Unfortunately, this picture is both incomplete and misleading. Once the uncertainty is included, we see that values as low as 0.07 may actually and reasonably be attributed to this individual's true BrAC.

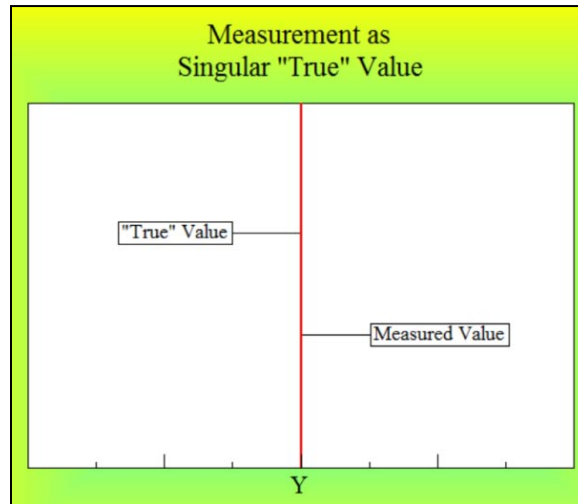


Of course, the uncertainty here does not dictate a particular result. It simply provides the decision maker with a complete and honest picture of the scientific “truth” so that all the evidence can be properly understood and weighed. Without more, though, it is obvious that an average juror may discover reasonable doubt hiding in the previously unreported uncertainty associated with these results. Clearly, “considering or not the uncertainty of a critical result can make the difference between acquittal and a guilty sentence.”¹⁵

The same general principles apply to scientific calculations based upon measured or experimentally determined values. Even where the relationship between such physical quantities is well understood, the values assigned to each component quantity are typically accompanied by uncertainty which translates into uncertainty associated with the final calculated result. If the results of such calculations are to be interpretable, then the uncertainty associated with each must be determined and communicated to those who intend to rely upon them.

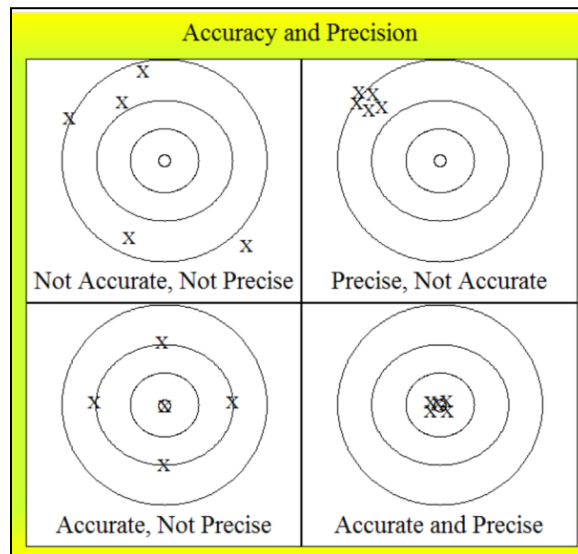
A BRIEF TOUR

ILLUSIONS OF CERTAINTY¹⁶



“It is scientific only to say what is more likely and what is less likely.”¹⁷

ACCURATE AND RELIABLE



“When uttered by an expert witness, the phrase ‘accurate and reliable’ is infused with talismanic connotation, beckoning one to trust the result presented. Unfortunately, it conveys little real information concerning how good a result actually is or what it means.”¹⁸

“Because ‘accuracy’ is a qualitative concept, one should not use it quantitatively, that is, associate numbers with it.”¹⁹

MEASUREMENT ERROR

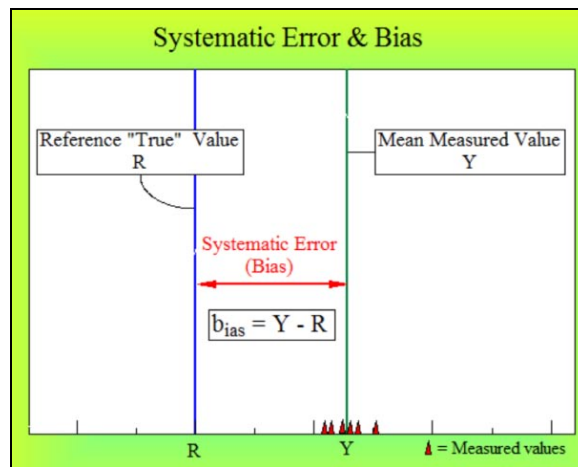
The objective of error analysis is to determine an estimate of a measurand's value that is as close as possible to *the* true value by identifying, accounting for and minimizing as many sources of measurement error as possible.²⁰

The Model

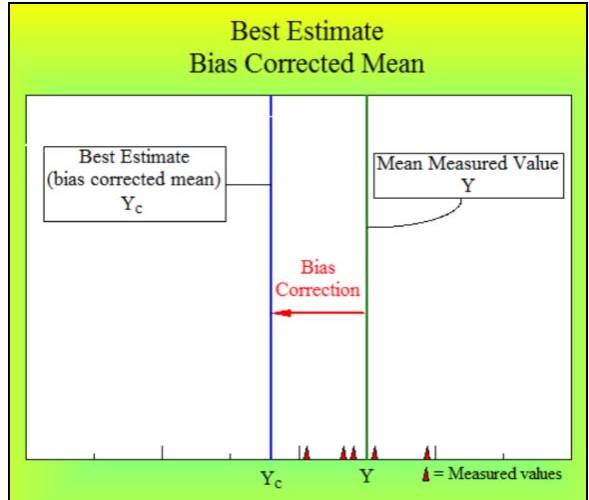
$$\bar{Y} = X + \epsilon$$

where
 \bar{Y} = mean of set of measurements
 ϵ = measurement error

Two Types of Error: Systematic and Random



$$\bar{Y} = X + \epsilon$$

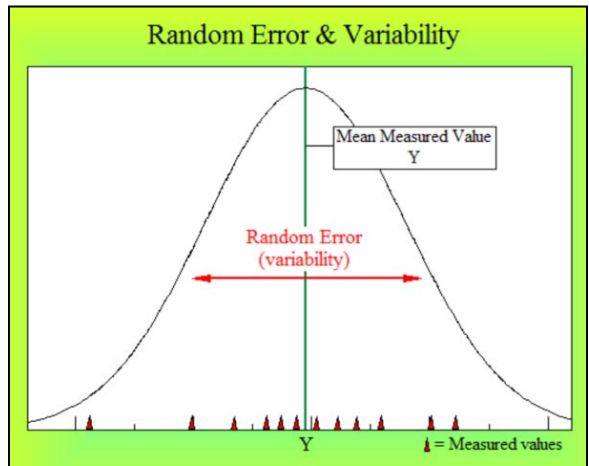


Correction Constant Bias

$$= -$$

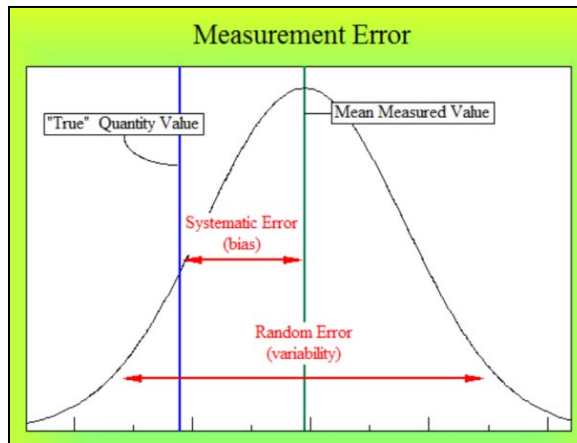
Correction Percent Bias

$$= \frac{\quad}{\quad} \%$$



Standard Deviation

$$= \frac{(y - y)}{n - 1}$$

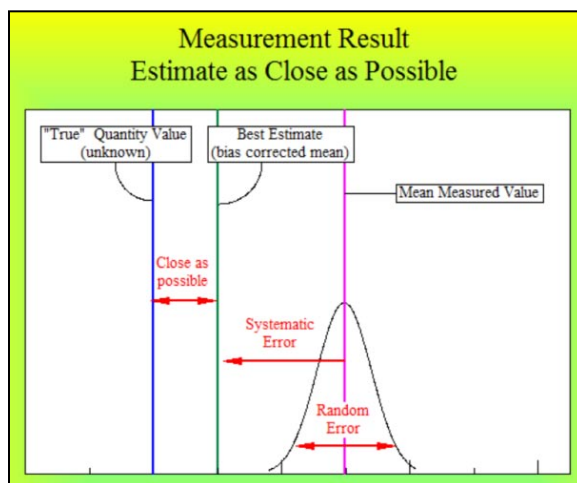


The Model

= -

Where

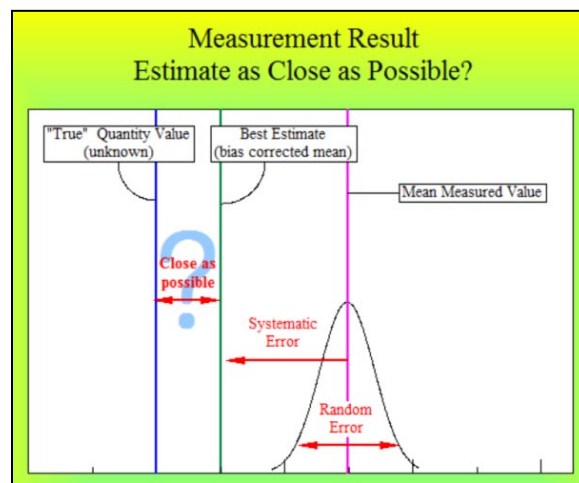
= +



Frequentist theory of probability: Probability defined as relative frequency of occurrence over the universe (population) of possible events.

