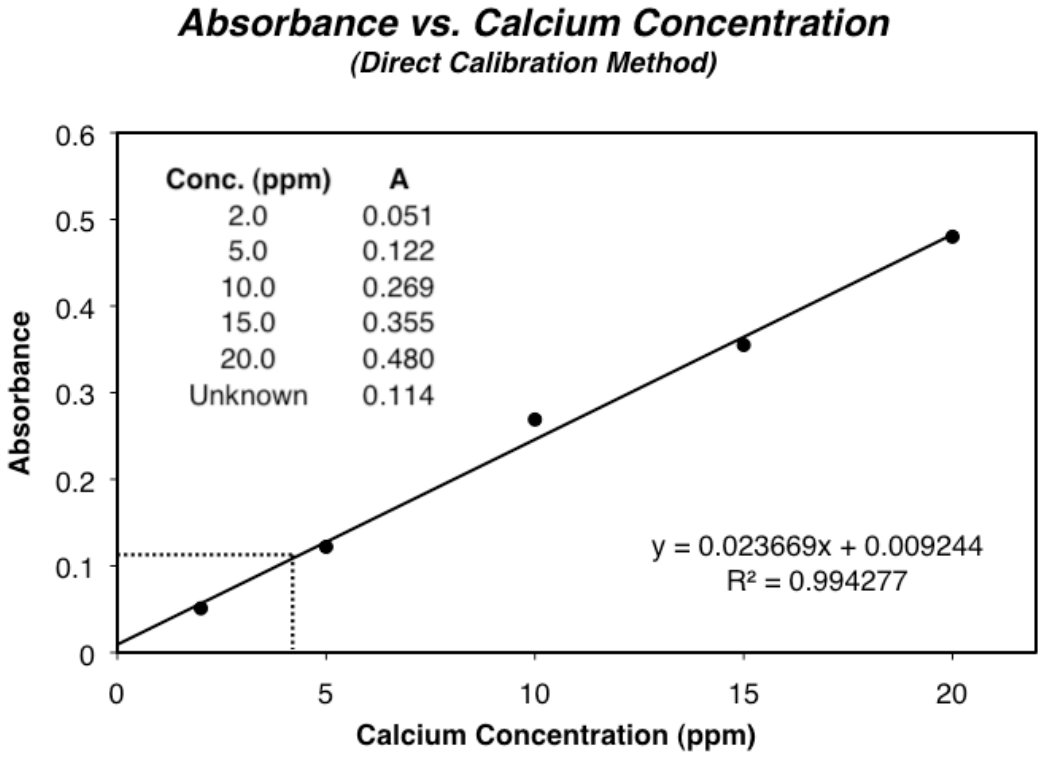
Calculating the Uncertainty of a Result Obtained by using a Linear Direct Calibration Curve

x is the independent variable (i.e., controlled, like concentrations of standards)

y is the dependent variable (i.e., measured, not controlled, such as absorbance)

Calibration: plot y on the ordinate (vert. axis) and x on the abscissa (horiz. axis).
Example below: absorbance on vertical, concentration on horizontal



Assume the instrument was adjusted to read zero for the blank. (Alternatively, define the dependent variable as {y – blank}, and plot {y – blank} versus x.)

Find the linear ($y = mx + b$) least-squares best fit to the calibration data.
Example: in Excel, use a linear "trendline" or the LINEST function

Once the linear calibration plot has been obtained, measure y for an unknown sample and find the x value that corresponds to it. If replicate samples have been prepared, find multiple values of x and calculate the mean.

What is the uncertainty (standard deviation) of x for the unknown sample?

$$s_x = \frac{s_y}{|m|} \cdot \sqrt{\frac{1}{k} + \frac{1}{n} + \frac{(y - \bar{y})^2}{m^2 \cdot \sum_{i=1}^{n} (x_i - \bar{x})^2}}$$

(for a direct calibration plot)

s_x = standard deviation of x, the analyte concentration in the unknown sample
s_y = standard deviation of y for the calibration points
m = slope of the least - squares best fit line
k = number of replicate measurements for the unknown sample
n = number of calibration standards used (i.e., number of points on the plot)
y = the measured y value for the unknown sample (or its mean y value, if $k > 1$)
$\bar{y}$ = mean value of y for the n calibration points
x_i = the values of x for the n individual calibration points
$\bar{x}$ = mean value of x for the n calibration points

In this case, the confidence interval = $\pm t \cdot s_x$ (It is independent of the number of replicate measurements k. Use n-2 degrees of freedom when looking up t.)

Complication: we can easily calculate values for all the variables in the equation except for s_y in the calibration plot. To find it, use Excel's **LINEST** function. The syntax is =LINEST(ycells,xcells,const,stats). If the logical variable "const" is set to TRUE or 1, then the data will be fitted with an equation containing a non-zero intercept (i.e., $y = mx + b$). If the variable "stats" is set to TRUE or 1, the standard deviations of y, m, and b will be calculated, along with the R^2 value.

Using the **LINEST** function: Somewhere on the spreadsheet that contains the x and y calibration data, select a 3-row X 2-column array of empty cells. This is where the results of the LINEST array function will appear. Then click inside the formula window, and type =LINEST(ycells,xcells,1,1), where "ycells" and "xcells" are selected by clicking and dragging over the cells that contain your y and x data values, respectively. After typing the closing parenthesis of the function, press CONTROL+SHIFT+ENTER or COMMAND+RETURN (depending on the type/version of computer and operating system). This is necessary because **LINEST** is an array function. The array of regression parameters and statistics will be calculated and placed in the six cells as below:

m	b
std.dev. of m	std.dev. of b
R^2	std.dev. of y(est)

An Example:

Values for calibration standards:	Ca^{2+} Conc. (x, ppm)	Absorbance (y)
	2.0	0.051
$n = 5$	5.0	0.122
$\bar{x}$ =10.4	10.0	0.269
$\bar{y}$ =0.2554	15.0	0.355
	20.0	0.480

From LINEST: $m = 0.023669$; $b = 0.0092439$; $s_y = 0.0151374$

Suppose a single unknown sample produces a measured absorbance (y) of 0.114. From $y = mx + b$ and the calibration data above, the unknown's Ca^{2+} concentration (x) must be 4.426 ppm (showing too many sig figs). Using the earlier equation with k=1, the standard deviation s_x for the sample's concentration is calculated to be 0.748 ppm. So, the result would be reported as 4.4 ± 0.7 ppm. To calculate the confidence interval, use n–2 degrees of freedom (= 3 in this case) to look up the t value (= 3.182 for 95% confidence). The 95% confidence interval is then $\pm 3.182 \cdot s_x = \pm 2.4$ ppm. These uncertainty values may seem rather large, but there are three good reasons for that. First, the calibration plot contains noticeable scatter. Second, the slope of the plot is somewhat small, so a small error in measuring absorbance (y) would produce a disproportionately large error in concentration (x). Finally, from the $(y - \bar{y})^2$ term in the equation, it is clear that a linear calibration curve produces the lowest uncertainty near its center. Unfortunately, the absorbance value for this unknown fell near the end of the plot.

Now suppose that 6 independent, replicate unknown samples are prepared instead of just one, and the mean of their absorbance values is 0.114. Using the same linear calibration curve, the mean concentration is obviously still 4.426 ppm. However, since k=6, the standard deviation s_x is reduced to 0.467 ppm and the result would be reported as 4.4 ± 0.5 ppm. The 95% confidence interval would then become $\pm 3.182 \cdot 0.467 = \pm 1.5$ ppm.