In the previous lecture, we learned about some of the properties of a particular ***linear*** function of random variables which did not depend on the explicit form of the P.D.F. If we also specify the P.D.F. we can go further...

Suppose there are two ***independent*** random variables x_1 and x_2 that have the same expectation value/true mean $\hat{x}$ and standard deviation σ_x (these may, as above, *e.g.* be repetitions of a single experiment).

We wish to calculate the P.D.F. of the ***sample mean*** $\bar{x} \equiv \tfrac{1}{2}(x_1 + x_2)$

We also define: $w \equiv \tfrac{1}{2}(x_1 - x_2)$

Then we also see that: $x_1 = \bar{x} + w$ and $x_2 = \bar{x} - w$.

Since x_1 and x_2 are (by assumption) ***independent***, their (joint) P.D.F. factorizes:

$$f(x_1, x_2) = f(x_1)\cdot f(x_2)$$

We now change variables from (x_1, x_2) to $(\bar{x}, w)$. Call the new P.D.F. $h(\bar{x}, w)$. What is $h(\bar{x}, w)$?

We know that: $h(\bar{x}, w)\,d\bar{x}\,dw = f(x_1, x_2)\,dx_1 dx_2$. From *p*. 10 of P598AEM Lecture Notes 3:

$$h(\bar{x}, w) = \frac{f(x_1, x_2)}{\left| J\left(\dfrac{\bar{x}\,w}{x_1 x_2}\right)\right|}$$ where the Jacobian determinant (here) is: $$|J| \equiv \begin{vmatrix} \dfrac{\partial \bar{x}}{\partial x_1} & \dfrac{\partial \bar{x}}{\partial x_2} \\ \dfrac{\partial w}{\partial x_1} & \dfrac{\partial w}{\partial x_2} \end{vmatrix} = \begin{vmatrix} +\dfrac{1}{2} & +\dfrac{1}{2} \\ +\dfrac{1}{2} & -\dfrac{1}{2} \end{vmatrix} = \left| -\frac{1}{2} \right|$$

Thus: $$h(\bar{x}, w) = \frac{f(x_1, x_2)}{\left|-\tfrac{1}{2}\right|} \frac{f(x_1, x_2)}{\tfrac{1}{2}} = 2 f(x_1, x_2) = 2 f(x_1)\cdot f(x_2)$$

Now if we want to find the P.D.F. of $\bar{x}$ alone, we must calculate the ***marginal distribution*** $H(\bar{x})$:

$$H(\bar{x}) = \int h(\bar{x}, w)\,dw = 2\int f(x_1, x_2)\,dw = 2\int f(x_1)\cdot f(x_2)\,dw$$

To go any further, we must specify a particular functional form for $f(x_1, x_2) = f(x_1)\cdot f(x_2)$.

Here, we assume that for both random variables x_1 & x_2 it is the Gaussian/normal distribution:

$$f(x_i) = \frac{1}{\sqrt{2\pi}\,\sigma_x} e^{-(x_i - \hat{x})^2 / 2\sigma_x^2} \quad \text{for } i = 1, 2.$$

Then:

$$
\begin{aligned}
H(\bar{x}) &= 2\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\,\sigma_x} e^{-(\bar{x}+w-\hat{x})^2/2\sigma_x^2} \cdot \frac{1}{\sqrt{2\pi}\,\sigma_x} e^{-(\bar{x}-w-\hat{x})^2/2\sigma_x^2} dw \\
&= \frac{2}{2\pi\,\sigma_x^2}\int_{-\infty}^{+\infty} e^{-[(\bar{x}+w-\hat{x})^2+(\bar{x}-w-\hat{x})^2]/2\sigma_x^2} dw \\
&= \frac{2}{2\pi\,\sigma_x{}^2}\int_{-\infty}^{+\infty} e^{-[2\bar{x}^2-4\hat{x}\bar{x}+2\hat{x}^2+2w^2]/2\sigma_x^2} dw \\
&= \frac{2}{2\pi\,\sigma_x{}^2} e^{-2(\bar{x}-\hat{x})^2/2\sigma_x^2}\int_{-\infty}^{+\infty} e^{-w^2/\sigma_x^2} dw \quad \Leftarrow \quad \text{use:} \quad \int_{-\infty}^{+\infty} e^{-ax^2}dx = 2\int_0^{+\infty} e^{-ax^2}dx = \sqrt{\tfrac{\pi}{a}} \\
&= \frac{1}{\sqrt{2\pi}\left(\sigma_x/\sqrt{2}\right)} e^{-(\bar{x}-\hat{x})^2/2\left(\sigma_x/\sqrt{2}\right)^2}
\end{aligned}
$$

We see that if we set: $\sigma_{\bar{x}} = \dfrac{\sigma_x}{\sqrt{2}}$ then: $H(\bar{x}) = \dfrac{1}{\sqrt{2\pi}\,\sigma_{\bar{x}}} e^{-(\bar{x}-\hat{x})^2/2\sigma_{\bar{x}}^2}$

So the ***sample mean*** $\bar{x}$ associated with $N = 2$ ***independent***, Gaussian/normally-distributed random variables x_1 & x_2 is ***also*** normally distributed, with expectation value/true mean $\hat{x}$.

