Tutorial Sheet 1 - Solutions

Sam Cooper and David Dye
MSE 101: Data Analysis

April 13, 2016

Problem 1.
Simplify the following expressions:

a) $5e^{\ln(A)} \quad \Rightarrow \quad 5A$

b) $7\ln(\mathrm{e}^{-A^2}) \quad \Rightarrow \quad -7A^2$

c) $3e^{\ln(5)+A} \quad \Rightarrow \quad 3(5+e^A)$

d) $\log_{10}(0.1 \times 10^A) \quad \Rightarrow \quad A-1$

e) $(\mathrm{e}^{(2\ln(A))})^2 \quad \Rightarrow \quad A^4$

f) $2^{2\log_2(A)} \quad \Rightarrow \quad A^2$

Problem 2.
Solve the following for x:

a) $7^x = 13 \quad \Rightarrow \quad x = \log(13)/\log(7)$

b) $3^x = \mathrm{e}^{x+2} \quad \Rightarrow \quad x = 2/(\ln(3)-1)$

c) $3\mathrm{e}^x = 7\mathrm{e}^{2x} \quad \Rightarrow \quad x = \log(3) - \log(7)$

d) $9 = 5 \times 2^x \quad \Rightarrow \quad x = \log_2(9/5)$

e) $7^{2x} - 9 = 0 \quad \Rightarrow \quad x = \log(3)/\log(7)$

f) $3\log_3(3^{3x}) = 3 \quad \Rightarrow \quad x = 1/3$

Problem 3.
Using the change of base rules, rewrite the following expressions in the natural logarithm (*i.e.* base e).

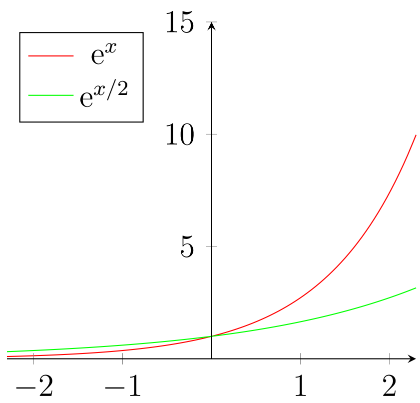
a) $\log_2(7) \quad \Rightarrow \quad (\ln(2))^{-1}\ln(7)$

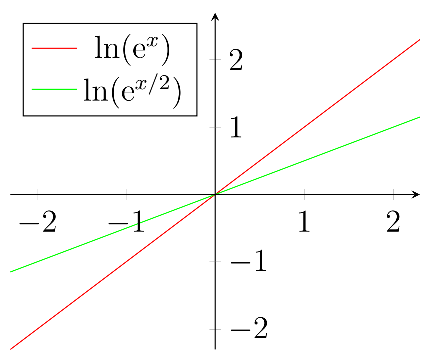
b) $\log_\pi(7) \quad \Rightarrow \quad (\ln(\pi))^{-1}\ln(7)$

c) $\log_{2e}(7) \quad \Rightarrow \quad (\ln(2\mathrm{e}))^{-1}\ln(7)$

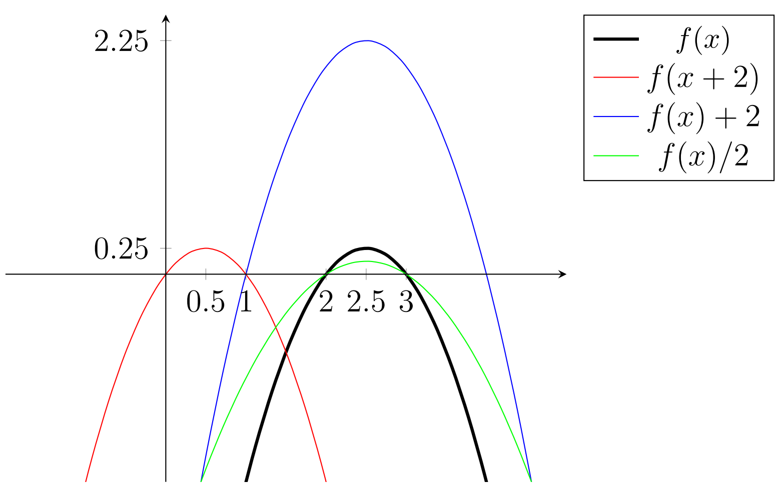
d) $\log_3(x) - \log_5(x) \quad \Rightarrow \quad ((\ln(3))^{-1} - (\ln(5))^{-1})\ln(x)$

Problem 4.
Plot the functions $y = \mathrm{e}^x$ and $y = \mathrm{e}^{x/2}$ on the same linear axes. Now take the natural log (*i.e.* ln) of each of these functions and plot them on a second graph.



