Statistics Review for Experimental Measurements

It is important to be able properly to estimate the uncertainty in an experimental result. This is true of both the actual measurement as well as any other results that may be derived from that measurement. An understanding of the statistical nature of the measurement will allow us to better approximate these uncertainties and to evaluate the accuracy of a given measurement.

1.0 Probability distributions

Anytime a measurement is made there is a range of possible values that have a finite probability of occurring. Understanding this range and the principles that govern that range will help in evaluating the result. Usually, we will estimate a probability distribution function for the measurement that will provide us with an understanding of this range of possible values and the likelihood of each occurring.

First, we need to define several terms. The first two apply to a collection of discrete data values. The second two apply to items determined from a statistical distribution that describes a particular continuous data set.

average: The technical name of this is the arithmetic mean but may also be called just the mean. It is given by

$$\bar{y} = \frac{1}{N} \sum_{i=1}^{N} y_i$$

where $\bar{y}$ is the average of the N values y_i.

standard deviation: Most commonly this refers to the *sample standard deviation* or the *corrected sample standard deviation*. It is given by

$$\sigma^2 = \frac{1}{N-1} \sum_{i=1}^{N} (y_i - \bar{y})^2$$

where σ is the standard deviation of the N data values y_i from the average value $\bar{y}$. σ represents the RMS (root mean square) deviation of the values from the average and provides an indicator of the uncertainty or the error in the values.

expected value: The is the most likely value to occur as indicated by the probability distribution representing some random variable. It is often referred to as the average since it usually corresponds to the arithmetic mean of the values as the number of repetitions of the measurement goes to infinity. It is also referred to as the probability-weighted average of all possible values. It is given by

$$\bar{s} \;=\; \sum_s s\, P(s)$$

for a collection of discrete values where $P(s)$ represents the probability distribution function for s and is normalized so

$$\sum_s P(s) \;=\; 1$$

and where the sum is over all possible values of s. If the possible values are continuous the expected value is given by

$$\bar{s} \;=\; \int_s s'\, P(s')\, ds'$$

with the normalization condition

$$\int_s P(s')\, ds' \;=\; 1\,.$$

variance: The variance is analogous to the standard deviation of a collection of data and is usually indicated by σ^2. The square root of the variance will be approximately equal to the standard deviation for a large collection of random values that are properly represented by the probability distribution used to calculate the variance. It is given by

$$\sigma^2 \;=\; \sum_s (s - \bar{s})^2\, P(s)$$

for a discrete distribution and

$$\sigma^2 \;=\; \int_s (s' - \bar{s})^2\, P(s')\, ds'$$

for a continuous distribution.

Since we are dealing with probabilities, the distribution functions should be properly normalized as noted in the above definitions. In the following discussion, we assume that $P(s)$ is always properly normalized.

We can also simplify the form for the variance:

$$\begin{aligned}
\sigma^2 &= \sum_s (s-\bar{s})^2\, P(s) \\
&= \sum_s (s^2 - 2\,\bar{s}\, s + \bar{s}^2)\, P(s) \\
&= \sum_s s^2\, P(s) - 2\,\bar{s} \sum_s s\, P(s) + \bar{s}^2 \sum_s P(s) \\
&= \sum_s s^2\, P(s) - 2\,\bar{s}^2 + \bar{s}^2 \\
&= \sum_s s^2\, P(s) - \bar{s}^2
\end{aligned}$$

using the definition of $\bar{s}$ and the normalization of $P(s)$.

It can be similarly shown that the same simplification applies to the integrals for a continuous probability distribution function:

$$\sigma^2 = \int_{s_1}^{s_2} s'^2 P(s')\,\mathrm{d}s' - \bar{s}^2$$

were s_1 and s_2 cover the entire range of possible values of s.

It is important to remember that this discussion of distributions assumes that the individual measurements are completely independent and uncorrelated so that any variations between successive measurements are random.