A Problem for Error Analysis

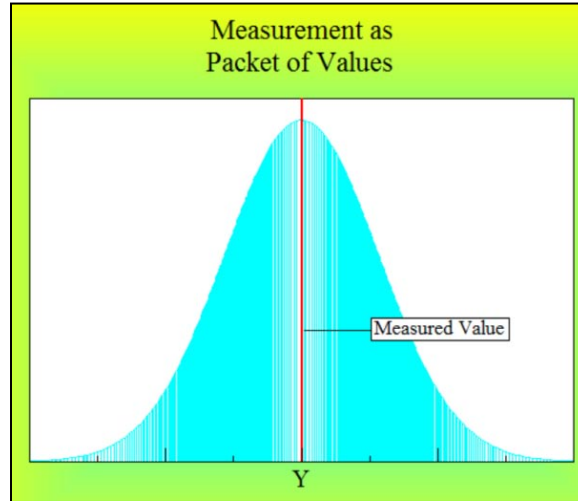
$$= \boxed{\frac{?}{\times}}$$



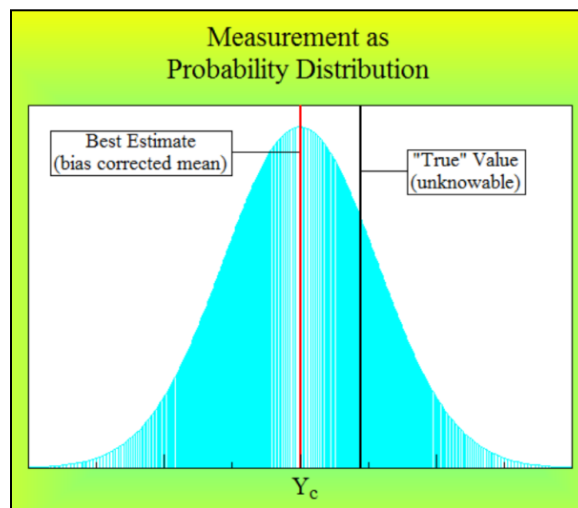
Measurement "error is an unknowable quantity in the realm of the *state of nature*."²¹

MEASUREMENT UNCERTAINTY

Error and Uncertainty are “not synonyms, but represent completely different concepts; they should not be confused with one another.”²²

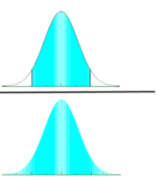


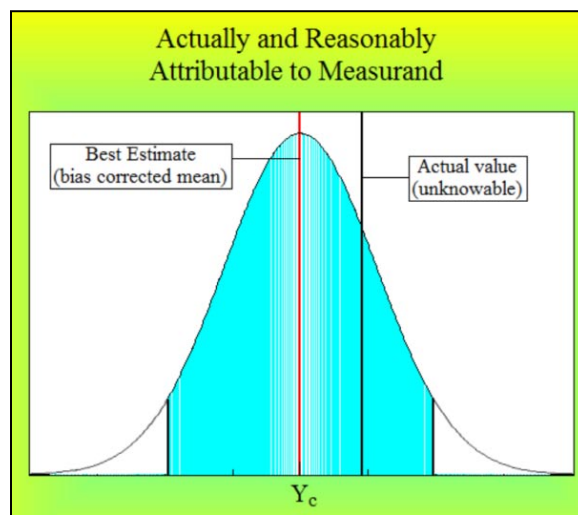
“...for a given measurand and a given result of measurement of it, there is not one value but an infinite number of values dispersed about the result that are consistent with all of the observations and data and one’s knowledge of the physical world, and that with varying degrees of credibility can be attributed to the measurand.”²³



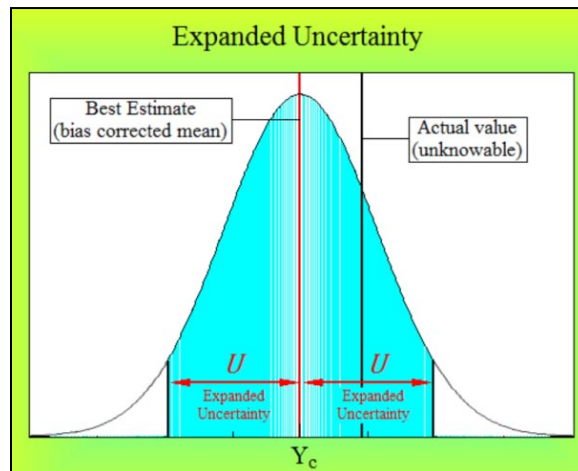
“The result of a measurement is a probability distribution that provides an unambiguous encoding of one’s state of knowledge about the measured quantity.”²⁴

The probability that a measurand's value lies within a specified range of values can be visualized as the area under the curve spanning the range. The proportion of the area under the curve spanning our range of values to the total area under the curve yields the probability that the measurand's value is contained within the specified region.

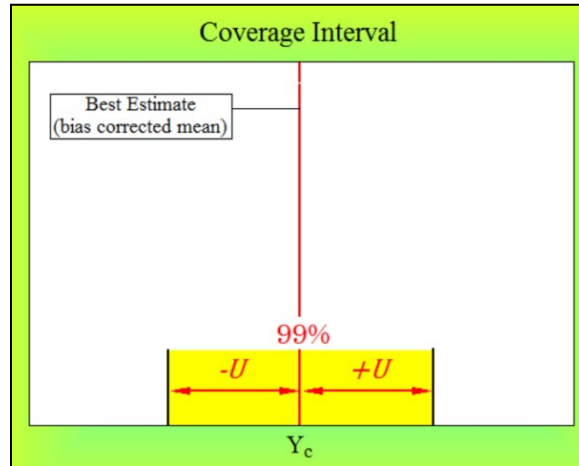
$$P_{rob} = \frac{\text{Area under curve in range}}{\text{Total area under curve}}$$




The expanded uncertainty, U , defines “an interval about the result of a measurement that may be expected to encompass a large fraction of the distribution of values that could reasonably be attributed to the measurand.”²⁵



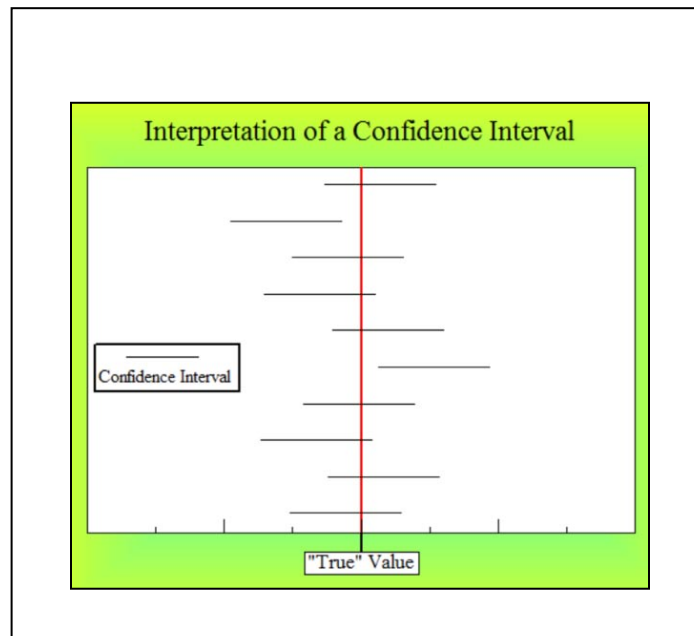
“Knowledge of the uncertainty associated with measurement results is essential to the interpretation of the results. Without quantitative assessments of uncertainty, it is impossible to decide...whether laws based on limits have been broken. Without information on uncertainty, there is a risk of misinterpretation of results. Incorrect decisions taken on such a basis may result in unnecessary expenditure in industry, incorrect prosecution in law, or adverse health or social consequences.”²⁶



Coverage Interval

$$- \quad Y_{99\%} \quad +$$

A coverage interval is an “interval containing the set of true quantity values of a measurand with a stated probability, based on the information available.”²⁷



“In general, the result of a measurement is only an approximation or estimate of the value of the specific quantity subject to measurement and thus the result is complete only when accompanied by a quantitative statement of its uncertainty.”²⁸ “Given the inherent variability of measurement, a statement of a measurement result is incomplete (perhaps even meaningless) without an accompanying statement of the estimated uncertainty of measurement.”²⁹ “When the value of a measurand is reported, the best estimate of its value and the best evaluation of the uncertainty of that estimate must be given.”³⁰

Measurement Result = Best Estimate \pm Uncertainty

= \pm (99%)

Bayesian theory of probability: Probability defined as an information based degree of belief that an event will occur.

