1 Review

- Specific weight $\Rightarrow \gamma = \rho g$

- Specific gravity $\Rightarrow S.G. = \frac{\rho}{\rho_{ref}}$

- Ideal gas law $\Rightarrow P = \rho R \Theta$

- Shear stress $\Rightarrow \tau = \mu \frac{du}{dz}$. Therefore Force/Area in a fluid with shear is equal to the absolute viscosity times the shear (strain rate).

Uncertainty of Experimental Data

2 Introduction

Kline & McClintock (1953) defined uncertainty as “A possible value the error might have” and this is the definition we still work with today. There are two fundamental types of experimental measurement – single-sample and multiple-sample measurements. Time and money permitting we would ideally make a measurement repeatedly, using different instruments, and then take a statistical view of the range of measurements that in theory should all be identical but in practice will not be. However, for most engineering experiments it is simply not practical to repeat an experiment enough times to provide uncertainty statistics. Even if multiple trials are feasible it generally is only possible with the same measurement equipment and sensors. Thus in the global sense, almost all experiments are single-sample. How might we ascribe the uncertainty interval to a measurement? If there are multiple independent measurements one possibility is in terms of the standard deviation, or perhaps the range as defined by the maximum value minus the minimum value. Kline & McClintock laid the foundation for single-sample uncertainty analysis and proposed that for single-sample uncertainty an uncertainty, or
confidence interval approach be taken. For the purposes of this class we will take the approach of Kline & McClintock to uncertainty analysis.

3 Basics

We begin with some definitions. We will use the term *variable* to mean a quantity being directly observed (e.g., measured) in the lab. This means if we are making a measurement of the flow depth by observing the pressure than we will consider pressure to be a variable. We will call depth a *result* – it is a value that is calculated from the basic variables by some sort of mathematical manipulation – in this case we would invoke our knowledge of statics to convert a pressure to depth by dividing by the specific weight of the fluid. Why is this distinction important? Because we will introduce error into our determination of the depth in two ways: (1) the error associated with measuring pressure and (2) the error associated with determining the specific weight of the fluid. Kline & McClintock give us a rationale way for combing the effects of these two errors as we will see shortly.

Consider a suite of variables under measurement which we will denote by $X_i$, $i = \{1, 2, 3, \ldots, I\}$ where $I$ is the number of variables in the suite. Let $\delta X_i$ be the uncertainty in a measured value of $X_i$ and let $\hat{X}_i$ be the measured value (or mean of $N$ measured values if the sample is multiple). We assume that the best estimate of $X_i$ is $\hat{X}_i$ and we can represent the uncertainty in our measurement as

$$X_i = \hat{X}_i \pm \delta X_i$$  \hspace{1cm} (1)

But would we mean by $\delta X_i$ – the uncertainty of course but quantitatively what do we mean? Had we sampled enough times to get a distribution of the $\delta X_i$ we would typically find a Gaussian distribution (but certainly not always!) In this case we could chose to set $\delta X_i$ equal to the standard deviation of the Gaussian distribution. This is a valid approach but not very flexible and also not very restrictive. The probability of a value landing between plus or minus one standard deviation of its mean is only about 68% or 2:1. A better method is to allow the experimenter to choose the probability level (or
odds) at which he or she would like to report the uncertainty interval. Hence, we choose to write

$$P[|X_i - \hat{X}_i| > \delta X_i] = \alpha$$

where $\alpha$ forms the $(1 - \alpha) \times 100\%$ confidence level (or $1/\alpha$:1 odds). For example, if we decide we want to report a 95% confidence interval (20:1 odds) we can express the above statement as $\hat{X}_i$ lies within $\pm \delta X_i$ of $X_i$ with at a confidence level of 95% (e.g., on average 1 in 20 times $\hat{X}_i$ will lie more than $\pm \delta X_i$ from $X_i$).