1.1 Binomial distribution

The binomial distribution represents the results from a system where any given trial can result in either a success or a failure – there are no other options. An example of this type of experiment would be flipping a coin. We will denote the probability that a given trial succeeds by the variable p, the probability of failure by $q = 1 - p$, and complete n trials. Then the probability of having s successes in those n trials is given by

$$\begin{aligned}
\text{probability of } s \text{ successes in } n \text{ trials} &= B_{n,p}(s) \\
&= \frac{n(n-1)\cdots(n-s+1)}{1\times 2\times\cdots\times s} p^s q^{n-s}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{n!}{s!(n-s)!} p^s q^{n-s} \\
&= \binom{n}{s} p^s q^{n-s}
\end{aligned}
$$

where $\binom{n}{s}$ is the binomial coefficient. This is properly normalized so

$$
\sum_{s=0}^{n} B_{n,p}(s) \;=\; 1\,.
$$

The expected value is

$$
\begin{aligned}
\bar{s} &= \sum_{s=0}^{n} s\, B_{n,p}(s) \\
&= \sum_{s=0}^{n} s\, \frac{n!}{s!(n-s)!} p^s q^{n-s} \\
&= \sum_{s=1}^{n} s\, \frac{n!}{s!(n-s)!} p^s q^{n-s} \\
&= \sum_{s=1}^{n} n\, p\, \frac{(n-1)!}{(s-1)!(n-s)!} p^{s-1} q^{n-s} \\
&= n\, p \sum_{j=0}^{n-1} \frac{(n-1)!}{j!(n-1-j)!} p^j q^{n-1-j} \\
&= n\, p \sum_{j=0}^{n-1} B_{n-1,p}(j) \\
\bar{s} &= n\, p\,.
\end{aligned}
$$

Because B is properly normalized, the sum in the next to last line evaluates to one (1). The variance is

$$
\begin{aligned}
\sigma^2 &= \sum_{s=0}^{n} s^2\, B_{n,p}(s) - \bar{s}^2 \\
&= \sum_{s=0}^{n} s^2\, \frac{n!}{s!(n-s)!} p^s q^{n-s} - \bar{s}^2 \\
&= \sum_{s=1}^{n} s^2\, \frac{n!}{s!(n-s)!} p^s q^{n-s} - \bar{s}^2
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_{s=1}^{n} s \, \frac{n!}{(s-1)!(n-s)!} p^s q^{n-s} - \bar{s}^2 \\
&= n\,p \sum_{s=1}^{n} s \, \frac{(n-1)!}{(s-1)!(n-s)!} p^{s-1} q^{n-s} - \bar{s}^2 \\
&= n\,p \sum_{j=0}^{n-1} (j+1) \, \frac{(n-1)!}{j!(n-1-j)!} p^j q^{n-1-j} - \bar{s}^2 \\
&= n\,p \left[\sum_{j=0}^{n-1} j \, \frac{(n-1)!}{j!(n-1-j)!} p^j q^{n-1-j} + \sum_{j=0}^{n-1} \frac{(n-1)!}{j!(n-1-j)!} p^j q^{n-1-j} \right] - \bar{s}^2 \\
&= n\,p \left[\sum_{j=1}^{n-1} j \, B_{n-1,p}(j) + \sum_{j=0}^{n-1} B_{n-1,p}(j) \right] - \bar{s}^2 \\
&= n\,p\,[(n-1)\,p + 1] - (n\,p)^2 \\
\sigma^2 &= n\,p\,(1-p)
\end{aligned}
$$

since the first sum represents the expected value of $j = (n-1)\,p$ and the second sum is again equal to one (1) since B is normalized.

1.2 Poisson distribution

The Poisson distribution is usually associated with counting randomly occurring events that have an average rate of occurrence over a particular interval (time, distance, area, etc.). Some examples of this type of measurement would be determining the rate of decay of a radioactive sample, counting photons in a low-level optical experiment, or counting how customers arrive at the teller in a bank.