“Uncertainty explicitly acknowledges our lack of information and instead of claiming to tell us ‘what reality is’ it simply says that, based on the information that was considered, this is ‘what we believe reality to be.’ This is the essence of the uncertainty paradigm. It does not claim to convey the truth of a thing: it claims only to convey what is believed about a thing.”³¹

“...uncertainty (in measurement) is a quantifiable parameter in the realm of the *state of knowledge about nature*.”³²

EXAMPLE

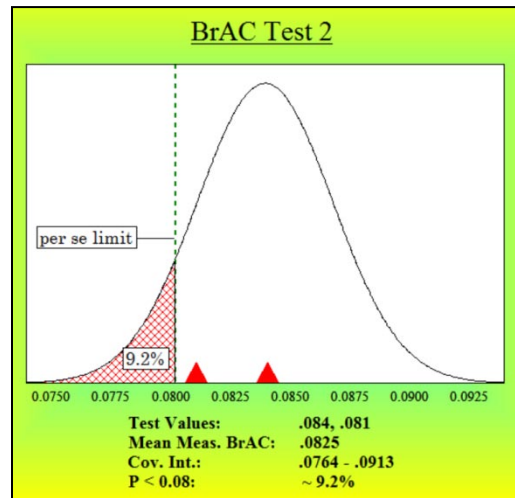
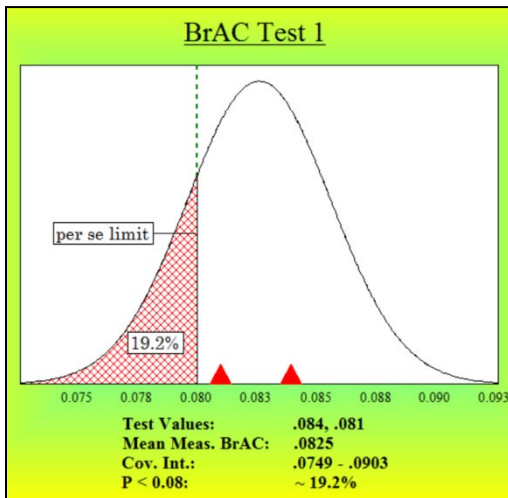
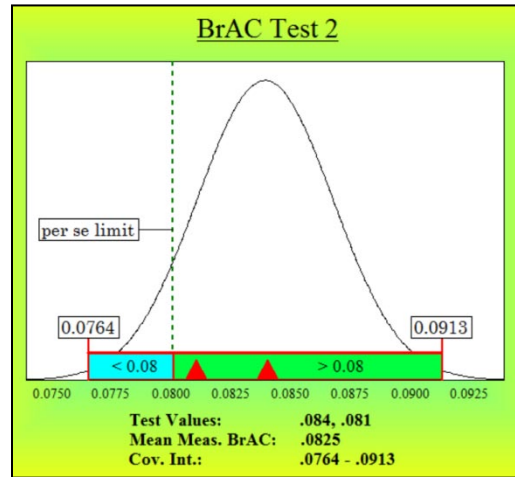
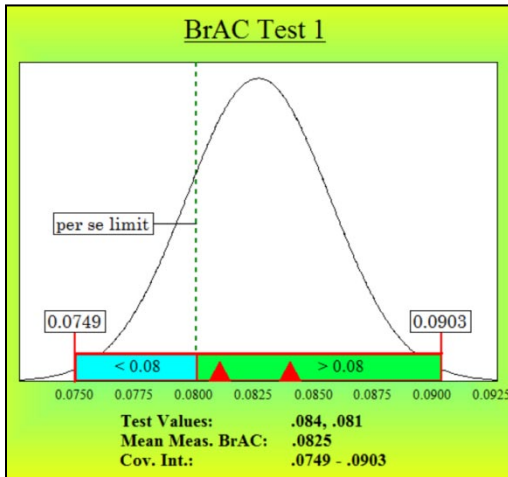
Identical test results, different meanings

```

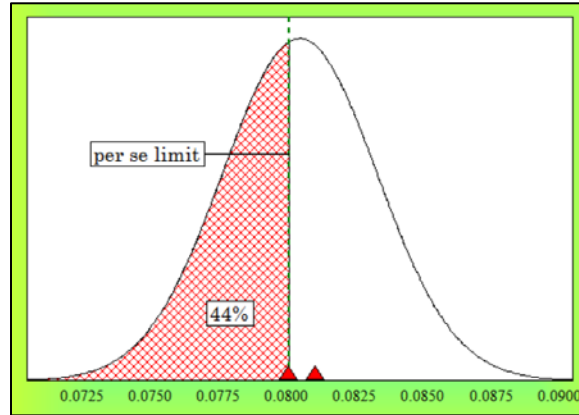
--- BREATH ANALYSIS ---
BLANK TEST      .000
INTERNAL STANDARD VERIFIED
SUBJECT SAMPLE  .084
BLANK TEST      .000
EXTERNAL STANDARD .082
BLANK TEST      .000
SUBJECT SAMPLE  .081
BLANK TEST      .000
    
```

```

--- BREATH ANALYSIS ---
BLANK TEST      .000
INTERNAL STANDARD VERIFIED
SUBJECT SAMPLE  .084
BLANK TEST      .000
EXTERNAL STANDARD .079
BLANK TEST      .000
SUBJECT SAMPLE  .081
BLANK TEST      .000
    
```



DETERMINING THE PROBABILITY THAT BRAC < 0.08 g/210L IN 6 EASY STEPS³³



Step 1: Start with the coverage interval.

Step 2: Determine the bias corrected mean.

$$= (\quad)$$

Step 3: Determine the expanded uncertainty.

$$= \quad -$$

Step 4: Determine the combined uncertainty.

$$= \quad \text{---}$$

Step 5: Determine the z-factor.

$$= (\quad)$$

Step 6: Determine probability from statistical table.

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064

EXERCISE³⁴

BREATH ALCOHOL UNCERTAINTY: WASHINGTON STATE

--- BREATH ANALYSIS ---	
BLANK TEST	.000
INTERNAL STANDARD	VERIFIED
SUBJECT SAMPLE	.082
BLANK TEST	.000
EXTERNAL STANDARD	.081
BLANK TEST	.000
SUBJECT SAMPLE	.088
BLANK TEST	.000

Some necessary relationships

$$1. \quad = \pm (99\%) \quad 3. \quad = \frac{\quad}{\quad} + \quad + \frac{\quad}{\quad}$$

where

$$2. \quad = \frac{\quad}{(\quad \cdot \quad)}$$

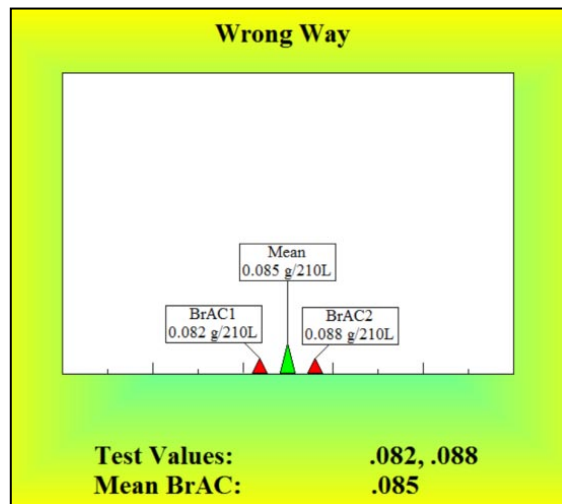
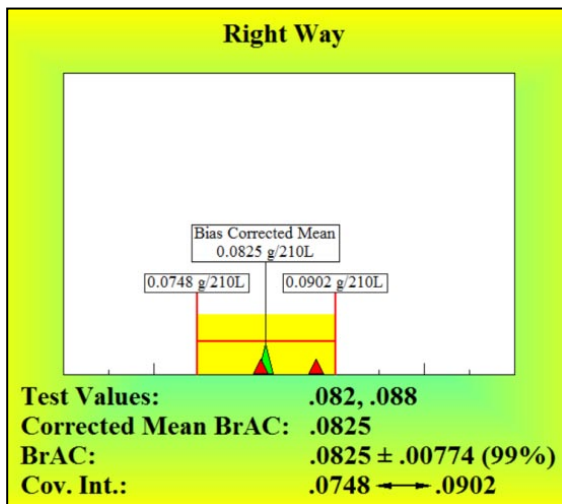
$$4. \quad / = (0.0249 \cdot \quad) + 0.00173$$

$$5. \quad = \quad \cdot$$

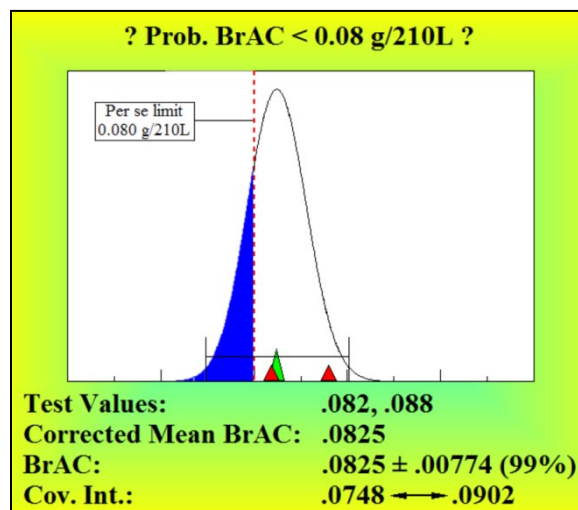
(Documents & Data
at back of packet)

Result?