It is generally assumed that uncertainty statistics will be Gaussian although this assumption should be kept in mind. Recall that any Gaussian distribution can be written such that it has a variance of 1 and a mean of zero by taking $z = (x - \mu)/\sigma$, where $\mu$ is the mean of the distribution and $\sigma$ is the standard deviation, yielding the standard normal probability distribution function

$$f(z) = \frac{1}{\sqrt{2\pi}}e^{-z^2/2}$$

A useful question is at what magnitude of $z$ is the probability that $z$ falls outside the range bounded by $\pm$ this value equal to $\alpha$. If we denote the particular value of $z$ as $z_\alpha$, then we can write our statement as

$$P[-z_\alpha \leq z \leq z_\alpha] = \int_{-z_\alpha}^{z_\alpha} f(z) \, dz = 1 - \alpha$$

For example, if $\alpha = 0.05$ then $z_\alpha = 1.96$, if $\alpha = 0.01$ then $z_\alpha = 2.33$. We typically refer to these as the $(1 - \alpha)100\%$ levels, or in the case of the examples the 95% and 99% levels, respectively.

The Gaussian assumption and the fact that the distribution is two-sided leads to choosing $z_{\alpha/2} \times \sigma$ as the bounds on the confidence interval. Taking $\alpha = 0.05$, as is commonly done, leads to the 95% uncertainty interval (20:1 odds) and

$$\delta X_i = 1.96\sigma$$

where $\sigma$ is the standard deviation of the possible measurements.
4 The propagation of error

The result $R$ is generally calculated in some way from the data, $\hat{X}_i$. Thus we can write

$$ R = R(\hat{X}_1, \hat{X}_2, \ldots, \hat{X}_I) $$

Now, $R$ can be Taylor series expanded in the $\hat{X}_i$’s. If the variation in the $\hat{X}_i$ are small than a linear approximation (truncation of all non-linear terms in the Taylor series expansion) provides a reasonable approximation of the propagation of the uncertainty into the result from uncertainty in the data. Thus we can write

$$ \delta R_{\hat{X}_i} = \frac{\partial R}{\partial \hat{X}_i} \delta \hat{X}_i $$

The partial derivative terms are known as sensitivity coefficients. Kline & McClintock demonstrated that the method to combine the $\delta R_{\hat{X}_i}$ optimally is the root-sum-square (RSS) technique given by

$$ \delta R = \left[ \sum_{i=1}^{M} \left( \frac{\partial R}{\partial \hat{X}_i} \delta \hat{X}_i \right)^2 \right]^{1/2} $$

The fundamental assumptions in the above are

- The measurements are independent
- The probability density function of the $\hat{X}_i$ are Gaussian
- The uncertainty interval in each measurement is the same

From Eq 8 it is clear that small terms have little effect. Hence Eq 8 can be used to determine which data in an experiment most impact the uncertainty in a result and hence where energy and money should be spent to maximally reduce uncertainty.

4.1 Single-Sample Uncertainty in Stochastic Measurements

Most of the data reported in fluid mechanics is the result of the ensemble averaging process. The determination of a mean value or an RMS value is conducted only once, but
from a large data set with \( N \) samples in it. Moffat (1988) presents a formal methodology based on single and multiple sample analysis to determine the uncertainty in these type of measurements. He proposes splitting the data into a fixed or bias error and a variable or random error. The bias error is the fixed or slowly varying error that is deterministic, but generally not separable from the measurement (e.g., calibration error). Essentially bias errors have a fixed mean and zero variance. Random errors arise from many sources (e.g., thermal noise, electronic noise, etc.) and are characterized as having zero mean but non-zero variance. The two error types, in general, are best considered separately since if the interest is in gradients of the data (e.g., differences over some time or space scale), the bias errors are eliminated leaving only the random error.

In single-sample experiments one alternative to determine the bias and random errors is to perform an auxiliary experiment, typically with 30 replicates, to determine these components. This is described in Moffat. An alternative, slightly less formal method, which we will invoke in CEE 331, is to estimate the bias error in the variables at the 95% level. This may be done in a number of ways including a conservative ‘worst case’ estimate, for example the worst case error in the measurement in length using a millimeter graded ruler is 1 mm or manufactures data on error (e.g., the pressure sensors absolute error is reported as 1% of full scale). The determined bias errors are then propagated into a results bias error using the RSS technique (Eq 8. For the purposes of laboratory reports in this course you must always report a bias error at the 95% confidence interval.