The Poisson distribution can be found from the limit of the binomial distribution if the number of trials, n, is large and the probability of success, p, is small. Typically this is satisfied if $n \geq 100$ and $(n\,p) \leq 10$. First, define $\mu = n\,p$, substitute $(1-p)$ for q, and factor $\mu^s/s!$ out of the binomial distribution

$$
\begin{aligned}
B_{n,p}(s) &= \frac{n!}{s!(n-s)!} \, p^s \, q^{n-s} \\
&= \frac{\mu^s}{s!} \, \frac{n!}{(n-s)!} \left(\frac{1}{n}\right)^s \left(1 - \frac{\mu}{n}\right)^{n-s} \\
&= \frac{\mu^s}{s!} \left[\frac{n!}{(n-s)!} \left(\frac{1}{n}\right)^s \right] \left[\left(1 - \frac{\mu}{n}\right)^n \right] \left[\left(1 - \frac{\mu}{n}\right)^{-s} \right]
\end{aligned}
$$

If we now take the limit of this equation as $n \to \infty$, and with a little help from

Mathematica, we find the limits of the three terms in square brackets to be

$$\begin{aligned}
\lim_{n\to\infty} \frac{n!}{n^s\,(n-s)!} &= 1 \\
\lim_{n\to\infty} \left(1-\frac{\mu}{n}\right)^{-s} &= 1 \\
\lim_{n\to\infty} \left(1-\frac{\mu}{n}\right)^{n} &= \mathrm{e}^{-\mu}\,.
\end{aligned}$$

Combining these results we get

$$P_\mu(s) = \mathrm{e}^{-\mu}\frac{\mu^s}{s!}$$

where $P_\mu(s)$ represents the probability that exactly s events occur in the specified interval.

The expected value is found by

$$\begin{aligned}
\bar{s} &= \sum_{s=0}^{\infty} s\,P_\mu(s) \\
&= \sum_{s=1}^{\infty} s\,\mathrm{e}^{-\mu}\frac{\mu^s}{s!} \\
&= \sum_{s=1}^{\infty} \mathrm{e}^{-\mu}\frac{\mu^s}{(s-1)!} \\
&= \sum_{s=1}^{\infty} \mu\,\mathrm{e}^{-\mu}\frac{\mu^{s-1}}{(s-1)!} \\
&= \mu\,\mathrm{e}^{-\mu}\sum_{j=0}^{\infty} \frac{\mu^j}{j!}
\end{aligned}$$

where $j = s - 1$. Note that the normalization of the distribution provides the relationship

$$\begin{aligned}
\sum_{j=0}^{\infty} \mathrm{e}^{-\mu}\frac{\mu^j}{j!} &= 1 \\
\sum_{j=0}^{\infty} \frac{\mu^j}{j!} &= \mathrm{e}^{\mu}
\end{aligned}$$

which, when substituted into the expression for $\bar{s}$ gives

$$\begin{aligned}\bar{s} &= \mu\,\mathrm{e}^{-\mu}\,\mathrm{e}^{\mu} \\ &= \mu\,.\end{aligned}$$

One of the surprising, but very useful, results for this distribution is that

$$\begin{aligned}\sigma^2 &= \sum_{s=0}^{\infty} s^2\,P_\mu(s) - \bar{s}^2 \\ &= \sum_{s=1}^{\infty} s^2\,\mathrm{e}^{-\mu}\frac{\mu^s}{s!} - \bar{s}^2 \\ &= \sum_{s=1}^{\infty} s\,\mathrm{e}^{-\mu}\frac{\mu^s}{(s-1)!} - \bar{s}^2 \\ &= \mu\sum_{s=1}^{\infty} s\,\mathrm{e}^{-\mu}\frac{\mu^{s-1}}{(s-1)!} - \bar{s}^2 \\ &= \mu\left[\sum_{j=0}^{\infty}(j+1)\,\mathrm{e}^{-\mu}\frac{\mu^j}{j!}\right] - \bar{s}^2 \\ &= \mu\left[\sum_{j=0}^{\infty} j\,P_\mu(j) + \sum_{j=0}^{\infty} P_\mu(j)\right] - \bar{s}^2 \\ &= \mu[\mu+1] - \mu^2 \\ \sigma^2 &= \mu\end{aligned}$$

where $j = s - 1$ again, the first sum finds the expected value of j and the second is equal to one (1) because this distribution is also normalized.