When we report a result simply as “a number,” it does not reflect the accuracy of the measurement and cannot be properly interpreted. Estimating and reporting measurement uncertainty with the number completes the picture and allows us to properly use the result to make reliable and defensible decisions.”³⁵ “All results for every forensic science method should indicate the uncertainty in the measurements that are made, and studies must be conducted that enable the estimation of those values.”³⁶ “For example, methods for measuring the level of blood alcohol in an individual or methods for measuring the heroin content of a sample can do so only within a confidence interval of possible values.”³⁷ In particular, breath alcohol “results need to be reported, along with a confidence interval that has a high probability of containing the true blood-alcohol level (e.g., the mean plus or minus two standard deviations).”³⁸



Probability “true” BrAC is less than 0.08 g/210L?



ERROR V. UNCERTAINTY

Measurement “error is an unknowable quantity in the realm of the *state of nature*...uncertainty (in measurement) is a quantifiable parameter in the realm of the *state of knowledge about nature*.”³⁹

The Issue of Below Threshold Interferents

$$= 0.01 / 210$$

Error analysis Bounded error

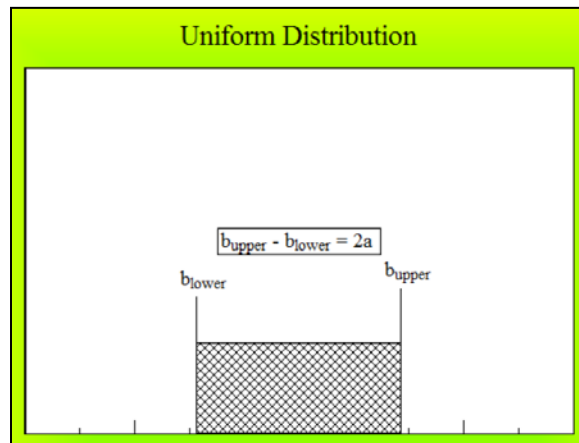
Uncertainty Analysis Model knowledge with probability distributions

Error Analysis

$$= 0.01 / 210$$

Result?

Uncertainty Analysis



$$= \frac{2}{2}$$

$$= \frac{1}{3}$$

=?

Result?

“Scientific measurements provide a rigorous method for obtaining information about the physical world. No matter how sophisticated the measurement, though, there are inherent limitations on how narrowly it can characterize the value of a quantity of interest. If we fail to account for those limitations, then any inferences we make based on the results of measurement, whether or not they ultimately prove correct, are themselves inherently flawed. In the context of the criminal justice system, such flawed inferences can result in the innocent being deprived of their liberty and the guilty being exonerated. This not only undermines the integrity of the system and the verdicts it produces, but ultimately our belief that we are part of a fundamentally fair and just society.

Measurement uncertainty is important because it provides an explicit scientific statement of the limitations governing the rational inferences that can be made based upon the result of a particular scientific measurement. If we ignore this statement, then we are ignoring that which makes the measurement scientifically rigorous, thereby defeating the purpose for our reliance on it in the first place. Although the truth of a thing can never be known with absolute certainty, by adhering to the basic principles of science when using measurement as a tool in our search we can at least ensure that the strength of our belief in a proposition is supported by the available information.

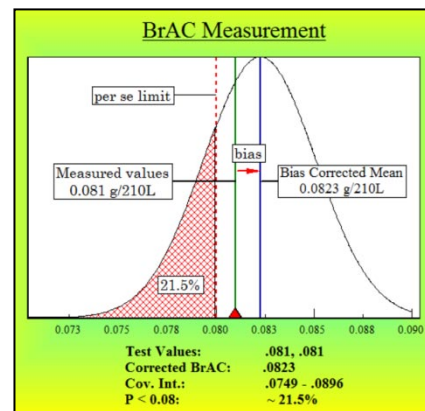
In the end, we are requiring no more of forensic science than we are of the juror charged with the task of making a determination of guilt or innocence in the context of the necessarily imperfect and incomplete information provided at trial. No reasonable juror would ever claim to know the truth with absolute certainty, nor does our system require them to do so. Their charge is simply to consider the information provided, and only that provided, and determine whether their degree of belief concerning the question of guilt is strong enough to establish the proposition beyond a reasonable doubt. As in the measurement context, though, the integrity of the determination is only as good as the information it is based on. By supplying jurors with a measurement’s uncertainty, we empower them to make rational inferences and determinations. Failing to do so calls into question everything our society, our constitution and simple fairness asks them to do. And for that we all suffer.

Only when we recognize with the same fearless honesty of a child, that a measurement result unaccompanied by its uncertainty is no less naked than the Emperor who wore no clothes, can we begin to have confidence in verdicts based on the results of forensic measurements and in the basic integrity of justice rendered thereupon.”⁴⁰

```

--- BREATH ANALYSIS ---
BLANK TEST          .000
INTERNAL STANDARD   VERIFIED
SUBJECT SAMPLE      .081
BLANK TEST          .000
EXTERNAL STANDARD   .080
BLANK TEST          .000
SUBJECT SAMPLE      .081
BLANK TEST          .000
  
```

=



TOOLS & CONCEPTS

Basic Measurement Concepts

Measurement: A set of empirical operations carried out to determine the quantity values that can reasonably be attributed to a quantity of interest. The objective of a measurement is to determine the value of the particular quantity being measured.

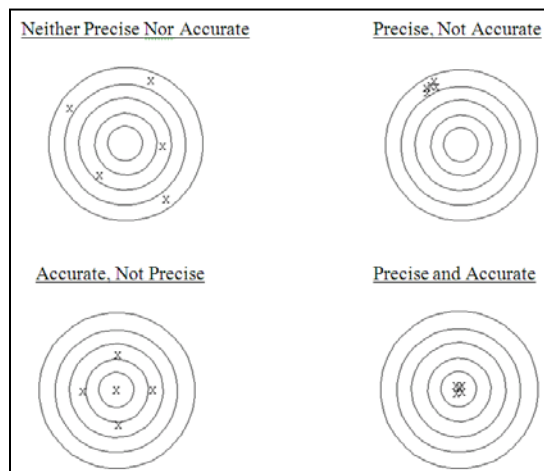
Measurand: The quantity intended to be measured.⁴¹

Direct Measurement: A measurement that senses the quantity of interest itself and maps it to a quantity value without the necessity of intermediate determinations.

Indirect Measurement: The determination of a quantity of interest through its relationship to other directly measured quantities.

Accuracy: The degree of agreement of a measured value with the “true” value of the quantity of interest. The degree of agreement expected from a measurement method/instrument is typically determined by comparing the mean of a set of measurements of a reference standard to the accepted value of the reference standard. Whether a measurement or instrument/method is deemed accurate is not an absolute judgment. Rather, accuracy is judged with respect to the use to be made of the data. What might be deemed accurate in one set of circumstances may not be accurate in another.

Precision: Precision is concerned with the variability or scatter of the individual results of replicate measurements. Measurements that are tightly grouped are considered precise while those with greater scatter are less so. As was the case with accuracy, precision is judged with respect to the use to be made of the data. What may be considered precise for one purpose may not be precise for another.



Measurement Interpretation – I: If a measurement value is to be interpretable, we must have an understanding of *how accurate* and *how precise* the measurement is. Absent such information, a measured value is simply a number, the meaning of which we know little about. Ideally, important measurements would be both accurate and precise. That is, not only would such measurements yield mean values in close agreement with a “true” value, but individual values having a high degree of agreement with each other.

An objective characterization of accuracy and precision are necessary in order to determine the value of the particular quantity being measured. Such objective characterization can be supplied by statistics.

Basic Statistical Concepts

Population: The entire set or universe of objects sharing specific traits defining a class of objects.

Sample: A subset of objects selected from the population.

Distribution: The set of possible values of a random variable related through their frequency of occurrence or belief based relative likelihood.

Parameter: A characteristic of a population’s distribution.

Statistic: A characteristic of a sample’s distribution.

Descriptive Statistics: Utilizes data to describe the properties of a sample, not to make predictions based upon it.

Inferential Statistics: Utilizes data to draw inference or make predictions. A typical example is the use of sample data to generate a sample statistic from which an inference concerning a population parameter may be made.

Probability – Frequentist Interpretation: Probability is interpreted as relative frequency of occurrence over all sample data sets. As such, probabilities are objectively determined as a function of sampling data. Population parameters have unique, fixed true values that are unknown. The randomness lies in the sampling process, not the parameter. Since population parameters are nonrandom, probability statements cannot be made about their values. Nor can probability statements be made about a characteristic of a unique event. The parameter or characteristic either is or is not a particular value. The level of confidence associated with an inference refers to the confidence in the sampling/inferential process, *not* the actual quantity of interest. It tells us how often, over repeated samplings, our inference will happen to correspond to the true value.