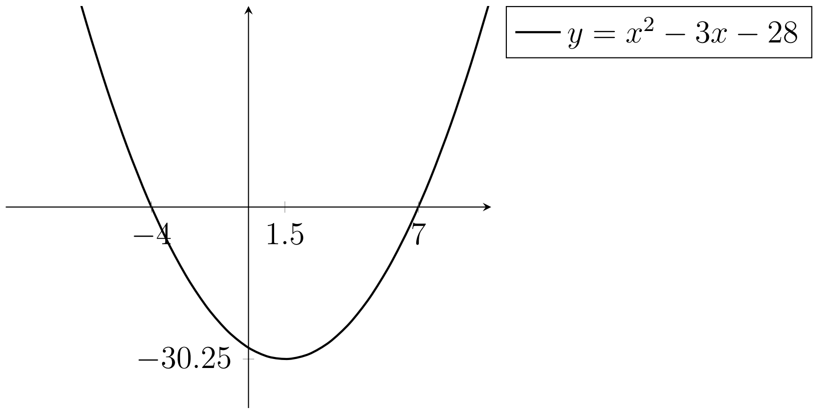


Problem 5.
One the same axes, sketch the function $f(x) = -x^2 + 5x - 6$ as well as the functions $f(x+2)$, $f(x) + 2$ and $f(x)/2$ (ideally in different colours!)

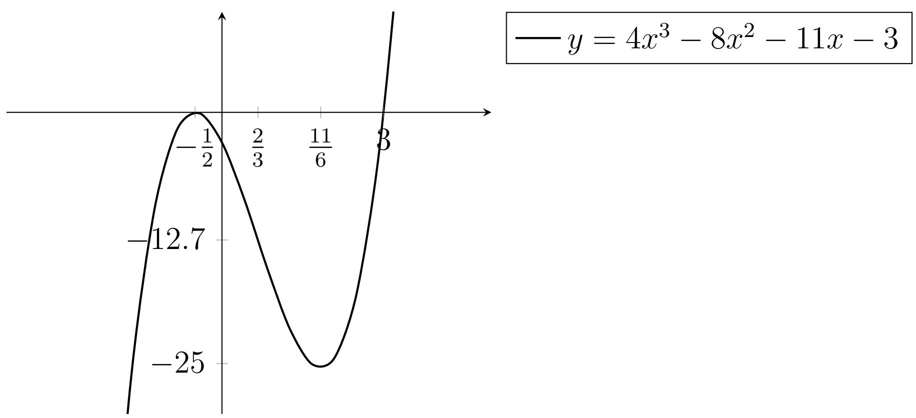


Problem 6. Sketch the following functions, labeling all roots, asymptotes, stationary points and inflection points. Also, state the domain and range of these functions:

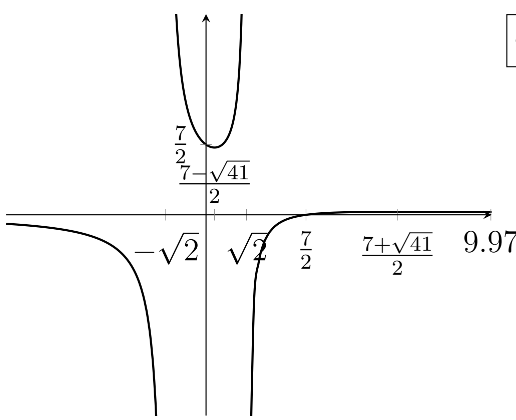
a) $y = x^2 - 3x - 28$



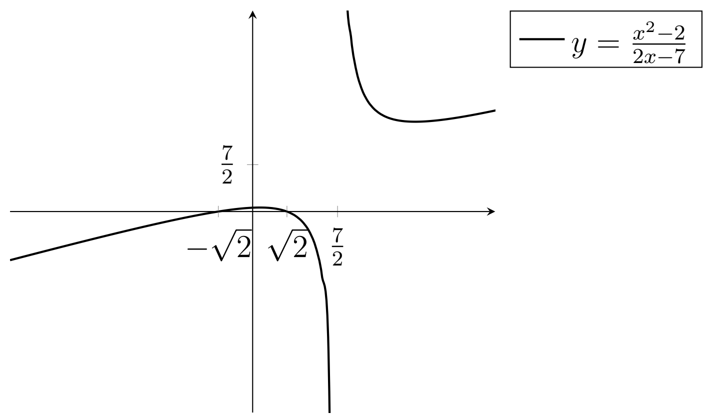
b) $y = 4x^3 - 8x^2 - 11x - 3$



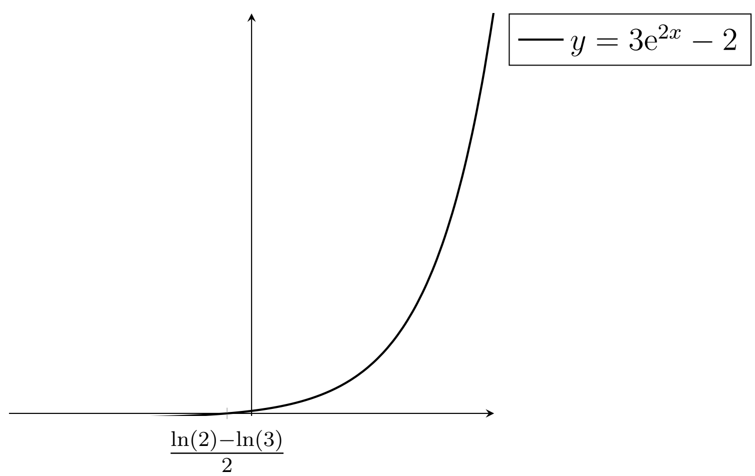
c) $y = \frac{2x-7}{x^2-2}$



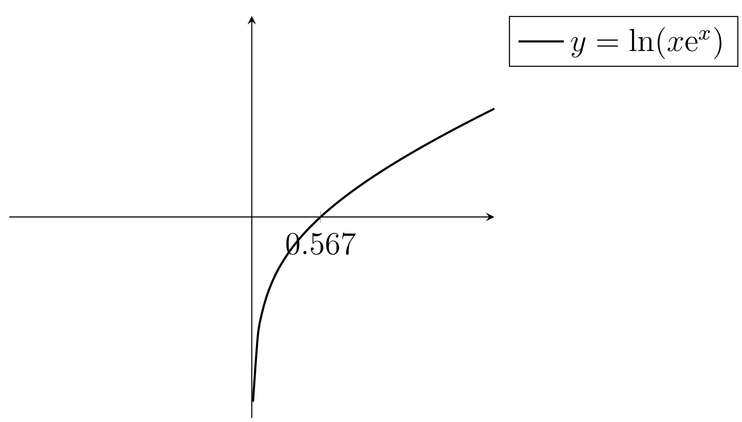
d) $y = \frac{x^2-2}{2x-7}$



e) $y = 3e^{2x} - 2$



f) $y = \ln(xe^x)$



Problem 7.

(a)

x	$f(x)$
1.00	-1.00
2.00	6.00
1.50	1.70
1.25	0.20
1.13	-0.43
1.19	-0.13
1.22	0.03
1.20	-0.05
1.21	-0.01
1.21	0.01

(b)

x	$f(x)$
0.0000	1.0000
1.0000	-6.0000
0.5000	-3.3787
0.2500	-1.6287
0.1250	-0.5617
0.0625	0.0852
0.0938	-0.2533
0.0781	-0.0887
0.0703	-0.0030
0.0664	0.0407
0.0684	0.0188
0.0693	0.0079
0.0698	0.0024
0.0701	-0.0003
0.0699	0.0010
0.0700	0.0004
0.0700	0.0000

x	$f(x)$
2.00	-4.00
3.00	58.00
2.50	10.15
2.25	0.85
2.13	-1.99
2.19	-0.69
2.22	0.05
2.20	-0.33
2.21	-0.14
2.21	-0.05

Problem 8.

