# Uncertainties03.wxmx

TABLE OF CONTENTS



## 1 Preface

In Uncertainties03.wxmx we discuss averaging weighted measurements and error propagation, with an emphasis on the physical sciences.

Edwin L. (Ted) Woollett
https://home.csulb.edu/~woollett/
Aug. 21, 2024

## 2 References

Data Science Articles:
https://www.abs.gov.au/statistics/understanding-statistics/
statistical-terms-and-concepts
https://www.abs.gov.au/statistics/understanding-statistics/
statistical-terms-and-concepts/correlation-and-causation

https://www.khanacademy.org/test-prep/praxis-math/
praxis-math-lessons/
gtp--praxis-math--lessons--statistics-and-probability/
a/gtp--praxis-math--article--correlation-and-causation--lesson

https://www.scribbr.com/methodology/population-vs-sample/
https://www.scribbr.com/methodology/correlation-vs-causation/

[HH] Ifan Hughes, Thomas Hase, Uncertainties: A Practical guide to Modern Error Analysis, 2010, Oxford Univ Press.

R. J. Barlow, Statistics: A Guide to the Use of Statistical Methods in the Physical Sciences (Manchester Physics Series), 1993, Wiley

Gerhard Bohm, Günter Zech,
Introduction to Statistics and Data Analysis for Physicists, 3rd revised edition, 2017,
available from CERN webpage:
https://s3.cern.ch/inspire-prod-files-d/da9d786a06bf64d703e5c6665929ca01

Frederick James, 'Statistical Methods in Experimental Physics', 2nd ed.,
2006, World Scientific.
Louis Lyons, Statistics for Nuclear and Particle Physics, 1986,
Cambridge Univ. Press,

Luca Lista, 'Statistical Methods for Data Analysis in Particle Physics',
Lecture Notes in Physics 909, 2016, Springer-Verlag,

Glen Cowan , Statistical Data Analysis, Clarendon Press, Oxford, 1998

```
(%i6) load(draw)$
      set_draw_defaults(line_width=2, grid = [2,2], point_type = filled_circle,
           head_type = 'nofilled, head_angle = 20, head_length = 0.5,
           background_color = light_gray, draw_realpart=false)$
      load (descriptive)$  load (distrib)$
       fpprintprec : 5$ ratprint : false$
```

Homemade functions fll, head, tail, Lsum are useful for looking at long lists.

```
(%i10) fll ( aL) := [ first (aL), last (aL), length (aL) ]$
       head(L) := if listp (L) then rest (L, - (length (L) - 3) ) else
           error("Input to 'head' must be a list of expressions ")$
       tail (L) := if listp (L) then rest (L, length (L) - 3 )  else
            error("Input to 'tail' must be a list of expressions ")$
       Lsum (aList) := apply ("+", aList)$
```

# 3 Averaging Weighted Measurements

We follow Barlow, Sec. 4.2.2.

"Suppose you have a set of measurements {x_i} of some quantity μ and that these measurements have different errors σ_i. To combine the values you obviously want to form an average in such a way that the better measurements (i.e. those with small σ) are given more weight than the poorer, large σ, measurements. To give a specific example, suppose that a voltage has been measured as 3.11 +/- 0.02 volts by a meter whose resolution is known to be σ1 = 0.02 volts, and 3.13 +/- 0.01 volts by another, better, meter whose resolution is known to be σ2 = 0.01 volts. How can these be usefully combined to give a single result?"

"Well, if you had taken four measurements with the poorer, 0.02 volt resolution, meter, and averaged them, then the average would have had a precision of 0.02/√4 = 0.01 volts. Four poor measurements are equivalent to one measurement twice as good, i.e. with half the error. You can run this argument backwards: one of the good measurements is equivalent to four of the inferior ones, and should thus be given four times the weight. This gives a result of
<V> = (1/5) * 3.11 + (4/5) * 3.13 = 3.126 volts."

```
(%i11) 3.11/5 + 3.13*4/5;
(%o11) 3.126
```

In Maxima, arithmetic calculations are evaluated from left to right.

