Outline

- Measurements, errors accuracy and precision
  - How good an estimator is the average, if the individual measurements have an error of $\sigma$?
  - Standard deviation of parent sample vs. daughter sample.
  - Error propagation.
  - Correlated errors and example.

So what is the variance of the average?

- If $N$ measurements are taken for supposedly physically the same quantity (like the length of an object, temperature of an object, intensity of constant light), what is the variance (error) of their average?
- Suppose that individual measurements have an error (due to resolution of your equipment, noise, etc. – for simplicity we assume that the error is common among all measurements) of $\sigma$.
- The variance of the average is given by $\sigma^2/N$.
- This makes sense when $N = 1$!
- In this context, the error of a measurement is given by $\sigma$ means that $<(x_i - x_{\text{truth}})^2> = \sigma^2$, where $<\ldots>$ is the average of many trials. “$i$” represents the trial index.
- So the error of the average will be given by $<(x_{\text{ave}} - x_{\text{truth}})^2>$.
- Plugging in the definition of the average, this can be modified to look like the following:

$$\left\langle \left( \frac{1}{N} \sum_{i=1}^{N} x_i - x_{\text{truth}} \right)^2 \right\rangle = \left\langle \frac{1}{N} \left\{ \sum_{i=1}^{N} (x_i - x_{\text{truth}}) \right\}^2 \right\rangle$$

$$= \left\langle \frac{1}{N^2} \left\{ \sum_{i=1}^{N} \sum_{j=1}^{N} (x_i - x_{\text{truth}})(x_j - x_{\text{truth}}) \right\} \right\rangle = \frac{1}{N^2} \left\{ \sum_{i=1}^{N} \sum_{j=1}^{N} (x_i - x_{\text{truth}})(x_j - x_{\text{truth}}) \right\},$$

where I change the order of averaging $<\ldots>$ and summing over $i$ and $j$.

Now I claim the $<(x_i - x_{\text{truth}})(x_j - x_{\text{truth}})> = 0$ unless $i = j$. This arises from our expectation that the $i$-th and $j$-th measurements are “independent.”
- This may not be the case if human factor is important – for example, if a person is thinking that “I think the previous measurement was on the lower side, I am going to compensate it with a higher value in the next measurement” type of psychology.
- $<(x_i - x_{\text{truth}})(x_j - x_{\text{truth}})> = 0$ because the two averages are both zero unless the measurements are biased (tend to get higher-than-truth or lower).
- Another way of describing the “independence” of the measurements is that if the $i$-th measurement is higher than the truth, there is equal chance that the $j$-th measurement is higher or lower than the truth.
- Given these manipulations,

$$= \frac{1}{N^2} \left\{ \sum_{i=1}^{N} \sum_{j=1}^{N} (x_i - x_{\text{truth}})(x_j - x_{\text{truth}}) \right\} = \frac{1}{N^2} \left\{ \sum_{i=1}^{N} (x_i - x_{\text{truth}})(x_i - x_{\text{truth}}) \right\} = \frac{1}{N^2} N\sigma^2 = \frac{\sigma^2}{N}.$$
What if error has random as well as "systematic bias?"

- This statement translate into this: $x_i - x_{\text{truth}}$ has two components: $\delta_i^{\text{random}}$ and $\delta_i^{\text{systematic}}$. i.e. $x_i = x_{\text{truth}} + \delta_i^{\text{random}} + \delta_i^{\text{systematic}}$, and even though $\langle \delta_i^{\text{random}} \delta_j^{\text{random}} \rangle = 0$, $\langle \delta_i^{\text{systematic}} \delta_j^{\text{systematic}} \rangle \neq 0$ because if $\delta_i^{\text{systematic}}$ is larger than 0 because Yuichi is not careful, there is a good chance that $\delta_j^{\text{systematic}}$ would also be greater than 0.