(a) $f(x) = x^5 - x^2 - 4$ so $f'(x) = 5x^4 - 2x$ so $x_{n+1} = x_n - \frac{x^5 - x^2 - 4}{5x^4 - 2x}$

x_0	0.500
x_1	-5.636
x_2	-4.505
x_3	-3.596
x_4	-2.863
x_5	-2.264
x_6	-1.759
x_7	-1.293
x_8	-0.733
x_9	0.900
x_{10}	3.757
x_{11}	3.019
x_{12}	2.438
x_{13}	1.994
x_{14}	1.681
x_{15}	1.500
x_{16}	1.440
x_{17}	1.434
x_{18}	1.434

x_0	-1.200
x_1	-0.579
x_2	1.979
x_3	1.671
x_4	1.496
x_5	1.439
x_6	1.434
x_7	1.434

(b)

x_0	1	-1	10
x_1	2.341	-0.8553	9.333
x_2	1.861	-0.8177	8.871
x_3	1.573	-0.8156	8.656
x_4	1.452	-0.8156	8.615
x_5	1.430		8.613
x_6	1.430		8.613

Problem 9.

a) $2x - (4x^3)/3 + (4x^5)/15 + O(x^6) + ...$

b) $2 - x^2 + x^4/12 - x^6/360 + ...$

c) $-1 - 2x - x^2 - (4x^3)/3 - (2x^4)/3 + ...$

d) $-1/3 - x/9 - x^2/27 - x^3/81 + ...$

Problem 10.

a) $sinh(1) + (x-1)cosh(1) + 1/2(x-1)^2 sinh(1) + 1/6(x-1)^3 cosh(1) + ...$

b) $2 - 2(x-pi)^2 + 2/3(x-pi)^4 - 4/45(x-pi)^6 + ...$

c) $(e + e^{(}2e)) + (1 + 2e^{(}2e))(x-e) + 2e^{(}2e)(x-e)^2 + 4/3e^{(}2e)(x-e)^3 + 2/3e^{(}2e)(x-e)^4 + ...$

Problem 11.

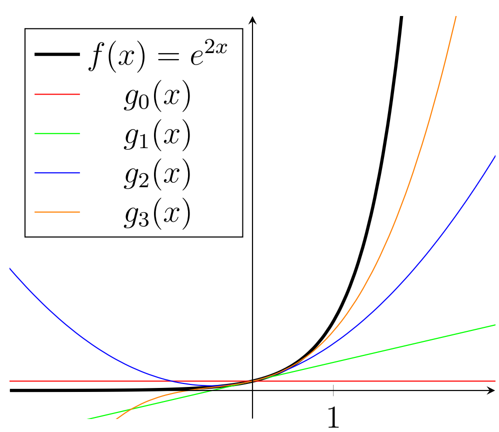


Figure 1: Taylor series expansions of the function $f(x) = \mathrm{e}^x$ at the point $x = 1$.

Problem 12.
$1 + x^2/2 - (7x^4)/24 + (31x^6)/720 + ...$

Problem 13.
$1 - x^2/2 + x^4/24 - x^6/720 + ...$

Problem 14.
Lots of 1/0... can't be done.

Problem 15.
$(x-1) - 1/2(x-1)^2 + 1/3(x-1)^3 - 1/4(x-1)^4 + ...$

Problem 16.

a) 0.9327

b) 0.0103

c) 0.5568

Problem 17.
Exact answer is $P = 0.691462$

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}}\left(x - \frac{x^3}{3} + \frac{x^5}{10} - \frac{x^7}{42} + \frac{x^9}{216} - ...\right)$$

g_0	1.15829%
g_1	-0.04370%
g_2	0.00138%
g_3	0.00003%

The x^3 term was required (four terms in total).

Error Analysis

Problem 18. Evaluate the uncertainty $\sigma_f(a, b, c, ...)$ given the following data. Comment, for each case, on effect of the magnitude of the different sources of error on your answer.

a) $f = a^2$ where, $a = 25 \pm 1$

$\partial f/\partial a = 2a$. Therefore $\sigma_f^2 = \sigma_a^2(2a)^2$ and so $\sigma_f = 2a\sigma_a = 2 \times 1 \times 25 = 50$, so $f = 630 \pm 50$. The fractional error in f is $50/625 = 8\%$, whereas the fractional error in a was only $1/25 = 4\%$ - a doubling.

b) $f = a - 2b$ where, $a = 100 \pm 3, b = 45 \pm 2$

$\partial f/\partial a = 1$ and $\partial f/\partial b = -2$. Therefore $\sigma_f^2 = \sigma_a^2 + 4\sigma_b^2 = 3^2 + 4 \times 2^2 = 25$ and so $f = 10 \pm 5$. The fractional error is 50%, which illustrates why subtracting similar-sized numbers is a bad idea.

c) $f = \frac{a}{b}(c^2 + \sqrt{d})$ where, $a = 0.100 \pm 0.003, b = 1.00 \pm 0.05, c = 50.0 \pm 0.5, d = 100 \pm 8$

$f = \frac{0.1}{1}(50^2 + \sqrt{100}) = 0.1(2500 + 10) = 251$. So d doesn't appear to be very important, for these values! Note that the fractional uncertainties in a, b, c and d are 3%, 5%, 1% and 8%, respectively.