"Generalising, the weight you give to a result when averaging should be proportional to the inverse square of the resolution."

With <V> = w1*3.11 + w2*3.13, and with the requirement w1 + w2 = 1 (the sum of the weights is unity), taking w1 = a/σ1^2, w2 = a/σ2^2, constant a to be determined, we first find a:

```
(%i12) soln : solve (a/0.02^2 + a/0.01^2 = 1);
```

(soln) $[a=\frac{1}{12500}]$

We then find the weights w1 and w2 using our solution for the constant a:

```
(%i13) at ([a/0.02^2, a/0.01^2], soln);
(%o13) [0.2,0.8]
```

We then convert the decimal weight values to ratios of integers:

```
(%i14) rat (%);
```

(%o14)/R/ $[\frac{1}{5},\frac{4}{5}]$

When dealing with measurements of the same quantity x, each with an error σ_k, the correct average to form is

<x> = Σ w_k * x_k

with w_k = (1/σ_k^2) / Σ (1/σ_i^2).

What about the resolution of this average voltage <V> = 3.126 volts? One measurement with the σ = 0.02 volts meter plus four measurements with the σ = 0.02 volts meter is equivalent to one measurement with the σ = 0.02 volts meter plus one measurement with the σ = 0.01 volts meter. The resolution of the average should thus be 0.02/√5 = 0.009 volts.

We can generalize the calculation of the resolution of an average by using

σ_av = sqrt [ Var(<x>) ], and by using

Var(<x>) = 1 / Σ ( 1/σ_i)^2 ) for the variance of the average <x>.

For our example, Var (<V>) = 1 / [1/0.02^2 + 1/0.01^2] = 1 / [50^2 + 100^2]

= 1 / 12,500 = 8 x 10^(-5),

whose square root is approximately 0.009.

```
(%i15) sqrt ( 1/ ( 1/0.02^2 + 1/0.01^2));
(%o15) 0.0089443
```

## 3.1 A Word of Caution

"Averaging results, whether weighted or not, needs to be done with due caution and commensense. Even though a measurement has a small quoted error it can still be, not to put too fine a point on it, wrong. If two results are in blatant and obvious disagreement, any average is meaningless and there is no point in performing it."

"Other cases may be less outrageous, and it may not be clear whether the difference is due to incompatibility or just unlucky chance. Certainly at some stage you are sure to have a result which disagrees with the rest by many (i.e. three or more) standard deviations. What is the correct way to deal with it? The first thing to do is go back as far as you can and check the readings. You are very likely to find a misplaced decimal point, or a pair of numbers transposed in the notebook. If you can easily retake the measurement then this should be done—and the moral is to plot your points as you go, so that you can catch these rogues at an early stage, before their origins get lost in the mists of history."

"If you cannot find an obvious mistake, then you probably have no choice but to throw the point away. However you should always do so with reluctance. If you have several such points, and/or if there are more points than you would expect with large ($> 2\sigma$) deviations, then you should be extremely suspicious, as there is probably some effect at work that you do not understand, and you should understand. It is usually a trivial matter, but it could be something new and fundamental. Distrust all algorithms that advise the automatic rejection of points outside certain limits as they can rapidly get out of hand; points should only be condemned after giving them a fair hearing."

# *4 Error Propagation in a single variable function*

## 4.1 Simple Linear Function Example

We follow Barlow, Sec. 4.3.1.

"Suppose that f is a simple linear function of x: $f = a x + b$, where a and b are exact constants and x has some distribution with variance V(x) or, equivalently, error $\sigma x$.
x represents a measurement, or perhaps an intermediate result in the analysis, and f could be the final result or another intermediate step."