$\sigma \ vs. \ \sigma_N$

- Now, here is a little note that $\sigma$ in the above expression is not the standard deviation you calculate for your finite samples of measurements $\langle (x_i - x_{\text{ave}})^2 \rangle = \sigma^2_N$, but $\langle (x_i - x_{\text{truth}})^2 \rangle = \sigma^2$. So what is the difference between them? It turns out that $\sigma^2_N = \frac{N-1}{N} \sigma^2$ or

$$\sigma^2 = \frac{N}{N-1} \sigma^2_N = \frac{N}{N-1} \frac{1}{N} \sum_{i=1}^{N} (x_i - x_{\text{ave}})^2 = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - x_{\text{ave}})^2.$$ Instead of dividing the sum of the squared terms by $N$, you should divide it by $(N-1)$ to get $\sigma$. This relation can be proven in the following way.

- Before I start my proof, it would be convenient to do the following. By considering $\delta_i = x_i - x_{\text{truth}}$, whenever $x_{\text{ave}}$ is not zero, you can invent a related quantity, whose average is zero. So from here on, I assume that the average is zero.

$$\langle \sigma^2_N \rangle = \left \langle \frac{1}{N} \sum_{i=1}^{N} (x_i - x_{\text{ave}})^2 \right \rangle = \left \langle \frac{1}{N} \sum_{i=1}^{N} \left( x_i - \frac{1}{N} \sum_{j=1}^{N} x_j \right)^2 \right \rangle = \left \langle \frac{1}{N^3} \sum_{i=1}^{N} \left( N x_i - \sum_{j=1}^{N} x_j \right)^2 \right \rangle \tag{1}$$

$$= \left \langle \frac{1}{N^3} \sum_{i=1}^{N} \left( (N-1)x_i - \sum_{j=1, j \neq i}^{N} x_j \right)^2 \right \rangle = \left \langle \frac{1}{N^3} \sum_{i=1}^{N} \left( (N-1)x_i \right)^2 - \left \{ \sum_{j=1, j \neq i}^{N} x_j \right \}^2 \right \rangle$$

- In the last step, I used the fact that $\langle x_i, x_j \rangle = 0$, which is true as long as measurements are independent of each other. Note that I am now assuming that $\langle x_i \rangle = 0$.

$$\langle \sigma^2_N \rangle = \left \langle \frac{1}{N^3} \sum_{i=1}^{N} \left( (N-1)x_i \right)^2 + \left \{ \sum_{j=1, j \neq i}^{N} x_j \right \}^2 \right \rangle = \frac{1}{N^3} \left[ (N-1)^2 \sum_{i=1}^{N} \langle x_i^2 \rangle + \sum_{i=1}^{N} \sum_{j=1, j \neq i}^{N} \langle x_i x_j \rangle \right]$$

$$= \frac{1}{N^3} \left[ (N-1)^2 (N \sigma^2) + N (N-1) \sigma^2 \right] = \frac{N-1}{N^2} [(N-1) + 1] \sigma^2 = \frac{N-1}{N^2} [(N-1) + 1] \sigma^2 = \frac{N-1}{N} \sigma^2$$
Error propagation

• From the theory of partial differential, if the errors of quantities, \(a, b, \ldots\) is known to be \(\Delta a, \Delta b, \ldots\) i.e. \(a = a_{\text{truth}} + \Delta a, b = b_{\text{truth}} + \Delta b\), and if a variable \(y\) can be calculated from \(a, b, \ldots\) then the error in \(y\) can be calculated by:

\[
\Delta y = \Delta a \left( \frac{\partial y}{\partial a} \right)_{b,c,\ldots} + \Delta b \left( \frac{\partial y}{\partial b} \right)_{a,c,\ldots} + \ldots
\]

• Now this is not realistic – we don’t usually know what the errors in \(a, b, \ldots\) are. If we do, we would have corrected the measured quantities so that we have exact values. But we often know the “expectation value” or “typical value.”

• i.e. \(\langle (\Delta a)^2 \rangle \equiv \sigma_a\) is known, and this is called the “error” in \(a\).

• Can we then estimate \(\sigma_y \equiv \langle (\Delta y)^2 \rangle\) from \(\sigma_a, \sigma_b, \ldots\)?

