If we <u>subtract</u> two variables we get a similar result:

$$z_i = x_i - y_i$$

then $\bar{z} = \bar{x} - \bar{y}$; $s_z^2 = s_x^2 + s_y^2 - 2 s_{xy}^2$

note (−) sign!

A positive covariance (note: s_{xy}^2 may be + or −) <u>reduces</u> the error in the difference between variables.

OK, how about multiplication and division?

Let $z_i = x_i / y_i$

where x_i and y_i are normally distributed.

Note that z is <u>not</u> normally distributed! It only approaches this <u>if</u> $s_x / \bar{x}$ and $s_y / \bar{y} \ll 1$!

Let's look at this more closely.

Let $x_i = \bar{x} + x_i'$

$\bar{x}$: mean, x_i': deviation

and $y_i = \bar{y} + y_i'$

Suppose $s_x/\bar{x} \ll 1$ and $s_y/\bar{y} \ll 1$

This means that x_i' and y_i' are small:

$$z_i = \frac{x_i}{y_i} = \frac{\bar{x} + x_i'}{\bar{y} + y_i'} = \frac{\bar{x}}{\bar{y}} \frac{\left(1 + x_i'/\bar{x}\right)}{\left(1 + y_i'/\bar{y}\right)}$$

$$\approx \frac{\bar{x}}{\bar{y}} \left(1 + \frac{x_i'}{\bar{x}}\right)\left(1 - \frac{y_i'}{\bar{y}} + O\left(\left(\frac{y_i'}{\bar{y}}\right)^2\right)\right)$$

$$\approx \frac{\bar{x}}{\bar{y}} \left(1 + \frac{x_i'}{\bar{x}} - \frac{y_i'}{\bar{y}} + O\left(\frac{x_i' y_i'}{\bar{x}\bar{y}}, \left(\frac{y_i'}{\bar{y}}\right)^2\right)\right)$$

We ignore the higher order terms!

$$\bar{z} \equiv \frac{1}{N}\sum_{i=1}^{N} \frac{x_i}{y_i} \approx \frac{\bar{x}}{\bar{y}}\,\frac{1}{N}\sum_{i=1}^{N}\left(1 + \cancel{\frac{x_i'}{\bar{x}}} + \cancel{\frac{y_i'}{\bar{y}}}\right)$$

Sum to zero

$$\therefore\ \bar{z} \approx \frac{\bar{x}}{\bar{y}}$$

Now for s_z^2:

$$s_z^2 \approx \underbrace{\left(\frac{\bar{x}^2}{\bar{y}^2}\right)}_{\approx \bar{z}^2}\left(\frac{s_x^2}{\bar{x}^2} + \frac{s_y^2}{\bar{y}^2} - 2\,\frac{s_{xy}^2}{(\bar{x}\,\bar{y})}\right)$$

Thus:

$$\frac{s_z^2}{\bar{z}^2} \cong \frac{s_x^2}{\bar{x}^2} + \frac{s_y^2}{\bar{y}^2} - 2\,\frac{s_{xy}^2}{\bar{x}\,\bar{y}}$$

so for division you add the fractional or relative variance!

For multiplication you get the same:

$$z_i \equiv x_i y_i$$

$$\bar{z} \approx \bar{x}\,\bar{y} \quad \text{\underline{provided}} \quad \frac{s_x}{\bar{x}}, \frac{s_y}{\bar{y}} \ll 1$$

and

$$\frac{s_z^2}{\bar{z}^2} \cong \frac{s_x^2}{\bar{x}^2} + \frac{s_y^2}{\bar{y}^2} + 2\,\frac{s_{xy}^2}{\bar{x}\,\bar{y}}$$

note (+) sign

What about a more general functional relationship?

Suppose $z_i = f(x_i, y_i)$ where f is some nasty function

Let $x_i = \bar{x} + x_i'$, $y_i = \bar{y} + y_i'$

We shall <u>expand</u> f in a 2-D

Taylor series (Note: the T.S. can be generalized to n-dimensions)

So:

$$f(x_i, y_i) = f(\bar{x}, \bar{y}) + x_i' \left.\frac{\partial f}{\partial x}\right|_{\bar{x},\bar{y}} + y_i' \left.\frac{\partial f}{\partial y}\right|_{\bar{x},\bar{y}} + O(x_i'^2, y_i'^2, x_i' y_i')$$

(Ignore 2^{nd} order terms!)

$$\therefore \bar{z} \equiv \frac{1}{N}\sum_{i=1}^{N} f(x_i, y_i)$$

$$\approx f(\bar{x}, \bar{y}) + \left.\frac{\partial f}{\partial x}\right|_{\bar{x},\bar{y}} \frac{1}{N}\sum_{i=1}^{N} x_i' + \left.\frac{\partial f}{\partial y}\right|_{\bar{x},\bar{y}} \frac{1}{N}\sum_{i=1}^{N} y_i'$$

sum to zero!

