—
—
—
ben@redfrontdoor.org

July 30, 2013

State Examinations Commission
Cornamaddy
Athlone
Co. Westmeath

Dear Sir/Madam,

Re: Project Maths sample paper: 2012, Ordinary Level, Paper 2

Many thanks for your email (sent to me 3rd July), and for your thorough and detailed engagement with my enquiry. I appreciate the substantial amount of time this must have taken.

I realise that you will not want to get drawn into a long discussion on two parts of one question, but you do mention that the question will be reviewed, so I hope I can offer a response to your email.

* * *

Compromise between rigour and accessibility. I think it is preferable if the exam questions can completely avoid the need to strike a balance between rigour and accessibility, perhaps by testing areas which can be cut down to a reasonably-sized yet still sound core. Of course, the externally-set curriculum constrains the Commission, and for applied topics in particular this compromise might be inevitable.

I can also accept that students are to assume data-sets in questions are similar to those seen before, usually measured from experiments. One could say that this approach gives an 'implicit prior' on the samples in questions.

Testing theoretical vs 'real data' understanding. However, by referring to abstract 'sets of data', and by giving exact, round values for the various μ and σ, the question to me suggested it was testing a more theoretical understanding of the concepts. The use of the absolute word 'must' then pushed me further towards this interpretation, suggesting that an ideal student would have in mind a proof for why the set must have negative values.

Cues suggesting a more 'real world' interpretation could be: a mention of 'experiment' or 'measurement' (although the word 'data' does hint at this to an extent already); and μ and σ values which are not round integers.

Which set *must* contain some negative numbers? Yes, your

$$a_1 = a_2 = \cdots = a_{999} = 10 - \frac{20}{\sqrt{999}}; \qquad a_{1000} = 10000 - 999a_1$$

was the solution I was approximating (although, obeying the golden rule of letters pointing out errors, I rounded one of my values incorrectly). Your other solution,

$$a_1 = a_2 = \cdots = a_{800} = 0; \qquad a_{801} = a_{802} = \cdots = a_{1000} = 50,$$

is indeed a pleasingly clean one.

If the four sets are combined, the median is most likely to be a value in which set? If we accept the 'implicit prior' argument, then there is only one remaining aspect making me uneasy. The median of the combined set, having $\sum N = 1210$, is half-way between the 605th and 606th values. If set A is discrete, then these two values may well be the same, and their mid-point (the median) in A. However, if set A is continuous, the 605th and 606th values will almost surely differ, and the median will not be in the set. This problem could be avoided by making $\sum N$ odd.

Possible re-wording of question. The following re-wording illustrates the points above:

The size, mean and standard deviation of four sets of experimental data A, B, C and D are given in this table:

Include 'experimental'. Make μ and σ values suggestive of real data. Make size of D odd to ensure total number of values is odd.

	A	*B*	*C*	*D*
size (N)	1000	100	100	15
mean (μ)	9·3	113·1	988·6	84·1
standard deviation (σ)	21·3	29·8	18·2	12·5

[Lead-in unchanged; use 'measurements' rather than 'numbers' in parts (a) and (b); part (c) unchanged.]

(d) The set that is almost certain to contain some negative measurements is set ____.

*Use 'is almost certain to', not '**must**'; 'measurements' not 'numbers'.*

[Part (e) unchanged.]

* * *

I am not suggesting that you are under any obligation to satisfy arbitrary members of the public in your sample papers, but I hope you might consider these points in the review of this question.

Again, thank you for time in this discussion.

Yours faithfully,

Ben North