Since $\bar{x}$ has standard deviation $\sigma_{\bar{x}} = \sigma_x/\sqrt{2} < \sigma_x$ we ***hope*** {quantitative statement later…} that $\bar{x}$ will lie closer to $\hat{x}$ than the individual measurements x_1 and/or x_2.

This result ***also*** holds for N ***independent***, Gaussian/normally-distributed random variables $x_1, x_2, ..., x_N$, with the result $\sigma_{\bar{x}} = \sigma_x/\sqrt{N}$ and $H(\bar{x})$ as given above.

This has been an example of one of the central tasks in "error analysis" – the determination (actually, the ***estimation***) of the variance of a function of random variables. For other than ***linear*** functions, life can be quite complicated…. *e.g.*:

Suppose we wish to estimate the variance of the ***product*** function $x_1 \cdot x_2$:

$$\text{var}(x_1 \cdot x_2) \equiv \sigma_{x_1x_2}^2 = E[(x_1 \cdot x_2 - E[x_1 \cdot x_2])^2]$$

To go further, we first need to compute the expectation value/true mean of $x_1 \cdot x_2$.

(Aside: we will (also) use the ***bracket*** notation for expectation values: $E[x] \equiv \hat{x} \equiv \langle x \rangle \equiv \mu$)

Thus: $E[x_1 \cdot x_2] = \int x_1 \cdot x_2 \; f(x_1, x_2)\, dx_1\, dx_2 = \langle x_1 \cdot x_2 \rangle$.

In order to carry out the integration, an explicit analytic form of the P.D.F. $f(x_1, x_2)$ is needed.

Suppose x_1 and x_2 are ***independent*** random variables. Then $f(x_1, x_2) = f(x_1) \cdot f(x_2)$, hence:

$$E[x_1 \cdot x_2] = E[x_1] \cdot E[x_2] = \langle x_1 \rangle \langle x_2 \rangle$$

thus:

$$\begin{aligned}\operatorname{var}(x_1 \cdot x_2) &= E[(x_1 \cdot x_2 - E[x_1 \cdot x_2])^2] = E[(x_1 \cdot x_2 - \langle x_1\rangle\langle x_2\rangle)^2] \\ &= E[x_1^2]\cdot E[x_2^2] - 2\langle x_1\rangle\langle x_2\rangle E[x_1]\cdot E[x_2] + \langle x_1\rangle^2\langle x_2\rangle^2 \\ &= E[x_1^2]\cdot E[x_2^2] - 2\langle x_1\rangle^2\langle x_2\rangle^2 + \langle x_1\rangle^2\langle x_2\rangle^2 \\ &= E[x_1^2]\cdot E[x_2^2] - \langle x_1\rangle^2\langle x_2\rangle^2\end{aligned}$$

But:

$$\operatorname{var}(x_1) \equiv \sigma_{x_1}^2 = E[x_1^2] - E[x_1]^2 = E[x_1^2] - \langle x_1\rangle^2$$

$$\operatorname{var}(x_2) \equiv \sigma_{x_2}^2 = E[x_2^2] - E[x_2]^2 = E[x_2^2] - \langle x_2\rangle^2$$

Then:

$$\begin{aligned}\operatorname{var}(x_1 \cdot x_2) \equiv \sigma_{x_1x_2}^2 &= E[x_1^2]\cdot E[x_2^2] - \langle x_1\rangle^2\langle x_2\rangle^2 \\ &= \left(\sigma_{x_1}^2 + \langle x_1\rangle^2\right)\left(\sigma_{x_2}^2 + \langle x_2\rangle^2\right) - \langle x_1\rangle^2\langle x_2\rangle^2\end{aligned}$$

Or:

$$\begin{aligned}\operatorname{var}(x_1 \cdot x_2) \equiv \sigma_{x_1x_2}^2 &= \sigma_{x_1}^2\sigma_{x_2}^2 + \langle x_1\rangle^2\sigma_{x_2}^2 + \langle x_2\rangle^2\sigma_{x_1}^2 + \cancel{\langle x_1\rangle^2\langle x_2\rangle^2} - \cancel{\langle x_1\rangle^2\langle x_2\rangle^2} \\ &= \sigma_{x_1}^2\sigma_{x_2}^2 + \langle x_1\rangle^2\sigma_{x_2}^2 + \langle x_2\rangle^2\sigma_{x_1}^2\end{aligned}$$