Differentiating;
$\frac{\partial f}{\partial a} = \frac{1}{b}(c^2 + \sqrt{d}) = \frac{f}{a}$, $\frac{\partial f}{\partial b} = -\frac{a}{b^2}(c^2 + \sqrt{d}) = \frac{f}{b}$, $\frac{\partial f}{\partial c} = \frac{2ac}{b}$ and $\frac{\partial f}{\partial d} = -\frac{a}{2b\sqrt{d}}$
Thus
$\sigma_f^2 = \left(\frac{251\times0.003}{0.1}\right)^2 + \left(\frac{251\times0.05}{1.00}\right)^2 + \left(\frac{2\times0.1\times50\times0.5}{1.00}\right)^2 + \left(\frac{0.1\times8}{2\times1.0\times\sqrt{100}}\right)^2$
and so $\sigma_f^2 = 7.5^2 + 12.5^2 + 5^2 + 0.04^2$, which gives $\sigma_f = 15.4$. Therefore $f = 250 \pm 20$.

The largest contributor to the uncertainty in f is the 5% uncertainty in b (two-thirds of the total), followed by the 3% uncertainty in a (about a quarter), and the 1% uncertainty in c (one-tenth of the total). The 8% uncertainty in d is insignificant! This serves to demonstrate that the governing equation and values can interact in ways that mean that a quick calculation is desirable in order to establish where experimental effort should be expended to improve the overall uncertainty in the answer.

d) $f = 1 - \frac{1}{a}$ where, $a = 50 \pm 2$

$\partial f/\partial a = 1/a^2$. So $\sigma_f = \sigma_a/a^2 = 2/50^2 = 8 \times 10^{-4}$. Therefore $f = 0.98 \pm 0.0008 = 0.98$, and the uncertainty in f is negligible.

Problem 19. The volume $V = xyz$ of a rectangular block is measured by determining the lengths of its sides x, y and z. From the scatter of the measurements a standard error of 0.01% is assigned to each dimension. What is the standard error in V if (a) the scatter is due to errors in reading the instrument or (b) if it is due to temperature fluctuations?

a) If $V = xyz$, then $\partial V = \partial x = V/x$, so $\sigma_V^2 = \left(\frac{\sigma_x}{x^2}\right)^2 + \left(\frac{\sigma_y}{y^2}\right)^2 + \left(\frac{\sigma_z}{z^2}\right)^2$, so $\sigma_V = \sqrt{3}\ 0.01\% = 0.02\%$.

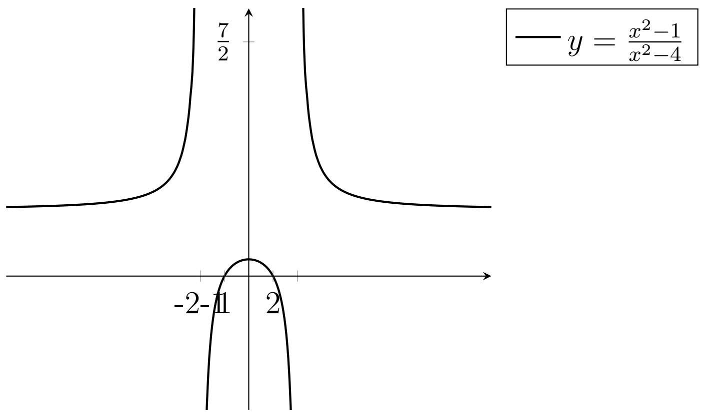
b) is trickier. The point is that in this case, the scatter in x, y and z will be perfectly correlated - when x goes up or down, so will y and z, *by exactly the same amounts.* So now, in effect $V = Cx^3$, where C is some constant given by the ratios of the lengths of y and z - $C = yz/x^2$. Then $\partial V/\partial x = 3Cx^2 = 3V/x$ and so $\sigma_V/V = 3(\sigma_x/x) = 0.03\%$.