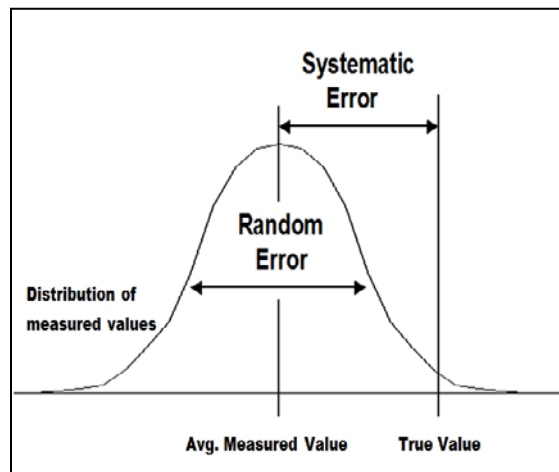
Probability – Bayesian Interpretation: Probability is interpreted as an information-based “degree of belief” that an event will occur. Bayesian inference employs sampling data and any other information deemed relevant in the decision making process so that probability (degree of belief) may be based upon both objective and subjective components. In this framework, the parameters themselves are considered random so that probability statements can be made directly about their values. The same holds for a characteristic of a unique event. Thus, probability statements made concerning the value of a parameter or characteristic *are* about the actual quantity of interest. It tells us the probability that this particular inference is “true”.

Measurement Error

Measurement Error: Traditionally, the quality of a measurement result was addressed through error analysis. This approach considered each measurand as having a unique true value. “The objective of measurement in the Error Approach is to determine an estimate of the true value that is as close as possible to that single true value. The deviation from the true value is composed of random and systematic errors.”⁴²

Systematic Error: The tendency of a set of measurements to consistently (on average) underestimate or overestimate the “true” value of the measurand by a given value or percentage. Most measurements have some amount of systematic error associated with them. Systematic error may be related to measuring methods, instruments or even empirically based calculations. It is a primary component of accuracy as it has a direct and regular impact on the degree of agreement of a measured value with the “true” value of the quantity of interest. Accordingly, “if a systematic error has not been accounted for, all [measured] values could be misleading.”⁴³ Fortunately, once identified systematic error can be corrected for. “It is assumed that the result of a measurement has been corrected for all recognized significant systematic effects and that every effort has been made to identify such effects.”⁴⁴

Random Error: The unpredictable/random fluctuation in measurement results under fixed conditions. Random error is associated with precision. Unlike systematic error, random error cannot be corrected for. It is an inherent aspect of all measurement results. Although random error cannot be completely eliminated, it can be minimized by making a large number of measurements.



Arithmetic mean: This is a simple average of measurement values. It is determined by adding all measured values together and then dividing the sum by the number of values included in the sum. It is typically used when all measured values are considered to be equally reliable.

$$= \frac{1}{n} \sum_{i=1}^n x_i$$

Bias: Quantitative measure of systematic error. Bias is typically treated as either having a constant magnitude across a range of measured values or being proportional to the measured value obtained. When proportional, the bias is commonly reported as a percent bias. For chemical measurements, it is not uncommon for the bias to be proportional to measured values. “Whenever the true value of the measured quantity is needed...bias can be a serious problem.”⁴⁵ Fortunately, once bias has been determined, systematic error can be easily accounted for. The bias of a method or instrument is ordinarily determined by comparing the mean of a set of measurements of a reference standard to its accepted value.

$$= \bar{x} - \mu$$

$$\% = \frac{\bar{x} - \mu}{\mu} \times 100$$

Standard Deviation: Quantitative characterization of the variability/dispersion of individually measured values about their mean. The standard deviation is the root mean square deviation of measured values from their mean. Precision/random error is typically expressed in terms of a standard deviation. The determination of the standard deviation varies slightly depending on the source of our data. If the standard deviation has been determined from a population, we use what is commonly referred to as a population standard deviation. On the other hand, when our data comes from a sample, we use what is commonly referred to as a sample standard deviation. Throughout the remainder of this section the distinction will not be noted unless necessary but it is assumed that whenever employed, the correct standard deviation is utilized.

$$= \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

$$= \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

Coefficient of Variation: The standard deviation expressed as a proportion relative to the mean of a set of measurements. The coefficient of variation can be useful when combining standard deviations or comparing the variability of separate measurements.

$$CV =$$

Bias Adjusted Mean/Best Estimate of True Value: The mean adjusted for bias. The bias adjusted mean is often considered the best estimate of the “true” value of the measurand. Whenever reporting the mean of a set of measurement, it should be corrected for bias. The correction applied depends upon whether the bias is constant or proportional.

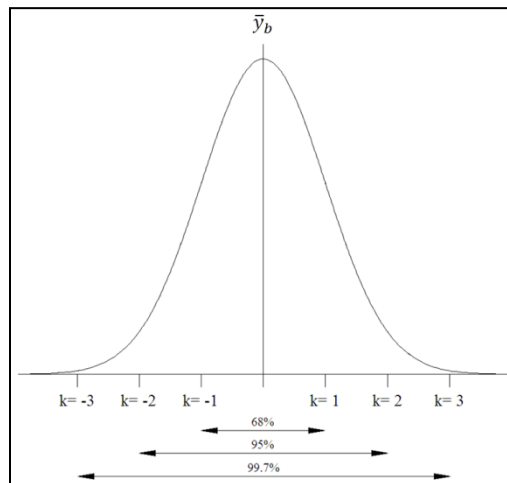
$$= -$$

$$= \frac{\quad}{1 + \%}$$

Confidence interval: A range of values symmetric about the bias adjusted mean constructed using a multiple of the standard deviation of the set of measurements and expected to cover the true value with a given level of confidence (likelihood).

$$\dots = \pm$$

The likelihood that the interval will overlap the true value is determined by the multiplier of the standard deviation (), known as a coverage factor, and the underlying distribution. If the underlying distribution is Gaussian (normal) the likelihood associated with = 1, 2 & 3, is given in the following figure.



One should be very careful with the interpretation of a confidence interval. The focus of the level of confidence is not the true value. That is, the level of confidence does not refer to the probability that the true value lies within the interval. It either does or does not. Rather, the subject of the level of confidence is the sampling procedure. It tells you that based upon the procedure utilized, you will be able to construct an interval that will overlap the true value a given percent of the time. In technical terms, “[t]he confidence reflects the proportion of

cases that the confidence interval would contain the true parameter value in a long series of repeated random samples under identical conditions.”⁴⁶ The confidence interval is based upon frequentist philosophy and the existence of a singular true value.

Standard deviation (error) of the mean: Quantitative characterization of the variability/dispersion of sample means. Due to the Central Limit Theorem, the following relationship holds regardless of the underlying population distribution as long as the sample size is large enough.

$$= \frac{\sigma}{\sqrt{n}}$$

Weighted mean: The weighted mean is an alternative way to determine the best estimate of the true value of a measurand. When combining multiple values determined for a given measurand, a weighted mean attaches more weight to those values considered more reliable.

$$= \frac{\sum w_i x_i}{\sum w_i}$$

Traditional weighted mean: Frequently, the values to be combined are the arithmetic means from several sets of measurements. The traditional weighted mean relies upon the precision associated with each set of measurements to determine the weight to accord the mean associated with each set. The greater the precision associated with a given mean, the more confidence we have in the value, and the more weight it is accorded in combining the means to determine a best estimate of the true value. In this case the above expression becomes:

$$= \frac{\sum \frac{x_i}{\sigma_i^2}}{\sum \frac{1}{\sigma_i^2}}$$

The weighted mean should be employed when the values to be combined are not equally reliable.

Standard deviation of the Traditional Weighted Mean:

$$= \frac{1}{\sqrt{\sum \frac{1}{\sigma_i^2}}}$$

Measurement Interpretation – II: If a measurement value is to be interpretable, we must have a *quantitative* determination of the *systematic* and *random error* associated with the measurement. Absent such information, a measured value is simply a number, the meaning of which we know little about. It has long been understood that no measurement result can be interpreted where only the value of the measurement itself is reported. Proper interpretation of a measured value requires knowledge and incorporation of the measurement’s systematic and random error into any reported values.

Unfortunately, as useful as traditional error analysis is, “[i]t is now widely recognized that, when all of the known or suspected components of error have been evaluated and the

appropriate corrections have been applied, there still remains an uncertainty about the correctness of the stated result, that is, a doubt about how well the result of the measurement represents the value of the quantity being measured.”⁴⁷ Put simply, it is not possible to know the true value of a measurand or the error of a measurement result and hence how close a measurement result is to the true measurand value.⁴⁸

Measurement Uncertainty

Measurement Uncertainty: “[F]or a given measurand and a given result of measurement of it, there is not one value but an infinite number of values dispersed about the result that are consistent with all of the observations and data and one’s knowledge of the physical world, and that with varying degrees of credibility can be attributed to the measurand.”⁴⁹ Accordingly, “[t]he objective of measurement in the Uncertainty Approach is not to determine a true value as closely as possible. Rather, it is assumed that the information from measurement only permits assignment of an interval of reasonable values to the measurand, based on the assumption that no mistakes have been made in performing the measurement.”⁵⁰

Contrary to the traditional approach, then, the measurand is not treated as having a unique “true” value. Instead, the measurand is deemed to consist of a set of “true” values. Measurement uncertainty is a quantitative statement characterizing the dispersion of values that can actually and reasonably be attributed “to a measurand based on the information available including systematic and random effects...and any other factors that may impact the measurement or test process or result.”⁵¹ Measurement uncertainty is based upon the Bayesian notion of probability as a measure of degree of belief.