From: $\operatorname{var}(x_1 \cdot x_2) \equiv \sigma_{x_1x_2}^2 = \sigma_{x_1}^2\sigma_{x_2}^2 + \langle x_1\rangle^2\sigma_{x_2}^2 + \langle x_2\rangle^2\sigma_{x_1}^2$,

dividing both sides of this expression by $\langle x_1 \cdot x_2\rangle^2 = \langle x_1\rangle^2 \cdot \langle x_2\rangle^2$ (for ***independent*** random variables) we get:

$$\boxed{\frac{\sigma_{x_1x_2}^2}{\langle x_1 \cdot x_2\rangle^2} = \frac{\sigma_{x_1}^2}{\langle x_1\rangle^2} + \frac{\sigma_{x_2}^2}{\langle x_2\rangle^2} + \frac{\sigma_{x_1}^2\sigma_{x_2}^2}{\langle x_1\rangle^2\langle x_2\rangle^2} = \frac{\sigma_{x_1}^2}{\langle x_1\rangle^2} + \frac{\sigma_{x_2}^2}{\langle x_2\rangle^2} + \frac{\sigma_{x_1}^2}{\langle x_1\rangle^2}\frac{\sigma_{x_2}^2}{\langle x_2\rangle^2}}$$

The first two terms on the RHS are familiar from our childhood – the third term is not!

If, for a particular experiment $\frac{\sigma_{x_1}^2}{\langle x_1\rangle^2}$ and $\frac{\sigma_{x_2}^2}{\langle x_2\rangle^2}$ were both ***small enough*** that the third term:

$$\frac{\sigma_{x_1}^2}{\langle x_1\rangle^2}\frac{\sigma_{x_2}^2}{\langle x_2\rangle^2} \ll \text{ either } \frac{\sigma_{x_1}^2}{\langle x_1\rangle^2} \text{ or } \frac{\sigma_{x_2}^2}{\langle x_2\rangle^2}$$

then we get the familiar result: $$\frac{\sigma_{x_1x_2}^2}{\langle x_1 \cdot x_2\rangle^2} \cong \frac{\sigma_{x_1}^2}{\langle x_1\rangle^2} + \frac{\sigma_{x_2}^2}{\langle x_2\rangle^2}$$

Thus, we see that the familiar formula is actually an ***approximation*** that assumes:

- ***small fractional*** uncertainties
- ***independence*** of x_1 and x_2.

Let us now ***try*** to calculate the P.D.F. for $x_1 \cdot x_2$, as we did for $x_1 + x_2$.

Suppose that x_1 and x_2 are ***independent*** random variables and both are described by Gaussian / normal P.D.F.'s with the same expectation value $\hat{x}_1 = \hat{x}_2 = \hat{x}$ and the same variance $\sigma_{x_1}^2 = \sigma_{x_2}^2 = \sigma_x^2$.

So: $$f(x_1) = f(x_2) = f(x) = \frac{1}{\sqrt{2\pi}\,\sigma_x} e^{-(x-\hat{x})^2/2\sigma_x^2}$$

Thus:

$$f(x_1, x_2) = f(x_1)\cdot f(x_2) = \frac{1}{\sqrt{2\pi}\,\sigma_x} e^{-(x_1-\hat{x})^2/2\sigma_x^2} \frac{1}{\sqrt{2\pi}\,\sigma_x} e^{-(x_2-\hat{x})^2/2\sigma_x^2} = \frac{1}{2\pi\,\sigma_x^2} e^{-[(x_1-\hat{x})^2+(x_2-\hat{x})^2]/2\sigma_x^2}$$

Now make a change of variables: $h(u,v)\,du\,dv = f(x_1, x_2)\,dx_1\,dx_2$ where: $u \equiv x_1 \cdot x_2$ and: $v \equiv x_1$.

(*n.b.* The choice for *v* is ***arbitrary*** as long as x_1 and x_2 can be ***functionally*** related to *u* and *v*. This particular choice for *v* is the simplest...)

Then: $x_1 = v$, $x_2 = u/v$ and: $h(u,v) = \dfrac{f(x_1, x_2)}{\left| J\left(\dfrac{uv}{x_1 x_2}\right)\right|}$

Using: $|J| = \begin{vmatrix} \dfrac{\partial u}{\partial x_1} & \dfrac{\partial u}{\partial x_2} \\ \dfrac{\partial v}{\partial x_1} & \dfrac{\partial v}{\partial x_2} \end{vmatrix} = \begin{vmatrix} x_2 & x_1 \\ 1 & 0 \end{vmatrix} = |-x_1|$ we get:

