Comments on MATH7019 Assignment 1 Submissions

1. *General Remarks*

   (a) **Nonsense Answers:** The number one thing I don't want to see is students giving answers that are *clearly* nonsense. Here is a selection.

   i. Student 1 is trying to find an approximation to the chain on the left. He find the chain on the right and it doesn't seem to bother him that they look very different indeed.

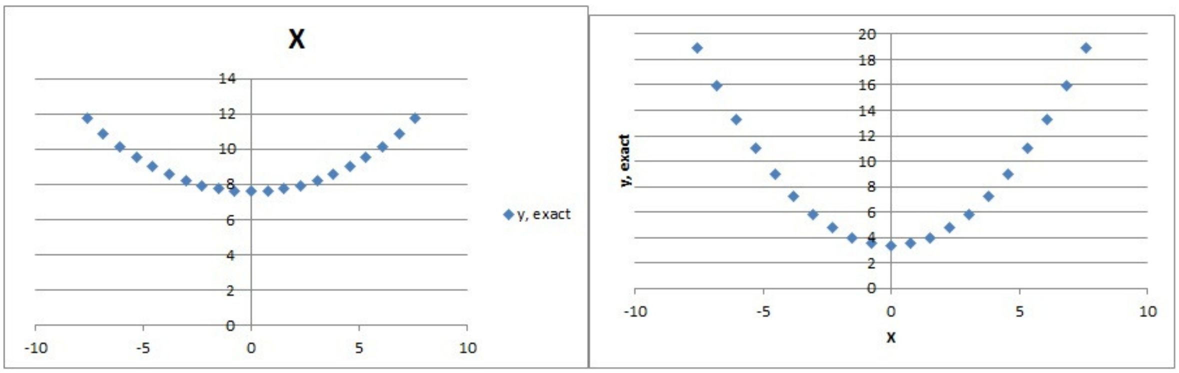

Figure 1: Note the second graph is very different to the first — at the boundaries the y-values are much different while at the minimum the y-values are also very different.

   ii. Student 2 is trying to extrapolate the maximum deflection of a beam of length 7 m. He comes up with an answer of 2.178 mm and doesn't realise this is clearly wrong as the beam of span 6 m already has a maximum deflection of 12.66 mm and so 2.178 mm is far too small.

   (b) **Errors in Assignments:** A lot of us noted that things were wrong... yes by all means do this in an exam if you don't have time to fix the problem. However ye had plenty of time to do this assignment and so plenty of time to figure out problems... if you couldn't figure it out you should have asked for help. For example, Student 7 was trying to extrapolate the maximum deflection on a 7 m beam and came up with an answer of 2.326 mm. In fairness, the student pointed out that this couldn't be correct as the maximum deflection for a 6 m beam was 5.000 mm and so the answer for the 7 m beam had to be greater than this. I gave the student half a mark for this but he should have tried to fix the problem.

   If you get a nonsense answer ask me what might be going wrong if you can't figure it out yourself.

2. *Mathematical Remarks*

   (a) **Percentage Error:** By convention, percentage error is given by:

$$\text{percentage error} = \frac{\text{error}}{\text{true value}}. \tag{1}$$

   For example, if the true value of the slope of the chain at a certain point is ≈ 0.5211, and the parabolic approximation is ≈ 0.5358, then the percentage error in this approximation is

$$\text{percentage error} = \frac{\Delta Q}{Q_0} = \frac{|0.5358 - 0.5211|}{0.5211} \approx 0.02821 = 2.821\%.$$

   (b) **Slope of a Curve:** The slope of (the tangent to) a curve is given by the derivative. This is first year material. The derivative gives you the slope of the tangent to the curve at any point x. Say, for example, that you had a (hanging-chain-approximating) curve

$$y = 0.141x^2 + 3.788.$$

   Then the slope at x is equal to

$$\frac{dy}{dx} = 0.141(2x) + 0 = 0.282x.$$

   We were interested in the slope of the chain at $x = a/2$, which for this chain above, was $a/2 = 1.9$. So $\frac{dy}{dx}$ gives the slope at any point x... and we are interested in $x = 1.9$:

$$\text{Ans: slope at } x = 1.9 = \left.\frac{dy}{dx}\right|_{x=1.9} = 0.282x|_{x=1.9} = 0.282(1.9) = 0.5358.$$

   (c) $y(x)$ **Notation:** When talking about a variable y depending on another variable x, for example

      i. the height of the chain y depends on the position x
      ii. the maximum deflection δ depends on the length of the beam L and
      iii. the passing percentage P depends on the grain size g
      iv. the biofuel consumption C depends on time t

   we say that y is a function of x and we write $y = f(x)$ or $y = y(x)$. This does *not* mean multiplication. For example,

      i. writing $y = y(x)$ just means y depends on x... for each value of x there is only one corresponding value of y.
      ii. $\delta(L) = a \cdot L^N$ does *not* mean $\delta \times L$ but that δ depends on L, δ is a function of L.
      iii. $P(g) = A \cdot e^{kg}$ means that P depends on g. Just dwelling on this, some students wanted to find the g such that $P = 50$ but substituted $g = 50$... no. It was $P = P(g) = 50$ so you had to solve:

$$50 = A \cdot e^{kg},$$

      where you had A and k from your log-linear least squares and you had to find the corresponding g.
      iv. $C(t) = m \cdot t + c$ does not mean "consumption times time"

3. *Curve-Fitting Remarks*

(a) $\sum c = Nc \neq (N-1)c$**:** So many students, for some unknown reason, took $\sum c = 17c$ if there were, for example, 18 data points. No. If there are N data points, $\sum c = Nc$. Always.