Therefore we need to watch out for variables which *seem* uncorrelated but in fact are correlated! I (DD) actually had this happen to me once - we were using neutron diffraction to measure residual stress profiles in samples, by measuring the change in lattice parameter. Unfortunately, where we were, it was -30°C outside in winter, and every two weeks the main (truck) doors to the reactor hall would be opened in order for a gas delivery to be made. The delivery guys were in the habit of leaving the door open for an hour, during which the temperature of the equipment in the building could drop by a noticeable amount - about 5 K. This gave rise to a thermal strain of about 85×10^{-6} - equivalent to a stress of about 20 MPa. For our measurements with an uncertainty of about that magnitude, this was easily enough to generate a confusing data point.

Problems

Problem 1.

(a)



(b) Domain $= (-\infty, -2) \cup (-2, 2) \cup (2, \infty)$ & Range $= (-\infty, 0.25] \cup -(1, \infty)$

(c) $(-2, -0.34439] \cup [0.34439, 2)$ Found by see where value of NR give the first iteration as $x_1 < 2$

Problem 2.

$g_3(x) = (e^6 - log(3)) + (2e^6 - 1/3)(x-3) + (1/18 + 2e^6)(x-3)^2 + ((4e^6)/3 - 1/81)(x-3)^3 + ...$

$f(2.8) \approx 269.4$ and $g_3(2.8) \approx 269.0$ therefore approx is roughly 0.1% less than true $f(2) \approx 53.9$ and $g_3(2) = 65/162 - e^6/3 - log(3) \approx -135.2$ therefore approx is roughly 350% less than true.

This is because accuracy of approximation depends on the distance from the expansion point.

Problem 3.

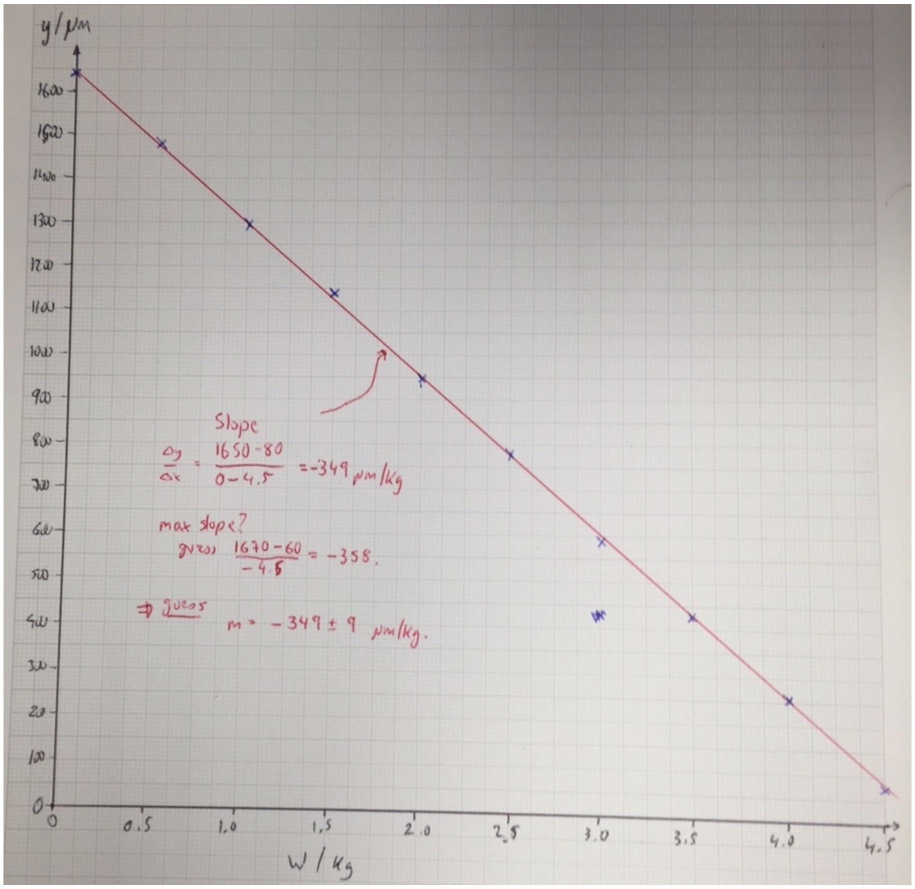
$\alpha = 3, \quad \beta = 10, \quad \gamma = 42$