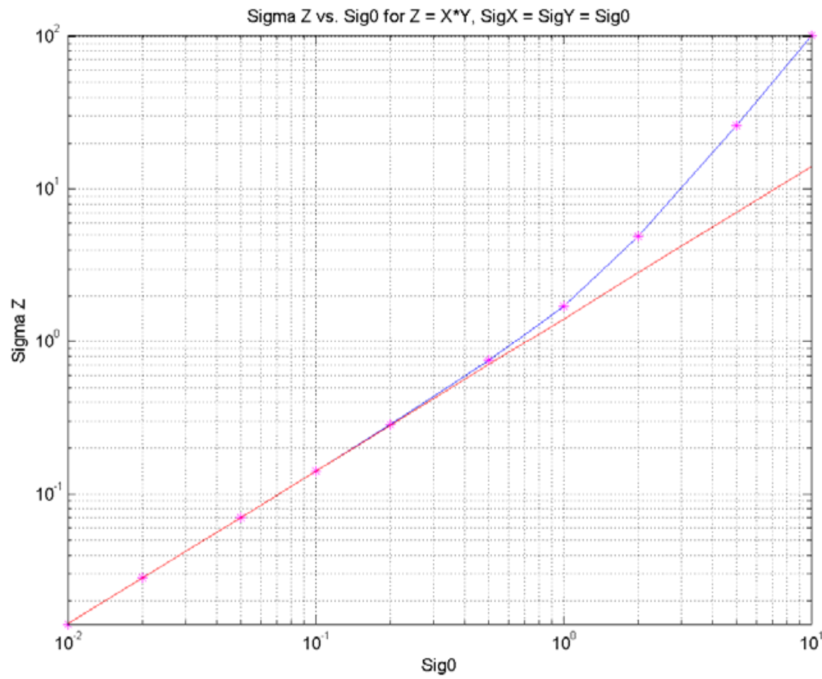
$$h(u,v) = \frac{1}{2\pi\sigma_x^2\,|x_1|} e^{-[(x_1-\hat{x})^2+(x_2-\hat{x})^2]/2\sigma_x^2} = \frac{1}{2\pi\sigma_x^2\,|v|} e^{-[(v-\hat{x})^2+(\frac{u}{v}-\hat{x})^2]/2\sigma_x^2}$$

If we want the P.D.F. for the variable $u \equiv x_1 \cdot x_2$ only, then we integrate over *v* to get:

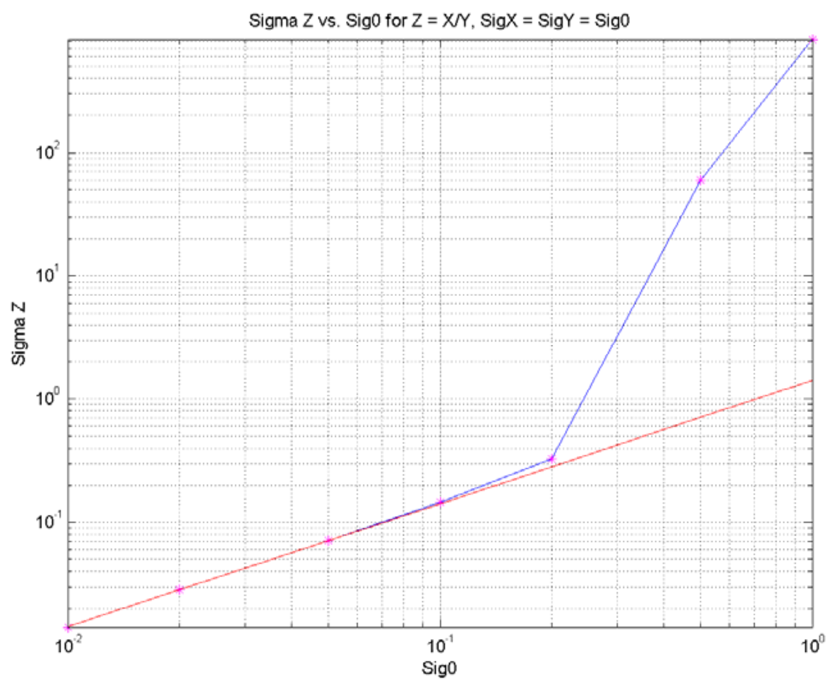
$$H(u) = \frac{1}{2\pi\sigma_x^2}\int_{-\infty}^{+\infty} \frac{dv}{|v|} e^{-[(v-\hat{x})^2+(\frac{u}{v}-\hat{x})^2]/2\sigma_x^2}$$

This is as far as we can go… carrying out the ***analytic*** integration of this integral is a mess... One can also see {here} that $H(u)$ is ***<u>not</u>*** the P.D.F. of a Gaussian/normal distribution!

In the two figures below, we show the overall uncertainty σ_Z *vs.* $\sigma_0 \left(= \sigma_x = \sigma_y\right)$ associated with the product (quotient) relation $Z = X \cdot Y$ ($Z = X/Y$) respectively, assuming X and Y are Gaussian-distributed ***independent*** random variables (*i.e.* are uncorrelated) with true means $\hat{X} = \hat{Y} = 1$ and with equal standard deviations $\sigma_x = \sigma_y = \sigma_0$, where $0.01 \le \sigma_0 \le 10.0$ {thus, the fractional sigmas on X and Y also vary from 0.01 to 10.0}. It can be seen that when $\sigma_x/\hat{X} \gtrsim 1$ or/and $\sigma_y/\hat{Y} \gtrsim 1$, the non-linear/cross-term in the boxed formula above on *p.* 3 of these lecture notes becomes increasingly important as the fractional sigmas become increasingly large.