The variance of f is given by

$$\begin{aligned} Var(f) &= \langle f^2\rangle - \langle f\rangle^2 = \langle (a*x+b)^2\rangle - \langle a*x+b\rangle^2 \\ &= a^2 \langle x^2\rangle + 2*a*b \langle x\rangle + b^2 - a^2 \langle x\rangle^2 - 2*a*b \langle x\rangle - b^2 \\ &= a^2 ( \langle x^2\rangle - \langle x\rangle^2 ) = a^2 Var(x). \qquad (4.18a) \end{aligned}$$

Taking the square root of both sides, we get

$$\sigma f = |a| \sigma x \qquad (4.18b)$$

"This makes good sense, b is a constant, and adding it to a variable does nothing to the spread, 'a' just multiplies the whole distribution by a factor, and increases the width accordingly. It also keeps the dimensions straight if necessary."

### 4.1.1 Distance Travelled at an Approximate Speed Example

Suppose an object is moving on a frictionless track horizontally with an approximate speed v = 200 ± 10 m/sec. If we take the time of travel to be exactly 6 sec, what is the best estimate of the horizontal distance travelled?
distance travelled = Δx = x - x0 = t * v = 6 sec * 200 m/sec = 1200 m.
The error in the distance travelled is σx = t * σv = 6 sec * 10 m/sec = 60 m.
Thus our best estimate of the distance travelled in six seconds is:
1200 +/- 60 m.

## 4.2 Nonlinear Function Case

"Now consider the more useful case where f is some general function of x. For small differences we can expand in a Taylor series" about the point x = x0:
(use ~ to mean approximately equal):

f(x) ~ f(x0) + (x - x0)* [df(x)/dx] | (x -> x0)
= a*x + b,
where a = [df(x)/dx] | (x -> x0) = constant, and
b = f(x0) - x0*[df(x)/dx] | (x -> x0) = constant.

This gives, using (4.18a), Var(f) ~ a^2*Var(x), or taking square roots
σf ~ |a| σx = |df/dx| σx
σf ~ |df/dx| σx

The derivative df/dx is evaluated at x = x0. The approximation in the result
σf ~ |df/dx| * σx
is equivalent to the assumption that the error in x is sufficiently small for f(x) to be represented by a straight line over the range of the measured values of x. The error in f(x) is therefore proportional to the error in x, the constant of proportionality being [df(x)/dx] | (x -> x0).

The use of the first order Taylor expansion of f(x) requires that f(x) be a 'smooth' function (infinitely differentiable at x = x0) and slowing changing with x over a change in x value of the order of σx. The first derivative of f(x) should be evaluated at the true value of x. If this is unknown then the measured value is used, but the difference between the two is insignificant for just this reason.

### 4.2.1 Trigonometry Example

If an angle θ is known with an error of 0.01 radians, then sin (θ) is known with an error of 0.01 |cos(θ)|. (Recall that sin (θ) is dimensionless and in Calculus we always assume the argument of the trig function is expressed in radians.)

```
(%i16) diff (sin (x), x);
```

(%o16) $\cos(x)$

### 4.2.2 Maxima's taylor function

For σ small, the first two terms in a Taylor expansion of f (A + σ) are:

```
(%i17) taylor ( f( A + σ ), σ, 0, 1);
```

(%o17)/T/ $$f(A)+\left(\left.\frac{d}{d\,\sigma}f(\sigma+A)\right|_{\sigma=0}\right)\sigma+\ldots$$

Using the calculus chain rule of differentiation, and denoting σ by e (a small number),
[ d f(A+e)/ de ] | (e = 0) = {[d (A + e)/ de] * [ d f (A + e) / d (A + e) ] } | (e = 0)
= [ d f(x)/ dx ] | (x = A)

So we can write
f (A + σ ) ~ f(A) + σ * d f(A)/ dA
or
f (<A> + σ ) ~ f(<A>) + σ * [ d f(A)/ dA ] | (A --> <A>).

# 5 *Covariance*

## 5.1 Covariance in General Use

Following investopedia.com section on covariance:

"Covariance vs. Variance
Covariance is related to variance, a statistical measure for the spread of points in a data set. Both variance and covariance measure how data points are distributed around a calculated mean. However, variance measures the spread of data along a single axis, while covariance examines the directional relationship between two variables."