Standard Uncertainty: The total uncertainty associated with any measurement result is typically the result of the combination of several smaller uncertainties associated with particular aspects of the measurement process. Each component of uncertainty that contributes to the uncertainty of a measurement result is known as a standard uncertainty. Each standard uncertainty is expressed and treated as, and may in fact be, a standard deviation.

Relative Standard Uncertainty: The standard uncertainty expressed as a proportion relative to the mean of a set of measurements. It can be useful when combining standard uncertainties or comparing the uncertainty of separate measurements.

$$= \frac{\quad}{\quad}$$

Type A Uncertainty: Component of uncertainty that has been determined by the statistical analysis of measured values. Determination is based on frequency distributions and any statistically valid method for data analysis. An example is the standard deviation determined from a set of measurements.

Type B Uncertainty: Component of uncertainty that has been determined by means other than the statistical analysis of measured values. Determination assumes *a priori* distributions based on relevant information and scientific judgment. Examples include information provided by instrument manufacturer, metrological certifications and reference publications.

Combined Uncertainty: The combination of all the standard uncertainties associated with a measurement. The individual standard uncertainties are combined in the same manner as standard deviations. Assuming the standard uncertainties are random and independent, the combined uncertainty is the root sum square of the standard uncertainties. The combined uncertainty is expressed and treated as, and may in fact be, a standard deviation.

$$=$$

When determining the combined uncertainty of a measurement it is critical to include all significant components of uncertainty. Failure to do so will cause an underestimate of the uncertainty misleading others to believe that the result is more precise than it actually is.

Expanded Uncertainty: Obtained by multiplying the combined uncertainty by a coverage factor.

$$=$$

A coverage factor is chosen such that when the expanded uncertainty is expressed as part of a complete measurement result it conveys a range of values that can actually and reasonably be attributed to a measurand with a given level of confidence. The level of confidence associated with a given coverage factor is determined by the measurement's underlying distribution. If the underlying distribution is Gaussian (normal) the level of confidence associated with $k = 1.64, 1.96$ & 2.576 , is given in the following table.

k	level of confidence
1.64	90%
1.96	95%
2.576	99%

Measurement Result: "In general, the result of a measurement is only an approximation or estimate of the value of the specific quantity subject to measurement, that is, the measurand, and thus the result is complete only when accompanied by a quantitative statement of its uncertainty."⁵² Moreover, "[i]t is assumed that the result of a measurement has been corrected for all recognized significant systematic effects and that every effort has been made to identify such effects."⁵³ Accordingly, a complete measurement result consists of the best estimate of the true value of the measurand, typically the bias adjusted mean, accompanied by the expanded uncertainty and its associated level of confidence.

$$= \pm (99\%)$$

This is interpreted to mean that the best estimate of the value attributable to the measurand is \bar{y} , and that $\bar{y} - U$ to $\bar{y} + U$ is the range of values that could actually be attributed to with a 99% level of confidence.

Coverage Interval: An “interval containing the set of true quantity values of a measurand with a stated probability, based on the information available.”⁵⁴ Ordinarily the coverage interval is derived from the expanded uncertainty and is symmetric about the mean so that it can be expressed as:

$$= \pm (99\%)$$

Note that the coverage interval is identical to the measurement result. Unlike the confidence interval, the coverage interval is based upon Bayesian philosophy so that it refers directly to the quantity of interest, the “true” value of the measurand. In this context, the level of confidence is the probability, understood as a degree of belief, “that the set of true quantity values of a measurand is contained within a specified coverage interval.”⁵⁵ It should also be noted that the coverage interval need not be symmetric about the mean.

Measurement Interpretation – III: For even the most carefully performed measurement, a unique “true” value for a measurand can never be determined. All that can ever be given is a set of values, all of which may actually and reasonably be assigned as “true” values. If a measurement value is to be interpretable, it must be *corrected for bias* and accompanied by a *quantitative estimate of its uncertainty*. Absent such information, a measured value is simply a number, the meaning of which we know little about.

“Knowledge of the uncertainty associated with measurement results is essential to the interpretation of the results. Without quantitative assessments of uncertainty, it is impossible to decide whether observed differences between results reflect more than experimental variability, whether test items comply with specifications, or whether laws based on limits have been broken. Without information on uncertainty, there is a risk of misinterpretation of results. Incorrect decisions taken on such a basis may result in unnecessary expenditure in industry, incorrect prosecution in law, or adverse health or social consequences.”⁵⁶

Functional Relationships, Measurement Functions and Propagation of Uncertainty

Algorithmic Determinations: When the quantity of interest cannot be measured directly, we must rely upon mathematical relationships between the quantity of interest and other measured and/or “given” values to calculate the quantity of interest. Each measured value and many “given” values have uncertainty associated with them. These uncertainties propagate through the calculation and are imparted to the value determined for the quantity of interest.

Measurement Function – General Form: A functional relationship between the quantity of interest and the input quantities (measured and/or “given” values) needed to calculate it.

$$= (,)$$

Best Estimate of True Value – General Form: The best estimate of a quantity value based on a measurement function is given by plugging in the best estimate for each of the input quantities

$$= (,)$$

Propagation of Uncertainty – General Form: For a quantity value based upon a general measurement function:

$$\begin{aligned} & \text{All Circumstances} \\ & \text{-----} \\ & = \text{---} \cdot \text{---} + 2 \text{---} \cdot \text{---} \\ & \text{-----} \\ & = \frac{(\text{---})}{\text{---}} \cdot \text{---} + \frac{(\text{---})}{\text{---}} \cdot \text{---} + 2 \cdot \frac{(\text{---})}{\text{---}} \cdot \frac{(\text{---})}{\text{---}} \\ & \quad \frac{(\text{---}, \text{---})}{\text{---}} \cdot \text{---} + \frac{(\text{---}, \text{---})}{\text{---}} \cdot \text{---} \end{aligned}$$

Independent Input Quantities

$$\begin{aligned} & \text{-----} \\ & = \text{---} \\ & \text{-----} \\ & = \frac{(\text{---}, \text{---})}{\text{---}} \cdot \text{---} + \frac{(\text{---}, \text{---})}{\text{---}} \cdot \text{---} \end{aligned}$$

Covariance: A measure of the association between two random variables. If two input quantities are independent then the covariance will be zero. When two input quantities are not independent this term appears in the propagation of uncertainty calculation to account for the dependence.

$$= \frac{1}{\text{---}} \cdot (\text{---}) (\text{---})$$

Measurement Function – Measured quantity multiplied by a constant:

$$\begin{aligned} & = \text{---} \cdot \text{---} \\ & = \text{---} \\ & = \text{---} \end{aligned}$$

Measurement Function – Variable raised to a constant power:

$$=$$

$$=$$

$$= \frac{\overline{|\quad|}}{\quad} = \frac{|\quad|}{\quad}$$

Measurement Function – Sums and differences:

$$= - + +$$

$$= - + +$$

$$= \frac{\text{Independent}}{+ + +}$$

$$\frac{\text{All Circumstances}}{+ + +}$$

Measurement Function – Products and quotients:

$$= \frac{\times \times}{\times \times}$$

$$= \frac{\times \times}{\times \times}$$

$$\frac{\text{Independent}}{\quad}$$

$$= \frac{\quad}{\quad} = \frac{- + - + + - + -}{\quad}$$

$$\frac{\text{All Circumstances}}{\quad}$$

$$= \frac{\quad}{\quad} \frac{\quad}{\quad} + \frac{\quad}{\quad} + \quad + \frac{\quad}{\quad} + \frac{\quad}{\quad}$$

Examples: BAC Results & Calculations

Breath Testing: Like any other measurement, forensic breath alcohol concentration tests have both bias and uncertainty associated with them. Both need to be determined and incorporated into a complete test result.

Best estimate for “true” BrAC (Bias adjusted mean): ()

Best estimate for “true” BrAC determined by computing the bias adjusted mean.

Machine bias: ()

Determined during calibration and which will be deemed proportional to the concentration measured.

Interferent bias: ()

Most breath test instruments are designed to detect the presence of interferences on an individual's breath. However, some are programmed such that they will only do so if the interference exceeds a particular level. There are several ways one might try to determine the average impact/bias such interference will have on a breath test below the level of detection but which will nonetheless contribute to the reported value. One could consult the literature to determine if there are published values. Another way is to postulate an underlying distribution based upon all the known information and determine the mean (expected) contribution due to bias based on the distribution. The bias due to this source will be a constant offset.

Best estimate for "true" BrAC (Bias adjusted mean):

$$= \frac{\quad}{1 + \quad} -$$

Combined uncertainty for BAC based on measurement:⁵⁷ ()

There are several sources of uncertainty that may be associated with a breath test. The ones considered here comprise only a subset and may or may not be relevant to your test. For ease of illustration they are treated as being independent.