1.3 The uniform distribution

When any value within a particular range of values is equally likely we have what is called a uniform distribution with a distribution function given by

$$P(s) = \begin{cases} 1/c & \text{for } (\bar{s} - c/2) \le s \le (\bar{s} + c/2) \\ 0 & \text{otherwise.}\end{cases}$$

The denominator in P is a normalization so that the integral of $P(s)$ over all s is equal to one (1). It is apparent that the expected value is $\bar{s}$ (you can do the integral if you would like). The variance is not as obvious:

$$\begin{aligned}
\sigma^2 &= \int_{\bar{s}-c/2}^{\bar{s}+c/2} s^2 \, \frac{1}{c} \, \mathrm{d}s - \bar{s}^2 \\
&= \left(\frac{c^2}{12} + \bar{s}^2\right) - \bar{s}^2 \\
&= \frac{c^2}{12} \, .
\end{aligned}$$

This distribution is very common when recording values with modern digital equipment since we can only record discrete values but we are usually measuring a continuous function. c is the minimum spacing between the possible values on a digital voltmeter or the spacing between levels returned from an analog-to-digital converter.

As an example of this distribution, if a digital voltmeter is used to measure a voltage, and it reads 0.001 V, the actual value could lie anywhere within the range of 0.0005 to 0.0015 V with equal likelihood. The average or expected value is 0.001 V and the standard deviation is $\sigma = c/\sqrt{12} = 0.001/\sqrt{12}$ since $c = 0.001$ is the interval between successive least-significant digits on the meter. ***With modern digital equipment, this uncertainty will always be present because we only record discrete values from continuous signals***.

1.4 The normal or Gaussian distribution

As noted in the discussion of measurements and noise above, the authors of the *Introductory Statistics* online text stated, "The normal, a continuous distribution, is the most important of all the distributions. It is widely used and even more widely abused. Its graph is bell-shaped. You see the bell curve in almost all disciplines. ... The normal distribution is extremely important, but it cannot be applied to everything in the real world." [1]

The normalized Gaussian distribution is given by

$$P(s) = \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-(s-\bar{s})^2/(2\sigma^2)} \, .$$

It is fairly easy to show using the integral equations that

$$\int_{-\infty}^{\infty} s \, P(s) \, \mathrm{d}s = \bar{s}$$

[1] OpenStax College, *Introductory Statistics*, OpenStax College, 19 September 2013, p. 361, available at http://cnx.org/content/col11562/latest/.

and

$$\int_{-\infty}^{\infty} s^2\, P(s)\, \mathrm{d}s - \bar{s}^2 \;=\; \sigma^2\,.$$

Many measurements of continuous variables in physics will exhibit a Gaussian distribution. If you are uncertain if your signal does have a Gaussian dependence, you can collect a large number of measurements. You then bin the values, plot a histogram of the number of times the value falls within the appropriate bin, and compare this to a Gaussian curve.

If your measurements are truly Gaussian the meaning of σ is very specific. If you integrate the distribution over some range that is centered on $\bar{s}$ you can get the fraction of measurements that will typically fall in that range:

$$f \;=\; \frac{1}{\sigma\sqrt{2\,\pi}} \int_{\bar{s}-\Delta}^{\bar{s}+\Delta} \mathrm{e}^{-(s-\bar{s})^2/(2\sigma^2)}\, \mathrm{d}s\,.$$

If we let $t = (s - \bar{s})/(\sqrt{2}\,\sigma)$ this integral becomes

$$\begin{aligned} f \;&=\; \frac{1}{\sigma\sqrt{2\,\pi}} \int_{-\Delta/(\sqrt{2}\,\sigma)}^{\Delta/(\sqrt{2}\,\sigma)} \mathrm{e}^{-t^2}\, \sqrt{2}\,\sigma\, \mathrm{d}t \\ &=\; \frac{1}{\sqrt{\pi}} \int_{-\Delta/(\sqrt{2}\,\sigma)}^{\Delta/(\sqrt{2}\,\sigma)} \mathrm{e}^{-t^2}\, \mathrm{d}t\,. \end{aligned}$$