"Covariance vs. Correlation
Covariance is also distinct from correlation, another statistical metric often used to measure the relationship between two variables. While covariance measures the direction of a relationship between two variables, correlation measures the strength of that relationship. This is usually expressed through a correlation coefficient, which can range from -1 to +1."

"A correlation is considered strong if the correlation coefficient has a value close to +1 (positive correlation) or -1 (negative correlation). A coefficient that is close to zero indicates that there is only a weak relationship between the two variables."

"In a financial context, covariance is used to examine how different investments perform in relation to one another. A positive covariance indicates that two assets tend to perform well at the same time, while a negative covariance indicates that they tend to move in opposite directions. Investors might seek investments with a negative covariance to help them diversify their holdings."

## 5.2 Barlow on covariance

Following Barlow, Sec. 2.6.1:

"Suppose each item of a data sample consists of a pair of numbers, {(x1, y1),(x2, y2), (x3, y3),…}. You can find their means <x> and <y>, and variances, V(x) and V(y), and standard deviations, σx and σy. However, there is more information there. You can also look at the two variables together- are they independent or do they depend on one another? This is described by the covariance between x and y...."

The covariance between variables x and y has the equivalent definitions

Cov (x, y) = (1/N) Σ (x_i - <x>)*(y_i - <y>)
= < (x - <x>) ( y - <y> ) >
= <x*y> - <x>*<y>,

and when Cov (x, y) is not zero, the values of x and y have some dependence on each other.

If values of x that are above average have a tendency to occur together with above-average y values (which implies that small x likewise tend to accompany small y), then Cov (x, y) will be positive.

If large x tend to go with small y the covariance is negative.

"The covariance between height and weight in a group of adults is presumably positive, as tall people tend to weigh more. Between weight and stamina it may well be negative, as overweight people may be out of condition in other ways. Between height and IQ it is probably zero, as there is no obvious reason for tall people to be more clever, or more stupid, than short people."

# *6 Correlation*

We follow Barlow, Sec. 2.6.2.

"The covariance is useful, but has dimensions. A covariance between height and weight of 7.6, say, means one thing in centimetre-grams and another in metre-kilograms. A better measure of the relation between two variables is the correlation coefficient, ρ. This is defined as:

ρ = Cov (x, y)/ [ σx * σy]."

"ρ is a number between – 1 and + 1. If ρ is zero then x and y are uncorrelated. A positive correlation means that if a particular x happens to be larger than the mean, then y will also (on average) be larger than the mean. For a negative ρ, a larger x will imply a smaller y. If ρ is 1 (or is –1) then x and y are completely correlated: if you know the value of one, that specifies precisely the value of the other."

"ρ is dimensionless, and is unaffected by shifts in the origin or by changes in the scale for x or y."

## 6.1 Barlow Prob. 2.4

We follow a group of 12 students who each take a course in classical mechanics with grades x_j and then each take a course in quantum mechanics with grades y_j. Given the two sets of twelve grades, calculate the average grade achieved in each course, <x> and <y>, the standard deviations σx and σy, the covariance Cov (x, y), and the correlation [coefficient].

```
(%i27) numer : true$
       x : [22, 48, 76, 10, 22, 4, 68, 44, 10, 76, 14, 56]$
       y : [63, 39, 61, 30, 51, 44, 74, 78, 55, 58, 41, 69]$
       mx : mean (x);
       sx : std (x);
       my : mean (y);
       sy : std (y);
       Cov : mean (x*y) - mx*my;
       Corr : Cov/ (sx*sy);
       numer : false$
(mx)   37.5
(sx)   25.912
(my)   55.25
(sy)   14.166
(Cov)  207.46
(Corr) 0.56516
```

"The averages are 37.50 and 55.25, with standard deviation 25.9 and 14.2. The covariance is 207.5 and the correlation is 0.57."