• By simply substituting the above expression in \(\sigma_y \equiv \langle (\Delta y)^2 \rangle\) we find:

\[
\sigma_y^2 = \langle (\Delta y)^2 \rangle = \left( \Delta a \left( \frac{\partial y}{\partial a} \right)_{b,c,\ldots} + \Delta b \left( \frac{\partial y}{\partial b} \right)_{a,c,\ldots} + \ldots \right) + \ldots
\]

\[
= \left( \left( \frac{\partial y}{\partial a} \right)_{b,c,\ldots} \left( \frac{\partial y}{\partial b} \right)_{a,c,\ldots} \ldots \right) \left( \begin{array}{cccc}
\langle \Delta a^2 \rangle & \langle \Delta a \cdot \Delta b \rangle & \ldots & \sigma_a^2 \\
\langle \Delta a \cdot \Delta b \rangle & \langle \Delta b^2 \rangle & \ldots & \sigma_{ab}^2 \\
\ldots & \ldots & \ldots & \ldots
\end{array} \right) \left( \begin{array}{cccc}
\left( \frac{\partial y}{\partial a} \right)_{b,c,\ldots} & \left( \frac{\partial y}{\partial b} \right)_{a,c,\ldots} & \ldots
\end{array} \right)
\]

where

\[
\begin{pmatrix}
\langle \Delta a^2 \rangle & \langle \Delta a \cdot \Delta b \rangle & \ldots \\
\langle \Delta a \cdot \Delta b \rangle & \langle \Delta b^2 \rangle & \ldots \\
\ldots & \ldots & \ldots
\end{pmatrix} = \begin{pmatrix}
\sigma_a^2 & \sigma_{ab}^2 & \ldots \\
\sigma_{ab}^2 & \sigma_b^2 & \ldots \\
\ldots & \ldots & \ldots
\end{pmatrix}
\]

is called error matrix.

• When the error in \(a\) and \(b\) are not correlated, \(\sigma_{ab} = 0\), and this particular off-diagonal component of the error matrix disappears. If \(a\) and \(b\) are independently measured quantity, this is probably the case. If \(a, b, \ldots\) are all uncorrelated, then

\[
\sigma_y^2 = \langle (\Delta y)^2 \rangle = \langle (\Delta a)^2 \rangle + \langle (\Delta b)^2 \rangle + \ldots
\]

• This form of error propagation is used often, partially because the independence of errors in \(a, b, \ldots\) is often assumed, not because it is proven, or even plausible.

Example of \(\sigma_{ab} \neq 0\)

• But if \(a\) and \(b\) are the \(x\) and \(y\) components of the momentum vector of a particle, this may not be so.

• Suppose you measure the coordinate of the path of a charged track at three places.
• where the curvature gives the momentum of the particle, and $\theta$ gives the direction.

• With appropriate approximations,

$$\frac{1}{R} \approx \frac{2x_2 + 2\Delta x_2 - \Delta x_1 - \Delta x_3}{l^2}$$

and $\theta \approx \frac{\Delta x_1 - \Delta x_3}{2l}$

$$\frac{\partial R^{-1}}{\partial x_1} = -\frac{1}{l^2}, \quad \frac{\partial R^{-1}}{\partial x_2} = \frac{2}{l^2}, \quad \text{and} \quad \frac{\partial R^{-1}}{\partial x_3} = -\frac{1}{l^2},$$

and $\frac{\partial \theta}{\partial x_1} = \frac{1}{2l}, \quad \frac{\partial \theta}{\partial x_2} = 0, \quad \text{and} \quad \frac{\partial \theta}{\partial x_3} = -\frac{1}{2l}$.

• This means that $\sigma_{R^{-1}}^2 = \left\langle (\Delta R^{-1})^2 \right\rangle = \left(\frac{1}{l^2}\right)^2 \sigma_{x_1}^2 + \left(\frac{2}{l^2}\right)^2 \sigma_{x_2}^2 + \left(\frac{1}{l^2}\right)^2 \sigma_{x_3}^2$

and $\sigma_{\theta}^2 = \left\langle (\Delta \theta)^2 \right\rangle = \left(\frac{1}{2l}\right)^2 \sigma_{x_1}^2 + \left(\frac{1}{2l}\right)^2 \sigma_{x_3}^2$.

• But in addition, $\sigma_{\theta R^{-1}}^2 = \left\langle (\Delta \theta)(\Delta R^{-1}) \right\rangle = \left(-\frac{1}{2l^3}\right) \sigma_{x_1}^2 + \left(\frac{1}{2l^3}\right) \sigma_{x_3}^2$.

which is not be zero unless the errors on $x_1$ and $x_3$ are exactly the same.