Since the integrand is symmetrical about zero (0) and the integral limits are also symmetrical about zero we can just use twice the integral from zero to the upper limit:

$$f \;=\; \frac{2}{\sqrt{\pi}} \int_{0}^{\Delta/(\sqrt{2}\,\sigma)} \mathrm{e}^{-t^2}\, dt\,.$$

If we now express Δ as a multiple of σ, $\Delta = n\,\sigma$, the integral becomes

$$\begin{aligned} f \;&=\; \frac{2}{\sqrt{\pi}} \int_{0}^{n/\sqrt{2}} \mathrm{e}^{-t^2}\, dt \\ &=\; \mathrm{erf}(n/\sqrt{2}) \end{aligned}$$

where $\mathrm{erf}(x)$ is the Gaussian error function that is well known. Most mathematical packages include this as one of the standard functions.

If we look at the fraction of the values that fall within a given number of standard deviations we find

$$\begin{aligned}
\Delta &= \sigma, \quad \mathrm{erf}(1/\sqrt{2}) = 0.6826894921370858 \\
\Delta &= 2\,\sigma, \quad \mathrm{erf}(2/\sqrt{2}) = 0.954499736103642 \\
\Delta &= 3\,\sigma, \quad \mathrm{erf}(3/\sqrt{2}) = 0.99730020393674 \\
\Delta &= 4\,\sigma, \quad \mathrm{erf}(4/\sqrt{2}) = 0.999936657516334 \\
\Delta &= 5\,\sigma, \quad \mathrm{erf}(5/\sqrt{2}) = 0.999999426696856 \\
\Delta &= 6\,\sigma, \quad \mathrm{erf}(6/\sqrt{2}) = 0.999999998026825
\end{aligned}$$

As you can see, virtually all the measurements of a signal that exhibits Gaussian behavior should fall within the range of $\bar{s} \pm 6\,\sigma$.

2.0 Comparison of distributions

It may be instructive to look at graphs of the distributions. First, we will compare the binomial distribution with the normal distribution. To emphasize the differences, I have chosen both to have a mean value of 4. Since the maximum variance of the binomial distribution would then be 2 ($\bar{s} = n\,p$, $\sigma^2 = n\,p\,q$ with the largest value at $p = q = 0.5$), both distributions in Figure 1 have been drawn with these values.

The Poisson distribution requires that the mean value and the variance be equal. In Figure 2 we use a mean value of 4 and a variance of 4 to compare the Poisson distribution and the normal distribution.

Figure 3 is included to show the uniform distribution with the same values (mean = variance = 4). In this case, the value of c (the width of the distribution and the reciprocal of the amplitude) is found from $\sqrt{12\,\sigma^2} = 6.9282$.

3.0 Propagating errors

Often a measured value will be used as a parameter in some equation to give the desired result. For example, if you have a temperature transducer that produces a voltage related to the temperature, you may be required to apply some calibration equation to the voltage to arrive at the actual temperature. That equation may simply be a linear function, or it can involve exponentials or logarithms, depending on the type of transducer. It is necessary to analyze how the uncertainty in the final temperature is related to any uncertainty in the measured voltage and uncertainties in the calibration parameters.