σ_Z *vs.* $\sigma_0\left(=\sigma_x=\sigma_y\right)$ for $Z = X \cdot Y$



σ_Z *vs.* $\sigma_0\left(=\sigma_x=\sigma_y\right)$ for $Z = X/Y$

Note that σ_Z for the quotient $Z = X/Y$ ***really*** goes wild for $\sigma_0 \gg 0.1$!

With the above examples in mind, we consider the general problem of an arbitrary, not necessarily linear function $g(x_1, x_2, ..., x_N)$ of N random variables $(x_1, x_2, ..., x_N)$. We ***assume*** that the $x_1, x_2, ..., x_N$ all exist, and that $g(x_1, x_2, ..., x_N)$ can be expanded in a Taylor series about the ***expectation values*** $(\hat{x}_1, \hat{x}_2, ..., \hat{x}_N)$:

$$g(x_1, x_2, ..., x_N) = g(\hat{x}_1, \hat{x}_2, ..., \hat{x}_N) + \sum_{i=1}^{N} (x_i - \hat{x}_i) \left.\frac{\partial g}{\partial x_i}\right|_{all\ x_i = \hat{x}_i} + \frac{1}{2!} \sum_{i=1}^{N} \sum_{j=1}^{N} (x_i - \hat{x}_i)(x_j - \hat{x}_j) \left.\frac{\partial g}{\partial x_i} \frac{\partial g}{\partial x_j}\right|_{all\ x_i = \hat{x}_i, x_j = \hat{x}_j} + ...$$

Note that the local slope(s) of $g(x_1, x_2, ..., x_N)$ {which are simply ***numbers*** !} are evaluated at their ***expectation values/true means*** $\frac{\partial g}{\partial \hat{x}_i} \equiv \left.\frac{\partial g}{\partial x_i}\right|_{all\ x_i = \hat{x}_i}$

Furthermore, we ***assume*** that ***all*** terms in $(x_i - \hat{x}_i)^2$, $(x_i - \hat{x}_i)(x_j - \hat{x}_j)$ .and. higher orders can be ***neglected***. *n.b.* This means that the results are valid ***only*** in the limit where the measurements $x_1, x_2, ..., x_N$ are never "very far" from their expectation values $\hat{x}_1, \hat{x}_2, ..., \hat{x}_N$. In turn, this implies that the corresponding standard deviations σ_{x_i} must be "small enough", as in the previous case, *i.e.* $\left(\sigma_{x_i}^2 / \hat{x}_i^2\right) \ll 1$. Under these assumptions, then $g(x_1, x_2, \cdots, x_N)$ becomes a ***linearized*** function of the random variables $(x_1, x_2, ..., x_N)$:

$$g(x_1, x_2, \cdots, x_N) \cong g(\hat{x}_1, \hat{x}_2, \cdots, \hat{x}_N) + \sum_{i=1}^{N} (x_i - \hat{x}_i) \frac{\partial g}{\partial \hat{x}_i}$$

Using this approximation we see that:

$$E[g] = E[g(x_1, x_2, \cdots, x_N)] \cong g(\hat{x}_1, \hat{x}_2, \cdots, \hat{x}_N) + \sum_{i=1}^{N} E[x_i - \hat{x}_i] \frac{\partial g}{\partial \hat{x}_i}$$

So we see that: $E[g] \cong g(\hat{x}_1, \hat{x}_2, \cdots, \hat{x}_N)$ since: $E[x_i - \hat{x}_i] = E[x_i] - \hat{x}_i = \hat{x}_i - \hat{x}_i = 0$.

Note that we have not yet made ***any*** assumptions about the independence or dependence of the random variables x_i, nor anything about possible analytical representations of their P.D.F.'s.

The ***variance*** of $g(x_1, x_2, \cdots, x_N)$, using the above assumptions, is:

$$\begin{aligned}
\mathrm{var}(g(x_1, x_2, \cdots, x_N)) = \mathrm{var}(g) \equiv \sigma_g^2 &\equiv E\left[(g - E[g])^2\right] \\
&\cong E[(g(x_1, x_2, \cdots, x_N) - g(\hat{x}_1, \hat{x}_2, \cdots, \hat{x}_N))^2] \\
&= E\left[\left(\sum_{i=1}^{N} (x_i - \hat{x}_i) \frac{\partial g}{\partial \hat{x}_i}\right)^2\right] \\
&= E\left[\sum_{i=1}^{N} (x_i - \hat{x}_i) \frac{\partial g}{\partial \hat{x}_i} \sum_{j=1}^{N} (x_j - \hat{x}_j) \frac{\partial g}{\partial \hat{x}_j}\right]
\end{aligned}$$