This leaves the issue of random error. If we are interested in mean values, say the mean pressure when we are using a pressure gage to measure the static pressure to determine the water depth, we often sample tens or hundreds or even thousands of values and take the mean. If we determine the standard deviation of these samples we can than report the uncertainty interval in the random error based on the standard deviation of these samples, just as described above (e.g., we can report \( \pm 1.96 \sigma \) as the 95% confidence interval on the random error).

For those of you interested in a powerful tool for the determination of error from a suite
of measurements that does not assume the data distribution is Gaussian read on. This is optional for CEE 331 but it is not terribly difficult and offers an even more accurate method to determine errors from multiple measurements.

4.2 Bootstrap Uncertainty Analysis

The bootstrap algorithm is surprisingly simply yet surprisingly powerful. Due to its heavily computational nature it was not even conceived until well into the computer age. It was first proposed by Bradley Efron in 1977 and has spread widely through all literature – wherever statistics has previously been used. For a thorough treatment see Efron & Tibshirani (1993). Diaconis & Efron (1983) give widely varied examples of the application of bootstrap analysis.

The bootstrap algorithm is incredibly simple. Given a set of \( N \) measurements we often like to calculate statistics from the measurements - the mean and the variance being some of our favorites but there are myriad other possibilities. The bootstrap is a method to directly estimate the probability density function of any statistic we can imagine that can be calculated from the measurements.

For example, consider the mean. Normally we would calculate the mean and have a single-sample estimate of the mean. The bootstrap algorithm gives us a way to replicate the estimate into a multiple-sample estimate and hence to directly estimate the probability distribution function for the mean of the measurement. This is done by selecting a set of \( N \) samples from the original \( N \) with replacement. This means that any individual measurement may be repeated multiple times within a given set. Now we can calculate our statistic, the mean in this case, from the sample. If we repeat this process \( M \) times we can now estimate the probability distribution function for the statistic under consideration - the mean.

As an example consider the bootstrap replication (at \( M=1000 \) replicates) of the mean value of the number of heads that appeared on a toss of a coin 8 times by 15 people (same
example as in next section). The probability density function would look something like the figure on the top of the next page.

![Probability Density Function](image)

giving us the sense that the statistic under study is Gaussian and that an uncertainty interval can be defined such that the area under the probability density function is 0.95.

### 4.3 Bootstrap Percentile Technique

The uncertainty interval in a statistic can be directly estimated at a given level from the ordered statistic (e.g., sorted from lowest to highest value of the statistic). Efron & Tibshirani (1993) suggest $M=1000$ replications to be sufficient for the 95% uncertainty interval. Thus for $M=1000$ the bootstrap replicates of the statistic are determined and then sorted lowest to highest. The $25^{\text{th}}$ and the $975^{\text{th}}$ values of the ordered statistic are then taken as the lower and upper bounds, respectively, of the 95% uncertainty interval.

As an example consider if the 15 members of the class each flipped a coin 8 times as a way of estimating the number of times a head would occur for 8 coin tosses. If our result was:

$$[ 2 \ 2 \ 7 \ 4 \ 4 \ 3 \ 6 \ 3 \ 4 \ 4 \ 3 \ 4 \ 3 \ 7 \ 4 ]$$ (9)

we would find the mean is exactly 4, but what is the uncertainty in our estimate? If we turn to the bootstrap we can calculate an ordered statistic that looks like the figure on
where the 25th statistic is 3.27 and the 975th is 4.80. Our best estimate for the number of heads is still the mean value that we measured, 4, and hence we report the uncertainty interval as 4.00 +0.80, - 0.73. Note that the 1000 drawn bootstrap replicates are random so the exact uncertainty interval is a random process itself (and the bootstrap could be used to look at the 95% uncertainty interval on the bounds themselves, which is left as an exercise for the student!)

The bootstrap is easily programmed in a language such as C or FORTRAN, however, we are lucky – it is already in Matlab as the function *bootstrap*. The Matlab code to perform the above analysis is simply:

```matlab
x = [2 2 7 4 4 3 6 3 4 3 4 3 7 4];

y = sort(bootstrp(1000,'mean',x));
mean(x)
ULlo = y(25) - mean(x)
ULhi = y(975) - mean(x)
plot(y)
```
References