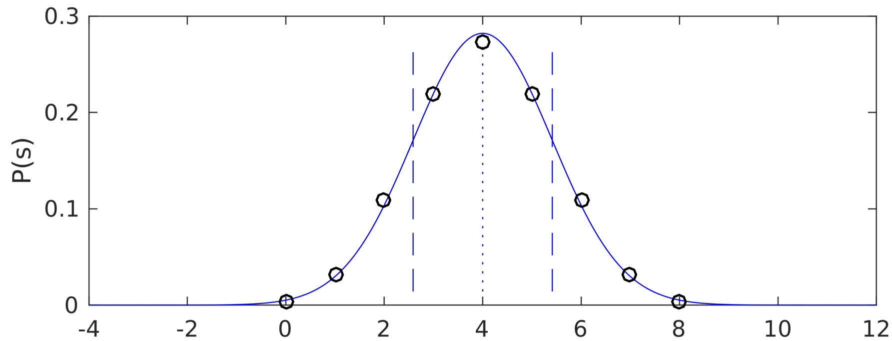

Figure 1: A graph of the binomial distribution (circles) overlaid on the normal distribution (solid line). Both have been calculated to have a mean value of 4 and a variance of 2. For the binomial distribution this corresponds to $n = 8$, $p = q = 0.5$. The vertical dotted line shows the mean value and the vertical dashed lines show the mean plus or minus σ ($\sqrt{2}$). You will notice that the binomial distribution has a lower peak value, is broader about halfway down, but goes to zero faster on the tails. At $s = 8$ and $s = 0$ the binomial distribution has a value of 0.003906 while the normal distribution has a value of 0.005167. The differences are not large.

The process of propagating errors through a calculation is fairly straightforward in many cases. For instance, if the value y is a function of the inputs x_1, x_2, $\ldots$, x_n

$$y = f(x_1, x_2, x_3, \ldots, x_n)\,,$$

and the uncertainties of the x_i are given by δ_i, we can use a Taylor expansion of y about $\bar{y}$, keeping only first order terms we get

$$\begin{aligned}\delta_y &= y - \bar{y} \\ &= \sum_{i=1}^{n} \delta_i \left(\frac{\partial f}{\partial x_i}\right)\,.\end{aligned}$$

Because the signs of the δ_i are random, this will usually incorrectly estimate the uncertainty since they can have alternating signs such that it sums to zero or they

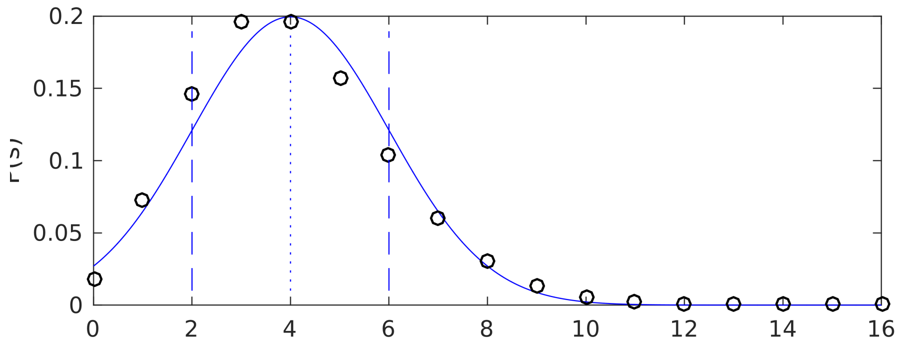

Figure 2: A graph of the Poisson distribution (circles) overlaid on the normal distribution (solid line). Both have been calculated to have a mean value of 4 and a variance of 4. The vertical dotted line shows the mean value and the vertical dashed lines show the mean plus or minus σ ($\sqrt{4}$). You will notice that the Poisson distribution has about the same peak value and that the peak is shifted toward lower values of s. On the right side of the peak (higher values of s) the Poisson distribution is lower than the normal distribution, consistent with the peak being shifted to the left. But the Poisson goes to zero more slowly as s increases. At $s = 12$ the Poisson and normal distribution values are 6.415×10^{-4} and 6.692×10^{-5} respectively. At $s = 16$ the values are 3.760×10^{-6} and 3.038×10^{-9} respectively.

---

can all have the same sign so that it sums to a large number. You can get a better estimate of the uncertainty by considering the results from a random walk where the square of the distance covered is approximately equal to the sum of the squares of each of the individual steps. This gives us

$$\delta_y^2 = \sum_{i=1}^{n} \delta_i^2 \left(\frac{\partial f}{\partial x_i} \right)^2 .$$

In this equation, we have assumed that the x_i are independent and there is no cross-correlation between different x_i. If this not true, we will have to add one or more