Note Maxima's behavior if you multiply lists of the same length (which we did to calculate the covariance of the lists x and y above).

```
(%i28) [a, b, c] * [A, B, C];
(%o28) [A a,B b,C c]
```

Maxima also has a 'dot product' of lists which uses a period:

```
(%i29) [a, b, c] . [A, B, C];
(%o29) C c+B b+A a
```

# 7 Error Propagation in a function of two or more variables

We follow Barlow, Sec. 4.3.2.

"Suppose that f is a function of two variables, x and y. As in the previous case, consider first a linear relation
f = a*x + b*y + c
where a, b, c are constants."
Var (f) = <f^2> - <f>^2
= < (a*x + b*y + c)^2 > - < a*x + b*y + c >^2

Using Var(x) = <x^2> - <x>^2, Var (y) = <y^2> - <y>^2, and the definition of Cov (x, y),
Cov (x, y) = (1/N) Σ (x_i - <x>)*(y_i - <y>)
= < (x - <x>) ( y - <y> ) >
= <x*y> - <x>*<y>, then f = a*x + b*y + c implies:

Var (f) = a^2*Var (x) + b^2*Var(y) + 2*a*b*Cov (x, y).

If we assume the variables x and y are measurements of independent variables, so <x*y> = <x>*<y> and Cov (x, y) = 0, we then get, with f = a*x + b*y + c,
Var (f) = a^2 * Var (x) + b^2 * Var (y), or
(σf)^2 = a^2 * (σx)^2 + b^2 * (σy)^2.

In the nonlinear case, we assume f(x,y) is approximately linear in both x and y for small values of (x - x0) and (y - y0) to use a two variable Taylor series expansion of f(x,y) about (x0, y0), with x a measured value of property x which has a true value of x0, and likewise y a different, independent measured property y with a true value y0,
f (x, y) ~ f (x0, y0) + (x - x0)*[∂f/∂x]_0 + (y - y0)*[∂f/∂y]_0
in which [∂f/∂x]_0 means [∂f/∂x] | (x -> x0, y -> y0, after partial differentiation of f(x,y) with respect to x), and thus a constant, and similarly for [∂f/∂y]_0.

Now let a == [∂f/∂x]_0 and b == [∂f/∂y]_0, and define a third constant c by:
c == f (x0, y0) - a * x0 - b*y0.
Then f (x,y) ~ a*x + b*y + c, in which a, b, and c are all constants.
This is exactly the linear example we started this section with, so we can immediately write the result (assuming Cov (x, y) = 0 ),
Var (f) = a^2 * Var (x) + b^2 * Var (y), or
(σf)^2 = [∂f/∂x]^2 * (σx)^2 + [∂f/∂y]^2 * (σy)^2,
with the understanding the first partial derivatives in this expression are each evaluated at the true values (x0, y0), or at least the best estimates of the true values.

"This says that the error on x and y, multiplied by suitable scaling factors, are added in quadrature. Adding two positive numbers in quadrature gives a smaller result than the usual arithmetic addition. Intuitively you can see that this is reasonable: errors of overestimation in x have a fair chance of being compensated by errors of underestimation in y."

Here we first add the numbers 3 and 4 with the usual arithmetic addition, and then add the same two numbers in quadrature:

```
(%i31) 4 + 3;
       sqrt (4^2 + 3^2);
(%o30) 7
(%o31) 5
```

"The extension to more than two variables is straightforward; all you do (provided they are all independent, of course) is add the errors in quadrature. Thus for a function of x, y, and z,
(σf)^2 = [∂f/∂x]^2 * (σx)^2 + [∂f/∂y]^2 * (σy)^2 + [∂f/∂z]^2 * (σz)^2.