$$= f(\bar{x}, \bar{y})$$

The <u>error</u> is $O(x_i'^2, y_i'^2, x_i' y_i' \cdot \nabla\nabla f)$

dyadic

Similarly, we get for the variance:

$$s_z^2 \cong \left(\frac{\partial f}{\partial x}\bigg|_{\bar{x},\bar{y}}\right)^2 s_x^2 + \left(\frac{\partial f}{\partial y}\bigg|_{\bar{x},\bar{y}}\right)^2 s_y^2$$
$$+ 2\left(\frac{\partial f}{\partial x}\bigg|_{\bar{x},\bar{y}}\right)\left(\frac{\partial f}{\partial y}\bigg|_{\bar{x},\bar{y}}\right) s_{xy}^2$$

again providing that the higher order terms can be neglected.

This formula reduces to the earlier ones for addition, subtr., mult. & div.

Only for add^n & subtr. is it <u>exact</u> because $\nabla\nabla f \equiv 0$ for this case and the higher order terms <u>vanish</u>.

Let's generalize this result to an n-dimensional (n-variable) system

Let $\underset{\sim}{x} = x_1, x_2, \ldots, x_n$

$z = f(\underset{\sim}{x})$

we need to find the variance of each variable x_1, x_2, etc. <u>and</u> the covariance of all pairs $S^2_{x_1 x_2}$, etc.

We can define the <u>matrix</u> of variance & covariance

$$V_{ij} = \frac{N}{N-1}\left(\overline{x_i x_j} - \bar{x}_i \bar{x}_j\right)$$

number of elements contributing to each x_i, x_j

Now <u>if</u> $\frac{V_{ii}}{\bar{x}_i^2} << 1$ for all i

(this corresponds to a small error in each variable, e.g. $\frac{S_x^2}{\bar{x}^2} << 1$, $\frac{S_y^2}{\bar{y}^2} << 1$, etc.)

And if we define:

$$\nabla f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_3}, \ldots \frac{\partial f}{\partial x_n}\right)$$

Then we get the result:

$$S_z^2 = (\underset{\sim}{\nabla} f)\ \underset{\approx}{V}\ (\underset{\sim}{\nabla} f)^T$$

Which gives us an easy way of calculating the variance in $z = f(\underset{\sim}{x})$!

Note: in this derivation the subscripts i and j denote different <u>variables</u> entirely, <u>not</u> different measurements contributing to the same average.

Let's look at the example $f(\underset{\sim}{x}) = x_1 x_2$

Let $x_1 \equiv x$, $x_2 \equiv y$

$$\therefore\ V_{ij} = \frac{N}{N-1}\left(\overline{x_i x_j} - \bar{x}_i \bar{x}_j\right)$$

so $$V_{xx} = \frac{N}{N-1}\left(\overline{xx} - \bar{x}\,\bar{x}\right)$$

$$= \frac{1}{N-1} \sum^{N} \left(x^2 - \bar{x}^2\right) = S_x^2$$

$$V_{xy} = \frac{N}{N-1}\left(\overline{xy} - \bar{x}\,\bar{y}\right) = \frac{1}{N-1}\sum^{N}(xy - \bar{x}\bar{y})$$
$$\equiv S_{xy}^2$$

$$\underset{\approx}{V} = \begin{pmatrix} S_x^2 & S_{xy}^2 \\ S_{xy}^2 & S_y^2 \end{pmatrix}$$

$$\underset{\sim}{\nabla} f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) \equiv (y, x)$$

$$\underset{\sim}{\nabla} f \; \underset{\approx}{V} \; (\underset{\sim}{\nabla} f)^T = (y \;\; x)\; \underset{\approx}{V} \begin{pmatrix} y \\ x \end{pmatrix}$$

$$= \left(y S_x^2 + x S_{xy}^2,\; y S_{xy}^2 + x S_y^2\right)\begin{pmatrix} y \\ x \end{pmatrix}$$

$$= y^2 S_x^2 + xy S_{xy}^2 + xy S_{xy}^2 + x^2 S_y^2$$

$$= y^2 S_x^2 + x^2 S_y^2 + 2xy S_{xy}^2$$

so

$$\frac{S_z^2}{(xy)^2} = \frac{S_x^2}{x^2} + \frac{S_y^2}{y^2} + 2\frac{S_{xy}^2}{xy} \quad !$$

Finally, we examine vector functions of vector variables:

Let $\underset{\sim}{z} = \underset{\sim}{f}(\underset{\sim}{x})$ be a vector function of $\underset{\sim}{x}$. We define:

$$\underset{\sim}{\mu}_z \equiv E(\underset{\sim}{z}), \quad \underset{\sim}{\mu}_x = E(\underset{\sim}{x})$$

and the matrix of covariance for each:

$$\underset{\approx}{\Sigma}_x^2 \equiv E\left[(\underset{\sim}{x} - \underset{\sim}{\mu}_x)(\underset{\sim}{x} - \underset{\sim}{\mu}_x)^T\right]$$

$$\underset{\approx}{\Sigma}_z^2 = E\left[(\underset{\sim}{z} - \underset{\sim}{\mu}_z)(\underset{\sim}{z} - \underset{\sim}{\mu}_z)^T\right]$$

we have taken $\underset{\sim}{z}$ and $\underset{\sim}{x}$ to be column vectors.