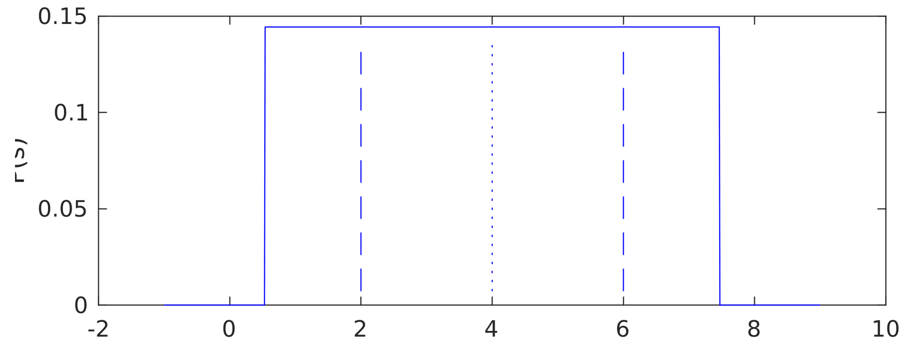

Figure 3: A graph of the normal distribution. This was calculated to have a mean value of 4 and a variance of 4 to match the values in Figure 2. The vertical dotted line shows the mean value and the vertical dashed lines show the mean plus or minus σ ($\sqrt{4}$). c, the width of the distribution and the inverse of the amplitude was calculated from the specified variance: $c = \sqrt{12 * \text{variance}} = 6.9282$.

---

terms to the sum of the form

$$\sum_{i=1}^{n} \sum_{j=1 (j \neq i)}^{n} \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} \delta_{ij}$$

where δ_{ij} is called the covariance of x_i and x_j and arises from correlations between those variables where a change in one of them causes a change in the other. If the δ_i values are uncorrelated we don't have the cross-correlation terms in the equation.

There are also some pitfalls in this equation if the function f is nonlinear or the uncertainties δ_i are large in some sense. This is because the propagation equation is based on a first-order Taylor expansion of the equation for y under the assumption that we only need to keep the lowest order powers of δ_i because $\delta_i^2 \ll \delta_i$..

3.1 Offset voltage

An application of this method can be seen if you have determined an offset voltage,

v_{off}, associated with an amplifier but that voltage has an uncertainty, δ_{off}, possibly due to noise on the signal. Then the corrected voltage will be given by

$$v_c = v - v_{off} .$$

Applying the above propagation technique to this equation is straightforward and results in the uncertainty

$$\delta_{v_c} = \sqrt{\delta_v^2 + \delta_{off}^2} .$$

3.2 Signal averaging

Returning to the case of averaging a signal over some period to reduce noise, it is straightforward to derive the effect on the standard deviation of the average.

If we are averaging N points together so

$$\bar{y} = \frac{1}{N} \sum_{j=1}^{N} y_j$$

we can also obtain the standard deviation of those N points given by

$$\sigma_y = \sqrt{\frac{\sum_{j=1}^{N} (y_j - \bar{y})^2}{N-1}}$$

(the sample standard deviation). If we took a single sample y_i, we would expect it to have a standard deviation of σ_y as long as the number of points, N, in the original data set is large enough and the acquisition sample rate fast enough to get a good sample of the signal and any noise present.

If we then take an average of N samples (which conveniently can be the same set of samples we used in determining the standard deviation of each of the samples) we would find

$$\bar{y} = \frac{1}{N} \sum_{j=1}^{N} y_j$$

$$\sigma_{\bar{y}} = \sqrt{\sum_{j=1}^{N} \left(\sigma_y \frac{\partial \bar{y}}{\partial y_j} \right)^2}$$

$$
\begin{aligned}
&= \sqrt{\sum_{j=1}^{N}\left(\sigma_y \frac{1}{N}\right)^2} \\
&= \sqrt{\frac{\sigma_y^2}{N^2}\sum_{j=1}^{N} 1} \\
&= \sqrt{\frac{\sigma_y^2}{N}} \\
\sigma_{\bar{y}} &= \frac{\sigma_y}{\sqrt{N}}\,.
\end{aligned}
$$

Averaging a slowly-varying signal can significantly improve the uncertainty in the resulting value.

You can combine other errors, in the same way, to arrive at a total error in your final value. Care in accounting for the possible errors in a measurement will significantly improve your understanding of the quality and usefulness of that measurement.

[Modified: January 16, 2019]