(b) **Fitting only to the Exponential Data:** I gave you 'S' shaped data:

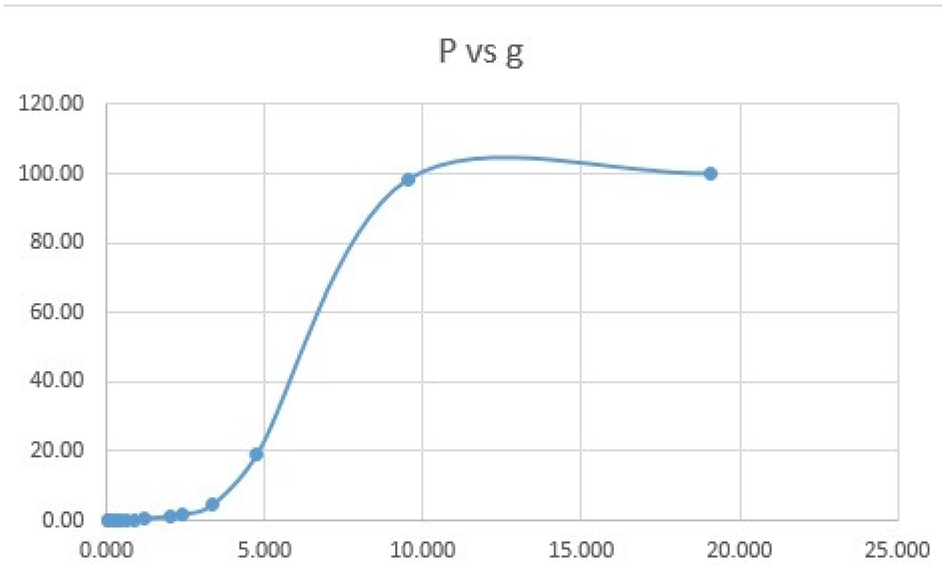

Figure 2: The entirety of the P vs g Data.

There is no point fitting an exponential to this because it looks nothing like an exponential.

If you do fit an exponential to it you get:

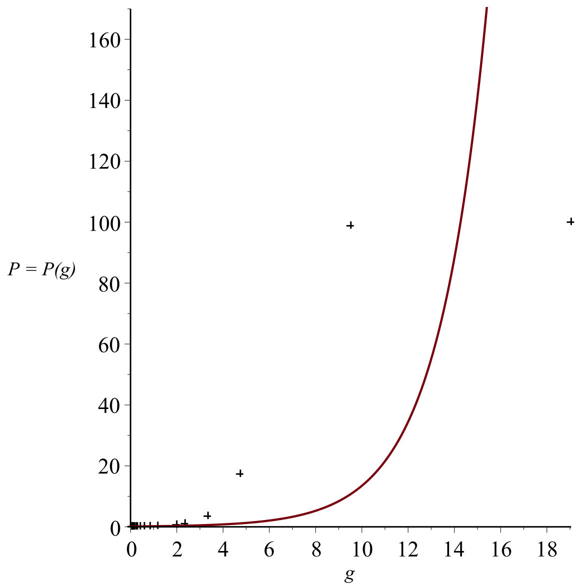

Figure 3: This isn't a good fit to the 'S' data above and predicts that half the sample has a grain size of 12.8 mm or less... which is nonsense because nearly 100% of the sample size has a grain size of about ten or less.

Instead either the $P = 100$ or both the $P = 100$ and $P \approx 100$ data points should have been thrown out and just the following, *relevant data* should have been looked at:

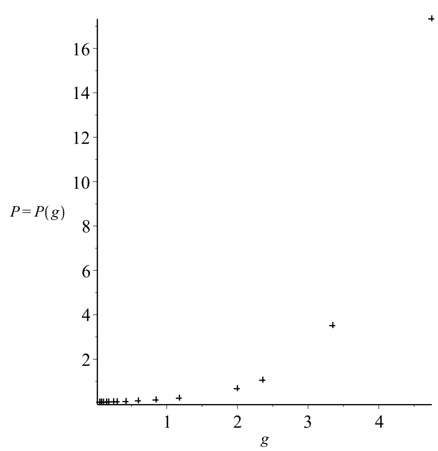

Figure 4: This data looks a lot like an exponential — we threw out the $P = 100$ and $P \approx 100$ data points.

Fitting an exponential to this yields, e.g. $g = 5.56$ mm for $P = 50$:

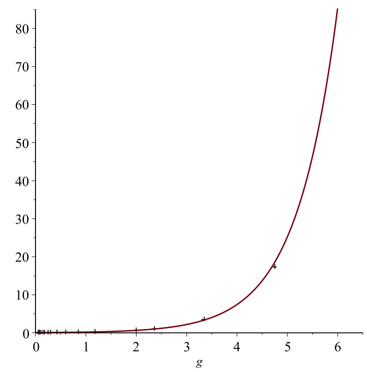

Figure 5: This is a much better fit to the data.

(c) **Answer the Question that was Asked:** The last question asked to comment on your predictions, based on the calculation of the Pearson correlation coefficient. Many people spoke about what the correlation coefficient being close to one said about the strength of the linear relationship between C and t... but this didn't answer the question that was asked. Other people didn't use the previous question... that is what "hence" refers to... Also people didn't state after Q. D5 whether the line formula in Excel matched the one they worked out by hand.