So: $$\operatorname{var}\left(g\left(x_1, x_2, \cdots, x_N\right)\right) \cong \sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\partial g}{\partial \hat{x}_i}\frac{\partial g}{\partial \hat{x}_j}E[(x_i-\hat{x}_i)(x_j-\hat{x}_j)] = \sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\partial g}{\partial \hat{x}_i}\frac{\partial g}{\partial \hat{x}_j}\operatorname{cov}\left(x_i, x_j\right)$$

where: $\operatorname{cov}\left(x_i, x_j\right) \equiv E[(x_i-\hat{x}_i)(x_j-\hat{x}_j)]$.

n.b. In the above definition of $\operatorname{var}(g)$ we have "liberalized" our previous notation to include the case where *i* = *j*, where $\operatorname{cov}\left(x_i, x_i\right) \equiv \operatorname{var}\left(x_i\right) \equiv \sigma_{x_i}^2$.

Also, one also needs to {always!} keep in mind that the "$\cong$" symbol above means that the equalities are true ***only*** in the limit that the 2^{nd} (.and. all higher-order) derivative terms in the Taylor's series expansion of $g\left(x_1, x_2, \cdots, x_N\right)$ can ***indeed*** be neglected…

In the special/limiting case where the x_i are ***independent*** random variables, noting that: $\operatorname{cov}\left(x_i, x_j\right) = 0$ for $i \neq j$ and that $\operatorname{cov}\left(x_i, x_i\right) \equiv \operatorname{var}\left(x_i\right) \equiv \sigma_{x_i}^2$ for $i = j$, then for ***independent*** random variables:

$$\operatorname{var}(g) \equiv \sigma_g^2 \cong \sum_{i=1}^{N}\frac{\partial g}{\partial \hat{x}_i}\frac{\partial g}{\partial \hat{x}_i}E[(x_i-\hat{x}_i)(x_i-\hat{x}_i)] = \sum_{i=1}^{N}\frac{\partial g}{\partial \hat{x}_i}\frac{\partial g}{\partial \hat{x}_i}\operatorname{cov}\left(x_i, x_i\right) = \sum_{i=1}^{N}\left(\frac{\partial g}{\partial \hat{x}_i}\right)^2 \operatorname{var}\left(x_i\right)$$

In more familiar notation: $\sigma_{g(x_1,x_2,..,x_N)}^2 \equiv \sigma_g^2 \cong \sum_{i=1}^{N}\left(\frac{\partial g}{\partial x_i}\right)^2 \sigma_{x_i}^2$

For ***independent*** random variables (*i.e.* no correlations), a few simple examples of $\sigma_{g(x_1,x_2,..,x_N)}^2 \equiv \sigma_g^2 \cong \sum_{i=1}^{N}\left(\frac{\partial g}{\partial x_i}\right)^2 \sigma_{x_i}^2$ follow:

• ***g* = *x*₁ ± *x*₂ :**

$$\frac{\partial g}{\partial x_1} = 1 \qquad \frac{\partial g}{\partial x_2} = \pm 1 \qquad \text{Thus: } \sigma_{x_1 \pm x_2}^2 = \sigma_{x_1}^2 + \sigma_{x_2}^2 \quad \Leftarrow \quad \textit{n.b. } \text{no terms above } 2^{\text{nd}} \text{ order}$$

• ***g* = *x*₁ · *x*₂ :**

$$\frac{\partial g}{\partial x_1} = x_2 \qquad \frac{\partial g}{\partial x_2} = x_1 \qquad \text{Thus: } \sigma_{x_1 \cdot x_2}^2 = \hat{x}_2^2 \sigma_{x_1}^2 + \hat{x}_1^2 \sigma_{x_2}^2 \qquad \text{or: } \frac{\sigma_{x_1 \cdot x_2}^2}{\left(\hat{x}_1 \cdot \hat{x}_2\right)^2} = \frac{\sigma_{x_1}^2}{\hat{x}_1^2} + \frac{\sigma_{x_2}^2}{\hat{x}_2^2}$$

• ***g* = *x*₁ / *x*₂ :**

$$\frac{\partial g}{\partial x_1} = \frac{1}{x_2} \qquad \frac{\partial g}{\partial x_2} = -\frac{x_1}{x_2^2} \qquad \text{Thus: } \sigma_{x_1/x_2}^2 \cong \frac{\sigma_{x_1}^2}{\hat{x}_2^2} + \hat{x}_1^2\frac{\sigma_{x_2}^2}{\hat{x}_2^4} \qquad \text{or: } \frac{\sigma_{x_1/x_2}^2}{\left(x_1/x_2\right)^2} \cong \frac{\sigma_{x_1}^2}{\hat{x}_1^2} + \frac{\sigma_{x_2}^2}{\hat{x}_2^2}$$

Now suppose that x_1 and x_2 are ***not*** independent random variables – *i.e*. there exist ***correlations*** between them. Then, with $\rho(x_1, x_2) \equiv \text{cov}(x_1, x_2) \big/ (\sigma_{x_1}\sigma_{x_2})$:

• ***g* = *x*₁ ± *x*₂ :**

$$\sigma^2_{x_1 \pm x_2} \cong \sum_{i=1}^{2}\sum_{j=1}^{2} \frac{\partial g}{\partial \hat{x}_i}\frac{\partial g}{\partial \hat{x}_j}\text{cov}(x_i, x_j) = \left(\frac{\partial g}{\partial \hat{x}_1}\right)^2 \sigma^2_{x_1} + \left(\frac{\partial g}{\partial \hat{x}_2}\right)^2 \sigma^2_{x_2} + 2\frac{\partial g}{\partial \hat{x}_1}\frac{\partial g}{\partial \hat{x}_2}\text{cov}(x_1, x_2)$$

$$= \sigma^2_{x_1} + \sigma^2_{x_2} \pm 2\,\text{cov}(x_1, x_2) = \sigma^2_{x_1} + \sigma^2_{x_2} \pm 2\sigma_{x_1}\sigma_{x_2}\rho(x_1, x_2)$$

• ***g* = *x*₁ · *x*₂ :**

$$\sigma^2_{x_1 \cdot x_2} \cong \sum_{i=1}^{2}\sum_{j=1}^{2} \frac{\partial g}{\partial \hat{x}_i}\frac{\partial g}{\partial \hat{x}_j}\text{cov}(x_i, x_j) = \left(\frac{\partial g}{\partial \hat{x}_1}\right)^2 \sigma^2_{x_1} + \left(\frac{\partial g}{\partial \hat{x}_2}\right)^2 \sigma^2_{x_2} + 2\frac{\partial g}{\partial \hat{x}_1}\frac{\partial g}{\partial \hat{x}_2}\text{cov}(x_1, x_2)$$

$$= \hat{x}_2^2\sigma^2_{x_1} + \hat{x}_1^2\sigma^2_{x_2} + 2\hat{x}_2\hat{x}_1\,\text{cov}(x_1, x_2) = \hat{x}_2^2\sigma^2_{x_1} + \hat{x}_1^2\sigma^2_{x_2} + 2\hat{x}_1\hat{x}_2\sigma_{x_1}\sigma_{x_2}\rho(x_1, x_2)$$

So: $$\frac{\sigma^2_{x_1 x_2}}{(\hat{x}_1 \cdot \hat{x}_2)^2} \cong \frac{\sigma^2_{x_1}}{\hat{x}_1^2} + \frac{\sigma^2_{x_2}}{\hat{x}_2^2} + 2\left(\frac{1}{\hat{x}_1 \cdot \hat{x}_2}\right)\text{cov}(x_1, x_2) = \frac{\sigma^2_{x_1}}{\hat{x}_1^2} + \frac{\sigma^2_{x_2}}{\hat{x}_2^2} + 2\left(\frac{\sigma_{x_1} \cdot \sigma_{x_2}}{\hat{x}_1 \cdot \hat{x}_2}\right)\rho(x_1, x_2)$$

• ***g* = *x*₁ / *x*₂ :**

$$\sigma^2_{x_1/x_2} \cong \sum_{i=1}^{2}\sum_{j=1}^{2} \frac{\partial g}{\partial \hat{x}_i}\frac{\partial g}{\partial \hat{x}_j}\text{cov}(x_i, x_j) = \left(\frac{\partial g}{\partial \hat{x}_1}\right)^2 \sigma^2_{x_1} + \left(\frac{\partial g}{\partial \hat{x}_2}\right)^2 \sigma^2_{x_2} + 2\frac{\partial g}{\partial \hat{x}_1}\frac{\partial g}{\partial \hat{x}_2}\text{cov}(x_1, x_2)$$

$$= \frac{\sigma^2_{x_1}}{\hat{x}_2^2} + \frac{\hat{x}_1^2\sigma^2_{x_2}}{\hat{x}_2^4} + 2\left(\frac{1}{\hat{x}_2}\right)\left(\frac{-\hat{x}_1}{\hat{x}_2^2}\right)\text{cov}(x_1, x_2) = \frac{\sigma^2_{x_1}}{\hat{x}_2^2} + \frac{\hat{x}_1^2\sigma^2_{x_2}}{\hat{x}_2^4} - 2\left(\frac{\hat{x}_1\sigma_{x_1}\sigma_{x_2}}{\hat{x}_2^3}\right)\rho(x_1, x_2)$$

So: $$\frac{\sigma^2_{x_1/x_2}}{(\hat{x}_1/\hat{x}_2)^2} \cong \frac{\sigma^2_{x_1}}{\hat{x}_1^2} + \frac{\sigma^2_{x_2}}{\hat{x}_2^2} - 2\frac{1}{\hat{x}_1 \cdot \hat{x}_2}\text{cov}(x_1, x_2) = \frac{\sigma^2_{x_1}}{\hat{x}_1^2} + \frac{\sigma^2_{x_2}}{\hat{x}_2^2} - 2\frac{\sigma_{x_1}\sigma_{x_2}}{\hat{x}_1 \cdot \hat{x}_2}\rho(x_1, x_2)$$

We now introduce the notation of an **under-bar** to denote a ***vector*** or ***matrix*** quantity:

Let $$\underline{x} \equiv \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix}$$ be a column vector ($N \times 1$ matrix) of N ***random variables*** x_i and:

Let $$\hat{\underline{x}} \equiv E[\underline{x}] \equiv \begin{pmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \vdots \\ \hat{x}_N \end{pmatrix}$$ be a column vector ($N \times 1$ matrix) of ***expectation values*/*true means*** $\hat{x}_i$.