$P(10.45 < X < 11.55)$

$$\begin{aligned}
P(x_a < X < x_b) &= P(X < x_b) - P(X < x_a) \\
&= \frac{1}{2}\left[\text{erf}\left(\frac{x_b - \mu}{\sigma\sqrt{2}}\right) - \text{erf}\left(\frac{x_a - \mu}{\sigma\sqrt{2}}\right)\right] \\
&= \frac{1}{2}\left[\text{erf}(-0.106066) - \text{erf}(-0.36534)\right] \\
&= \frac{1}{2}\left[-0.1192 - -0.3946\right] \\
&= 0.1377
\end{aligned}$$

0.1377×4000 hairs $= 550$ hairs.

Problem 4. This is the 3-point bend test problem, using a suspended weight and measuring the deflection of the centre of the bar. When we plot the graph we obtain



We try to plot a line with as many points above as below (which is quite tricky). Notice that the axes are labelled correctly. The gradient is then taken from the line, which gives a slope of $-349\,\mu\mathrm{m\,kg^{-1}}$. Estimating, we don't have a lot of wiggle-room in the line we pick - the biggest gradient I could justify was -358. Therefore a first-order estimate of the gradient would be $-349 \pm 9\,\mu\mathrm{m\,kg^{-1}}$

When we do the calculation (see Excel attachment) we obtain a gradient of $-349 \pm 4\,\mu\mathrm{m\,kg^{-1}}$ and an intercept of $1650 \pm 10\,\mu\mathrm{m}$.

This is a much lower estimate of the uncertainty than the one we obtained by eye, and pleasingly the value of the gradient obtained by eye was within the uncertainty of the computed fit. Therefore we conclude (from a sample of 1) that fitting gradients by hand is both accurate and expedient, much of the time.

Squires' suggests very strongly that we should always plot the data and examine it critically before we do any statistical fitting.

Problem 5. To solve this problem, go into Matlab and go through the following steps;

1. Import data x (times) and y (temperatures) from weld.csv, using the built in 'import' tab (you'll need to rename the file on the wordpress site - wordpress has a limitation of not allowing plain text filed).

2. Go to curve fitting and fit arbitrary function

```
20+n*1033*exp(-64.10*(c-0.0015*x))* ...
    besselk(0,64.10*sqrt((c-0.0015*x)^2+b^2))
```

   Here 1033, -64.10 and 0.0015 are the numerical values of the prefactors in the Equation for the temperature, given our material properties and weld data. n,b and c are our fitting parameters. Then, when we fit using the curve fitting app, we obtain estimates for n,b and c of 45%, 3.8 mm and 3.0 cm.

3. define

```
x=transpose(x)
y=transpose(y)
for j=1:10
  % first build a vector of possible efficiency values n
   n(j) = 0.2 + 0.05*j;
end
for k=1:20
  % and a vector of horizontal positions of the thermocouple b
   b(k)= k*0.00025;
end
S=zeros(10,20)    % initialise S
for i=1:101
  for j=1:10
    for k=1:20
      yfit(i,j,k)=20+n(j)*1033*exp(-64.10*(0.03-0.0015*x(i)))* ...
          besselk(0,64.10*sqrt((0.03-0.0015*x(i))^2+b(k)^2));
      % now build the residual for all time/temps i, for each j, k
      s(i,j,k)=(y(i)-yfit(i,j,k))^2;
      % and then we accumulate these into S, our chi-squared
      S(j,k)=S(j,k)+s(i,j,k);
    end
  end
end
```

4. look at the results

```
min(min(S))          % gives you the minimum chi-squared
surf(S)                  % makes a figure of S -  χ^2
S                          % write out all the elements of S
```

5. The point is that the thermal efficiency (n) and the y-position of the thermocouple (b) are quite strongly coupled together, so the minimum in b is quite broad and shallow. For example:

```
plot(x,yfit(:,5,15),x,yfit(:,6,18),x,yfit(:,3,8))
```

You can also do lots of fun things from here, but if you can plot how $\chi^2(n, b, c)$ and $y = f(x; n, b, c)$ vary, then you can understand the whole function. It might also be fun to graph the besselk function K_0 and thinng about how it multiplies with $\exp(-x)$, for example.