Reference material: ()

The standard uncertainty associated with the reference material utilized to calibrate machine.

Machine precision: ()

The precision of the breath test machine determined at the time of its calibration and expressed as a standard uncertainty.

Bias: ()

The standard uncertainty associated with the value determined for the bias.

Sampling: ()

The standard uncertainty due to circumstances arising during the collection of breath samples.

Combined uncertainty for BAC based on measurement:

$$= \sqrt{\quad^2 + \quad^2 + \quad^2 + \quad^2 + \quad^2 + \quad^2}$$

Complete Result:

$$= \pm$$

Breath as an indirect measure of blood: When the breath alcohol concentration is being utilized as an indirect measure of blood alcohol concentration, the breath result must be converted into one for blood. This involves a conversion factor () between breath and blood alcohol concentration which the literature illustrates has a great deal of uncertainty associated with it. This uncertainty must also be factored into a reported result.

Functional Relationship:

$$= .$$

Combined Uncertainty:

$$= \frac{(\quad, \quad)}{\quad} + \frac{(\quad, \quad)}{\quad} + 2 \cdot \frac{(\quad, \quad)}{\quad} \cdot \frac{(\quad, \quad)}{\quad}$$

$$= (\quad \cdot \quad) + (\quad \cdot \quad) + 2 \cdot \quad \cdot \quad$$

Complete Result:

$$= \pm$$

Widmark's Formula: For the determination of blood alcohol content based on the number of drinks consumed.

Functional Relationship 1:⁵⁸ Assuming post-absorptive.

$$= \text{---} -$$

The Variables:

BAC at time t
 Number of drinks
 Density of alcohol
 Ethanol per drink

Body weight
 Volume of distribution
 Alcohol elimination rate
 Time since drinking began

Combined Uncertainty:⁵⁹

$$= \frac{\dots}{\dots} + 2 \cdot \frac{\dots}{\dots} \cdot \frac{\dots}{\dots}$$

$$= \frac{\dots}{\dots} + \frac{\dots}{\dots} + \frac{\dots}{\dots} + \frac{\dots}{\dots} + \frac{\dots}{\dots} + \frac{\dots}{\dots} + \frac{\dots}{\dots} + \frac{\dots}{\dots} + 2 \cdot \frac{\dots}{\dots} \cdot \frac{\dots}{\dots}$$

Complete Result:

$$= \frac{\dots}{\dots} \pm \dots$$

Functional Relationship 2: Accounting for rate of absorption.

$$= \frac{(1 - \dots)}{\dots}$$

The Variables:

- | | |
|-------------------------|---------------------------|
| BAC at time t | Body weight |
| Number of drinks | Volume of distribution |
| Density of alcohol | Alcohol elimination rate |
| Ethanol per drink | Time since drinking began |
| Alcohol absorption rate | |

Combined Uncertainty:⁶⁰

$$= \frac{\dots}{\dots} + 2 \cdot \frac{\dots}{\dots} \cdot \frac{\dots}{\dots}$$

$$= \frac{(1 - \dots)}{\dots} + \frac{\dots}{\dots} + \frac{\dots}{\dots} + \frac{\dots}{\dots} + \frac{\dots}{\dots} + \frac{\dots}{\dots} + \frac{\dots}{\dots} + \frac{\dots}{\dots} + 2 \cdot \frac{\dots}{\dots} \cdot \frac{\dots}{\dots}$$

Complete Result:

$$= \frac{(1 - \dots)}{\dots} \pm \dots$$

DOCUMENTS & DATA
FOR EXERCISES

UNCERTAINTY DECISIONS

1. *City of Kent v. McDaniel*, No. K81862 *Order Suppressing Defendant's Breath Alcohol Measurements in the Absence of a Measurement for Uncertainty* (Kent Muni. Ct. WA – 5/4/11)(counsel Ted Vosk).
2. *People v. Jabrocki*, No. 08-5461-FD (79th Dist. Ct. Mason Co. MI – 5/6/11)(counsel Michael Nichols).
3. *State v. Barnes*, No. 9Y6304935 *Court's Ruling on Motion to Suppress BAC* (Clark Co. Dist. Ct. WA – 12/16/10)(counsel Ted Vosk; co-counsel Nicole Dalton).
4. *State v. Fausto*, No. C076949 *Order Suppressing Defendant's Breath Alcohol Measurements Under in the Absence of a Measurement for Uncertainty* (King Co. Dist. Ct. WA – 9/20/10)(counsel Ted Vosk; co-counsel Kevin Trombold & Andy Robertson).
5. *State v. Herbolzheimer*, No. XY181244 *Order Denying Motion to Suppress* (Skagit Co. Dist. Ct. WA – 6/4/10)(counsel Ted Vosk; co-counsel Kevin Trombold & Roy Howson).
6. *State v. Weimer*, No. 7036A-09D *Memorandum Decision on Motion to Suppress* (Snohomish Co. Dist. Ct. WA – 3/23/10)(counsel Ted Vosk; co-counsel Eric Gaston).

UNCERTAINTY STANDARDS FREE ONLINE

1. BIPM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)* (2008)(http://www.bipm.org/utis/common/documents/jcgm/JCGM_100_2008_E.pdf).
2. BIPM, *Evaluation of measurement data — An introduction to the “Guide to the expression of uncertainty in measurement” and related documents* (2009)(http://www.bipm.org/utis/common/documents/jcgm/JCGM_104_2009_E.pdf).
3. NIST, *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results* – NIST TN 1297 (1994) (<http://physics.nist.gov/Pubs/guidelines/TN1297/tn1297s.pdf>).
4. EURACHEM, *Quantifying Uncertainty in Analytical Measurement* CG-4 (2000)(<http://www.eurachem.org/guides/pdf/QUAM2000-1.pdf>).
5. EURACHEM, *Measurement uncertainty arising from sampling: A guide to methods and approaches* (2007)(http://www.eurachem.org/guides/pdf/Ufs_2007.pdf).

RECOMMENDED BOOKS

1. Gullberg, *Statistical Applications in Forensic Toxicology*, Medical-Legal Aspects of Alcohol, Ch. 18 (James Garriott ed., 5th ed. 2009).
2. Kirkup, L. *An Introduction to Uncertainty in Measurement* (2006).
3. Vosk, *Measurement Uncertainty*, Encyclopedia of Forensic Sciences (Elsevier – *In Press*).
4. Vosk, *Measurement Uncertainty: Forensic Metrology and Result Interpretation Part I & II*, Understanding DUI Scientific Evidence, Chapters 7 & 8 (4th ed. Aspatore, 2011).
5. Vosk, *Uncertainty in Forensic Breath Alcohol Testing*, Intoxication Test Evidence, Ch. 56 (E. Fitzgerald ed., 2nd ed. 2009).
6. Vosk, *Computational Aspects of Measurement Uncertainty in Washington State Breath Alcohol Tests*, Wash. Prac., DUI Practice Manual Appendix (Callahan ed., *In Press*).
7. Vosk & Emery, *Forensic Metrology: A Primer on Scientific Measurement for Lawyers, Judges and Forensic Scientists* (CRC Press – *In Preparation* (due out summer 2012)).

RECOMMENDED PAPERS/ARTICLES

1. Ehrlich, C., Dybkaer, R., and Wöger, W., *Evolution of philosophy and description of measurement (preliminary rational for VIM 3)* 12 ACCRED. QUAL. ASSUR. 201 (2007).
2. Estler, W.T., *Measurement as Inference: Fundamental Ideas* 48(2) ANNALS CIRP 611 (1999).
3. Gullberg, R. G., *Estimating the measurement uncertainty in forensic breath-alcohol analysis*, 11 ACCRED. QUAL. ASSUR. 562 (2006).
4. Gullberg, *Breath Alcohol Measurement Variability Associated with Different Instrumentation and Protocols*, 131(1) FORENSIC SCI INT. 30 (2003)
5. Jones, A. W., *Dealing with Uncertainty in Chemical Measurements*, 14(1) NEWSL. OF THE INT. ASSOC. FOR CHEM. TEST. 6 (2003).
6. Kacker, *Evolution of modern approaches to express uncertainty in measurement* 44 METROLOGIA 513 (2007).
7. Kristiansen, J. and Petersen, H.W., *An Uncertainty Budget for the Measurement of Ethanol in Blood by Headspace Gas Chromatography*, 28(6) J. ANAL. TOX. 456 (2004).
8. Vosk, *Trial by Numbers: Uncertainty in the Quest for Truth and Justice* 56 THE NACDL CHAMPION 54 (Nov. 2010)(reprinted with permission in 40(3) THE VOICE FOR THE DEFENSE 31 (April 2011)).