We also define a column vector ($N \times 1$ matrix) of "***residuals***" and its $1 \times N$ row vector ***transpose***:

$$\underline{R}_x \equiv \left(\underline{x} - \hat{\underline{x}}\right) \equiv \left(\underline{x} - E[\underline{x}]\right) \equiv \begin{pmatrix} x_1 - \hat{x}_1 \\ x_2 - \hat{x}_2 \\ \vdots \\ x_N - \hat{x}_N \end{pmatrix} \quad \text{and:} \quad \underline{R}_x^T \equiv \left(\underline{x} - \hat{\underline{x}}\right)^T \equiv \left(\underline{x} - E[\underline{x}]\right)^T \equiv \begin{pmatrix} x_1 - \hat{x}_1 \\ x_2 - \hat{x}_2 \\ \vdots \\ x_N - \hat{x}_N \end{pmatrix}^T$$

We can then additionally define an $N \times N$ matrix by taking the ***outer product*** of the ***residual*** vector with its ***transpose*** (a row vector):

$$\underline{R}_x \underline{R}_x^T \equiv \underbrace{\left(\underline{x} - \hat{\underline{x}}\right)}_{N\times 1} \underbrace{\left(\underline{x} - \hat{\underline{x}}\right)^T}_{1\times N} = \underbrace{\begin{pmatrix} (x_1-\hat{x}_1)(x_1-\hat{x}_1) & (x_1-\hat{x}_1)(x_2-\hat{x}_2) & \cdots & (x_1-\hat{x}_1)(x_N-\hat{x}_N) \\ (x_2-\hat{x}_2)(x_1-\hat{x}_1) & (x_2-\hat{x}_2)(x_2-\hat{x}_2) & \cdots & (x_2-\hat{x}_2)(x_N-\hat{x}_N) \\ \vdots & \vdots & \ddots & \vdots \\ (x_N-\hat{x}_N)(x_1-\hat{x}_1) & (x_N-\hat{x}_N)(x_2-\hat{x}_2) & \cdots & (x_N-\hat{x}_N)(x_N-\hat{x}_N) \end{pmatrix}}_{N\times N}.$$

The ***expectation value*** of ***this*** matrix: $\hat{\underline{V}}_x \equiv E[\underline{R}_x \underline{R}_x^T] \equiv E[\left(\underline{x} - \hat{\underline{x}}\right)\left(\underline{x} - \hat{\underline{x}}\right)^T]$ is an $N \times N$ matrix known as the "***covariance matrix***" (or "***variance matrix***", if purely diagonal) or simply / generically as the "***error matrix***".

$$\hat{\underline{V}}_x \equiv E[\underline{R}_x \underline{R}_x^T] \equiv E[\left(\underline{x} - \hat{\underline{x}}\right)\left(\underline{x} - \hat{\underline{x}}\right)^T]$$
$$= \begin{pmatrix} \sigma_{x_1}^2 & \mathrm{cov}(x_1, x_2) & \ldots & \mathrm{cov}(x_1, x_N) \\ \mathrm{cov}(x_2, x_1) & \sigma_{x_2}^2 & \vdots & \mathrm{cov}(x_2, x_N) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{cov}(x_N, x_1) & \mathrm{cov}(x_N, x_2) & \ldots & \sigma_{x_N}^2 \end{pmatrix} = \begin{pmatrix} \sigma_{x_1}^2 & \sigma_{x_1 x_2}^2 & \ldots & \sigma_{x_1 x_N}^2 \\ \sigma_{x_2 x_1}^2 & \sigma_{x_2}^2 & \vdots & \sigma_{x_2 x_N}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{x_N x_1}^2 & \sigma_{x_N x_2}^2 & \ldots & \sigma_{x_N}^2 \end{pmatrix}$$

Note that (by definition) the $N \times N$ matrix $\hat{\underline{V}}_x$ is ***real*** and ***symmetric***, *i.e.* note that:

$$\mathrm{cov}(x_j, x_i) = \sigma_{x_j x_i}^2 \quad = \quad \mathrm{cov}(x_i, x_j) = \sigma_{x_i x_j}^2 \quad = \quad E[\left(x_i - \hat{x}_i\right)\left(x_j - \hat{x}_j\right)]$$