## 7.1 Distance s with both approximate a and v0

Suppose an object is moving on a track horizontally with an approximate initial speed v0 = (200 ± 10) m/sec and also an approximate constant acceleration a = (12 +/- 2) m/sec^2. If we take the time of travel to be exactly 6 sec, what is the best estimate of the horizontal distance travelled?

distance travelled == s (v0, a) = x - x0 = v0 * t + (1/2)*a*t^2
= 6 sec * 200 m/sec + (1/2)* (12 m/s^2) * (6 sec)^2 = 1416 m.
The distance travelled is a function of the independent variables (v0, a)
with ∂s/∂v0 = t = 6 sec, and ∂s/∂a = (1/2)*t^2 = 18 sec^2

The square of the error in the distance travelled, σs, is then given by

σs^2 = [∂s/∂v0]^2 (σv0)^2 + [∂s/∂a]^2 (σa)^2
= (6 sec)^2 * (10 m/sec)^2 + (18 sec^2)^2 * (2 m/sec^2 )^2
= 3600 m^2 + 1296 m^2 = 4896 m^2
Hence
σs = 69.97 m ~ 70 m,
and s = (1416 +/- 70) m.

## 7.2 Nonlinear Example with 3 Variables

Let
y = A*sin (θ) + B*cos (θ)
and assume A, B, and θ are all independent variables.
Then the error in y is the square root of

σy^2 = sin (θ)^2 * σA^2 + cos (θ)^2 * σB^2 + [ A*cos (θ) - B*sin (θ) ]^2 * σθ^2 .

## 7.3 Ohm's Law Example

The steady state current I is given by Ohm's Law:
I = V / R = voltage/resistance
The usual equation for the square of the error in the current
σI^2 = [ ∂I/∂V ]^2 * σV^2 + [ ∂I/∂R ]^2 * σR^2
can be expressed as
[ σI/I ]^2 = [ σV/V ]^2 + [ σR/R]^2.

### 7.3.1 Barlow, Prob. 4.4

If a voltage is determined by measuring a current of 1120 ± 10mA through a resistance of 1400 ± 30 Ω, what is its value and error?

1 volt = 1 A * 1 Ω, 1 mA = 10^(-3) A.

V = I*R = 1120 mA * 1400 Ω = 1.12 A * 1400 Ω = 1568 volts.
∂V/∂I = R, ∂V/∂R = I,
σV^2 = R^2 * σI^2 + I^2 * σR^2
= (1400)^2 * (10^(-2) )^2 + (1.12)^2 * (30)^2
= 196 + 1128.96 = 1324.66
σV = 36.4 ~ 40 volts.
V = (1570 +- 40) volts

## 7.4 Barlow, Prob. 4.6; Trig Functions

A.) If θ = 0.56 ± 0.01, what are the errors on sin θ, cos θ, and tan θ?

```
(%i32) Sig (ff, xx ) := subst (0.56, xx, abs (diff (ff, xx))*0.01)$

(%i34) x;
       kill (x)$
(%o33) [22,48,76,10,22,4,68,44,10,76,14,56]

(%i35) Sig (sin(x), x);
(%o35) 0.0084726

(%i36) Sig (cos (x), x);
(%o36) 0.0053119

(%i37) Sig (tan (x), x);
(%o37) 0.013931
```

The errors on sin (x), cos (x), and tan (x), if x = 0.56 +/- 0.01 are respectively 0.008, 0.005, and 0.014.

B.) If θ = 1.56 ± 0.01, what are the errors on sin θ, cos θ, and tan θ?
This is about 89.4 degrees.

```
(%i38) Sig (f, x ) := subst (1.56, x, abs (diff (f, x))*0.01)$
```

```
(%i41) Sig (sin(x), x);
       Sig (cos(x), x);
       Sig (tan(x), x);
```

(%o39) $1.0796\ 10^{-4}$

(%o40) 0.0099994

(%o41) 85.796

The errors on sin (x),  and cos (x), if x = 1.56 +/- 0.01 are
respectively 0.0001 and 0.01.
"The errors on tan(x) are so large that the 'small errors' assumption breaks down and so no sensible error can be assigned."

# 8 Appendix

partial derivative template
In wxMaxima, press escape key once, then p, then down-key once to get to partial, then press enter.

∂
∂F/∂x