Author's Bio

TED VOSK was kicked out of his house as a teenager and lived on the streets, homeless, for the better part of four years. Eventually tiring of his situation, he gained admittance to Eastern Michigan University where he earned the status of national Goldwater Scholar before graduating with honors in Theoretical Physics and Mathematics. Thereafter Ted entered the Ph.D. program in Physics at Cornell University before moving on to Harvard Law School where he obtained his JD. Ted is currently a legal and science writer, criminal defense attorney, and legal/forensic consultant. A member of the American Academy of Forensic Sciences and Mensa, he has written, broadcast, presented and taught around the country on topics ranging from string theory to the doctrine of constitutional separation of powers. He has been part of the most significant DUI cases in Washington State over the past decade. For this work, he has received the President's Award from the Washington Association of Criminal Defense Lawyers and the Certificate of Distinction from the Washington Foundation for Criminal Justice. He has been published in multiple legal periodicals and treatises on DUI law and is currently authoring a textbook on Forensic Metrology. Ted lives in Washington with his wife Kris. Whether it's chasing down active volcanoes, traveling with friends or simply sharing a sunset in the mountains, the two of them strive to suck every last ounce of marrow out of life. Quiet moments are often spent with their dog and cats simply enjoying the wonder of it all.

¹ Richard Feynman, *The Meaning of it All* 15 (Perseus Books 1998).

² Richard Feynman, *The Character of Physical Law* 165-166 (MIT Press 1965).

³ BIPM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 3.1.2, Appendix D.4 (2008); Kirkup, *An Introduction to Uncertainty in Measurement* 33 (Cambridge University Press 2006); NIST, *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*, NIST TN 1297 § 2.1 (1994); Kristiansen, *An Uncertainty Budget for the Measurement of Ethanol in Blood by Headspace Gas Chromatography*, 28(6) J. ANAL. TOX. 456 (2004).

⁴ BIPM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 5.2 (2008); BIPM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 0.1 (2008).

⁵ Estler, *Measurement as Inference: Fundamental Ideas* 48(2) ANNALS CIRP 611, 623 (1999)

⁶ BIPM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 2.2.3 (2008); BIPM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.26 (2008); EURACHEM, *Quantifying Uncertainty in Analytical Measurement* CG-4, § 2.1.1 (2000); Hofmann, *Common Sources of Error in Measurement*, Handbook Of Measuring System Design V.1, p. 289, 293 (Sydenham & Thorn ed., 2005).

⁷ Ehrlich, *Metrological Timelines in Traceability*, 103 J. Res. Natl. Inst. Stand. Technol. 93, 94 (1998).

⁸ Desimoni, *About considering both false negative and false-positive errors when assessing compliance and non-compliance with reference values given in compositional specifications and statutory limits*, 13 ACCRED. QUAL. ASSUR. 653, 653 (2008); BIPM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 0.1 (2008); Richter, *Reporting measurement uncertainty in chemical analysis*, 13 ACCRED. QUAL. ASSUR. 113, 113 (2008); Watson, *Pharmaceutical Analysis - A Textbook for Pharmacy Students and Pharmaceutical Chemists*, 2 (2nd ed. Elsevier 2005).

⁹ ISO, *Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation*, ISO/TS 21748 DRAFT REVISION, v (2009); ISO, *Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation*, ISO/TS 21748, v (2004).

¹⁰ NIST *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*, NIST TN 1297 § 2.1 (1994); BIPM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 7.1.4 (2008).

¹¹ Bono, *ISO/IEC 17025:2005: Section 5.4.6: Estimation of Uncertainty – Is Anyone Certain What This Means?* p.7, Presentation at the 61st Annual Meeting of the American Academy of Forensic Sciences (2/17/2009).

¹² NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 116-117 (2009); Kristiansen, *An Uncertainty Budget for the Measurement of Ethanol in Blood by Headspace Gas Chromatography*, 28(6) J. ANAL.

TOX. 456 (2004); Jones, *Dealing with Uncertainty in Chemical Measurements*, 14(1) NEWSL. OF THE INT. ASSOC. FOR CHEM. TEST. 6 (2003).

¹³ Bono, *ISO/IEC 17025:2005: Section 5.4.6: Estimation of Uncertainty – Is Anyone Certain What This Means?* p.7, Presentation at the 61st Annual Meeting of the American Academy of Forensic Sciences (2/17/2009).

¹⁴ NAS, 117.

¹⁵ Bich, *Interdependence between measurement uncertainty and metrological traceability* ACCRED. QUAL. ASSUR. (IN PRESS - 2009).

¹⁶ All graphic illustrations were created utilizing ProStat Software.

¹⁷ Feynman 165-166 (1965).

¹⁸ Vosk, *Measurement Uncertainty: Forensic Metrology and Result Interpretation*, Understanding DUI Scientific Evidence (4th ed. *In Press*).

¹⁹ NIST 1297 App. D 1.1.1 note 2 (1994).

²⁰ VIM § 0.1 (2008).

²¹ Kacker 517 (2007).

²² GUM § 3.2.2 n.2 (2008).

²³ GUM App. D.5.2 (2008).

²⁴ Estler (1999).

²⁵ GUM § 2.3.5 (2008).

²⁶ ISO, *Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation*, ISO/TS 21748 DRAFT REVISION, v (2009); ISO, *Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation*, ISO/TS 21748, v (2004).

²⁷ VIM § 2.36 (2008).

²⁸ NIST 1297 § 2.1 (1994); GUM § 3.1.2 (2008).

²⁹ EURACHEM QUAM p.1, § 1.0 (2000).

³⁰ GUM App. E.2.1 (2008).

³¹ Vosk, *Measurement Uncertainty: Forensic Metrology and Result Interpretation*, Understanding DUI Scientific Evidence (4th ed. *In Press*).

³² Kacker 517 (2007).

³³ See also, Vosk, *Computational Aspects of Measurement Uncertainty in Washington State Breath Alcohol Tests*, Wash. Prac., DUI Practice Manual Appendix (Callahan ed., *In Press*); Vosk, *Trial by Numbers: Uncertainty in the Quest for Truth and Justice* 56 THE NACDL CHAMPION 54, Fn. 30 (Nov. 2010)(reprinted with permission in 40(3) THE VOICE FOR THE DEFENSE 31 (April 2011)). Determining the z-factor assumes a Gaussian distribution. The same method can be used with a t-distribution as long as we know the initial value of t utilized to create the expanded uncertainty.

³⁴ See also, Vosk, *Computational Aspects of Measurement Uncertainty in Washington State Breath Alcohol Tests*, Wash. Prac., DUI Practice Manual Appendix (Callahan ed., *In Press*); Vosk, *Uncertainty in Forensic Breath Alcohol Testing*, Intoxication Test Evidence, Ch. 56 (E. Fitzgerald ed., 2nd ed. 2009); Vosk, *Using the database and QAP to determine the likelihood that a test result greater than a .08 reflects a true BAC less than the per se limit*, § 11.7 Defending DUI's in Washington – 2008 Update (Cowan & Fox, 3rd ed. 2007).

³⁵ Bono 7 (2009).

³⁶ NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 184 (2009).

³⁷ NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 116-117 (2009).

³⁸ NAS, *Strengthening Forensic Science in the United States: A Path Forward*, 117 (2009).

³⁹ Kacker 517 (2007).

⁴⁰ Vosk & Emery, *Forensic Metrology: A Primer on Scientific Measurement for Lawyers, Judges and Forensic Scientists* (CRC Press – In Preparation).

⁴¹ JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.3 (2008).

⁴² JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 0.1 (2008).

⁴³ Les Kirkup, *An Introduction to Uncertainty in Measurement* 33 (Cambridge University Press 2006).

⁴⁴ BIPM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 3.2.4 (2008); NIST, *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*, NIST TN 1297 § 5.2 (1994).

-
- ⁴⁵ NIST, *NIST Special Publication 260-100*, 4 (1993).
- ⁴⁶ ISO, *Statistics — Vocabulary and symbols — Part 1: General statistical terms and terms used in probability*, ISO 3534-1, § 1.28 (2007).
- ⁴⁷ BIPM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 0.2, 3.2.2 – 3.2.3 (2008).
- ⁴⁸ BIPM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, §3.2.1, 3.3.2 (2008); Ehrlich, *Evolution of philosophy and description of measurement* 12 ACCRED. QUAL. ASSUR. 201, 210 (2007).
- ⁴⁹ BIPM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 5.2 (2008).
- ⁵⁰ JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 0.1 (2008).
- ⁵¹ JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.26 (2008); ASTM, *Standard Terminology Relating to Quality and Statistics*, E 456 § 3 (2008).
- ⁵² NIST *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*, NIST TN 1297 § 2.1 (1994); BIPM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 7.1.4 (2008).
- ⁵³ BIPM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*, § 3.2.4 (2008); NIST, *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*, NIST TN 1297 § 5.2 (1994).
- ⁵⁴ JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.36 (2008).
- ⁵⁵ JCGM, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)*, § 2.37 (2008).
- ⁵⁶ ISO, *Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation*, ISO/TS 21748 DRAFT REVISION, v (2009); ISO, *Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation*, ISO/TS 21748, v (2004).
- ⁵⁷ The example assumes that the uncertainties are independent.
- ⁵⁸ This example is taken from Gullberg, *Estimating the uncertainty associated with Widmark's equation as commonly applied in forensic toxicology* 172 FOR. SCI. INT. 33 (2007), and fills in some blanks not explicitly shown therein.
- ⁵⁹ Assuming only σ_1 and σ_2 are not independent.
- ⁶⁰ Assuming only σ_1 and σ_3 are not independent.