Suppose we know $\underset{\sim}{x}$ and $\underset{\approx}{\Sigma}_x^2$. We wish to calculate $\underset{\approx}{\Sigma}_z^2$.

We do a multivariable Taylor series expansion of $\underset{\sim}{z}$ about $\underset{\sim}{\mu}_x$:

$$\underset{\sim}{z} = \underset{\sim}{f}(\underset{\sim}{x}) = \underset{\sim}{f}(\underset{\sim}{\mu}_x) + \underset{\sim}{\nabla}\underset{\sim}{f}\Big|_{\underset{\sim}{\mu}_x} \cdot (\underset{\sim}{x} - \underset{\sim}{\mu}_x)$$
$$+ \tfrac{1}{2}\underset{\sim}{\nabla}\underset{\sim}{\nabla}\underset{\sim}{f} : (\underset{\sim}{x} - \underset{\sim}{\mu}_x)(\underset{\sim}{x} - \underset{\sim}{\mu}_x)^T + \ldots$$

If the deviation between $\underset{\sim}{x}$ and $\underset{\sim}{\mu}_x$ is small (e.g. that $\underset{\approx}{\Sigma}_x^2$ is small), or $\underset{\sim}{f}(\underset{\sim}{x})$ is linear in $\underset{\sim}{x}$, then we can ignore the quadratic or higher order terms in the Taylor series!

Thus:

$$\underset{\sim}{z} \cong \underset{\sim}{f}(\underset{\sim}{\mu}_x) + \underset{\sim}{\nabla}\underset{\sim}{f}\Big|_{\underset{\sim}{\mu}_x} \cdot (\underset{\sim}{x} - \underset{\sim}{\mu}_x)$$

and thus:

$$\underset{\sim}{\mu}_z \equiv E(\underset{\sim}{z}) \cong \underset{\sim}{f}(\underset{\sim}{\mu}_x) + \underset{\sim}{\nabla}\underset{\sim}{f}\Big|_{\underset{\sim}{\mu}_x} \cdot E(\underset{\sim}{x} - \underset{\sim}{\mu}_x)$$

but $E(\underset{\sim}{x} - \underset{\sim}{\mu}_x) \equiv 0$ by definition!

Thus $\underset{\sim}{\mu}_z \cong \underset{\sim}{f}(\underset{\sim}{\mu}_x)$ + quadratic terms!

Putting this together, we get:

$$\underset{\sim}{z} - \underset{\sim}{\mu}_z \simeq \nabla \underset{\sim}{f}\Big|_{\underset{\sim}{\mu}_x} (\underset{\sim}{x} - \underset{\sim}{\mu}_x)$$

So:

$$\underset{\approx}{\Sigma}_z^2 \equiv E\left((\underset{\sim}{z} - \underset{\sim}{\mu}_z)(\underset{\sim}{z} - \underset{\sim}{\mu}_z)^T\right)$$

$$\cong E\left[(\underset{\sim}{\nabla} \underset{\sim}{f})\,(\underset{\sim}{x} - \underset{\sim}{\mu}_x)(\underset{\sim}{x} - \underset{\sim}{\mu}_x)^T\,(\underset{\sim}{\nabla} \underset{\sim}{f})^T\right]$$

$$= (\underset{\sim}{\nabla} \underset{\sim}{f})\; E\left((\underset{\sim}{x} - \underset{\sim}{\mu}_x)(\underset{\sim}{x} - \underset{\sim}{\mu}_x)^T\right)(\underset{\sim}{\nabla} \underset{\sim}{f})^T$$

$$= (\underset{\sim}{\nabla} \underset{\sim}{f})\; \underset{\approx}{\Sigma}_x^2\; (\underset{\sim}{\nabla} \underset{\sim}{f})^T$$

<u>provided</u> that $\frac{1}{2} \underset{\sim}{\nabla} \underset{\sim}{\nabla} \underset{\sim}{f} : \underset{\approx}{\Sigma}_x^2 \ll \underset{\sim}{f}$

so that we can neglect the quadratic term in the Taylor series expansion.

We will use this formulation again in calculating linear regression error!