

**A response to the first examination for the NCEA Level 2 Achievement Standard:
2.12: 91267 Apply probability methods in solving problems
New Zealand Statistical Association Education Committee: July 2013**

The direction taken in this exam, and a recommendation:

The exam writers have done well in producing this paper. Probability is an area with many teaching and learning difficulties and the writers have made good progress with these. They have introduced new topics (comparison of theoretical and experimental distributions, risk, relative risk, etc). We gather that this has enabled a more positive set of results for students. We hope that this exam is part of an ongoing process in which the teaching of the area moves ever closer to the intent of the curriculum.

We are aware that the country has a shortage of resources for the new ideas in the curriculum, and we are keen to have this issue addressed. We hope that the Ministry's earlier Tier two work on Probability is readily available.

We strongly recommend that present and future exam writers help to move us to a new way of teaching and learning Probability. This would involve:

- Scenarios that resemble real and important situations that matter (such as food masses)
- Realistic distributions, with graphs and summary statistics you would find in practice
- Suggested theoretical distributions, fitted to the data (either by the students or by the examiner)
- Assessment of the assumptions that would justify the use of these distributions
- An expectation that experimental distributions are often lumpy, skewed, granular and non-normal; but full of interesting evidence and story
- Relative risk issues that matter (again, the writers did well here)
- Investigations into joint probabilities that both matter and challenge the mind (there's scope for creativity in providing examples here)
- Release from detailed calculations or graph-making that software can do
- A step away from purely mathematical mental exercises with Probability rules.

This involves a mindshift into new territory for us all. The recipe above is just a start. We'll need clear principles and examples that lead strongly. Here are some illustrations about questions on risk and relative risk:

- For a given situation, the question asks the student to find the relative risk(s) and then to write a newspaper headline that interprets this relative risk in everyday language without using the phrase 'relative risk' in the headline.
- In a given situation, relative risk is provided or is asked to be calculated and then asks what else should be considered to fully appreciate the implications of the relative risk. (i.e., the need to consider absolute risk in order to fully appreciate the practical implications of the relative risk.)

More detailed statements follow from three individual members of the NZSA Education Committee. The statements suggest that New Zealand research and debate are needed in the area of Probability teaching, learning and assessment.

Person 1

QUESTION ONE

Parts (a) to (d) are fairly straightforward.

- (e) This is marked as achieved but many students could find this question confusing. The wording is a little obscure and students could miss the point. They may work out an answer rather than just compare the claim with the answer to (b) or (d), which was what was intended. If the examiner wishes students to comment on a result rather than calculate an answer, this needs to be better signposted, either by changing the wording or by the placement of the question.

- (f) There are a variety of responses that could be made. The mark schedule lists:
- The data is not bell shaped (not symmetrically distributed about a mean).
 - It is bi-modal.
 - Mean and median not the same.
 - Most of the data is within two standard deviations.
 - The mean is not 41.

The schedule didn't mention that the mode(s) are not the same as the mean and median. Expecting all these responses for (e) seems excessive. There is no allowance for two correct responses giving an Achieved grade.

QUESTION TWO

- (a) (iii) This seems very easy for Excellence.
(b) (ii) The use of conditional probability could throw lots of students as it is only taught formally at Level 3.

QUESTION THREE

- (a) (iii) Answers to this question could gain a Merit grade fairly easily.

(a) (v) and (b) (ii) These two are almost the same: two Excellence questions, essentially assessing the same skill. However, in (b), students are required to fill in the table AND find probabilities, which is harder than (a) (i).

Person 2

QUESTION ONE

(e) Apparently the 'machine' has the same mean and standard deviation as the start of the question. This was not obvious to me from the text.

(f) The graph is drawn in half-grams on graph paper. I'd like to see the exam resemble likely practice more closely. The weighing machine probably gives masses to 0.1 of a gram (not 0.5). The simplest graph for the factory person with the clip-board would be a dotplot, marking each mass where it fell. The factory person may use software, in which case they can do a dotplot or histogram, with classes of 0.1. The more detail shown, the more we can interpret.

The bumps are good; they're big enough to not be there by chance. We have two equal modes it seems, which is a good discussion point. An awful lot of the bars seem to be below 40g. This is the most important thing for the markers: they could be in big trouble. This issue is not really highlighted in the Assessment Schedule.

At the start we have 'The weights are normally distributed ...' which gets more realistic in (f): 'When the machine is working properly'. I wonder if the phrasing from the start should have been more realistic: 'The workers set the machine so that mean and standard deviation should be (and we'll assume for now that the distribution is normal.)'

(d), (e) and (f) questions really do matter to the makers and eaters (if the data really is normal). But (a), (b), and (c) probably don't. They just test skills at slicing up the normal, for no apparent reason.

We could have lots more questions like (f), but with the data on a dotplot as practitioners would (i.e. at its most detailed level). Students can then be asked to draw the supposed or fitted normal onto it (as the software would do), and asked what the data is doing and why. Sketching the Normal is a bit tricky: centre and spread are easy but height is not. The exam could possibly have it drawn in: either expected (intended) or fitted versions. Data like this is mostly, in reality, collected over time and Excellence students should say: "Plot it over time".

Can we lead the world away from the Normal? The Curriculum says: 'such as Normal'. At least the uniform distribution should be familiar to students as well.

QUESTIONS TWO and THREE

The same applies here: the writers did good work at using questions that matter, which is hard. There is still mental gymnastics in the questions. In practice, does anyone have to back-calculate counts into a blank table? Probably not, but it may be a useful way of provoking and assessing thought.

Person 3

QUESTION ONE

There is widespread awareness that the Standard covers uniform as well as Normal distributions. One of the most widely used mock exams last year had a question based on the uniform distribution. Just because the first exam had a Normal distribution doesn't mean other distributions won't be covered in other years. For this reason, I am against putting the model onto the graph, as the student should select the model.

One of the risks when asking writers to create real world problems is that writers often go beyond their area of knowledge to create a problem which makes sense mathematically or statistically but is not in line with the real world. I have seen a handful of candy production distributions, designed so that the weight was at least that stated on the packaging, considerably more than 96% of the time. The introduction to Question One is a description of all Nutty Bars, which indicates that over 10% of bars are underweight. I suspect that Nutty Bars would be prosecuted under the Weights and Measures Act.

I agree that parts (a), (b), and (c) should have some reason for asking the question.

QUESTION THREE

I really like the structure of part (b), providing a table for students to fill in but not requiring them to use it. As I see it, one of the things holding back Probability education in New Zealand is the continued teaching of two way Venn Diagrams instead of two way tables as a first choice of tool for solving probability problems. Giving the table is a gentle nudge in the right direction at NCEA Level 2, whereas at Level 3 they will need to select their own method.

Links:

NCEA level 2 Probability (2012):
Achievement Standard

<http://www.nzqa.govt.nz/ncea/assessment/search.do?query=Probability&view=achievements&level=02>

at:

<http://www.nzqa.govt.nz/nqfdocs/ncea-resource/achievements/2012/as91267.doc>

Exam from 2012:

<http://www.nzqa.govt.nz/ncea/assessment/search.do?query=Mathematics&view=exams&level=02>,

at:

<http://www.nzqa.govt.nz/nqfdocs/ncea-resource/exams/2012/91267-exm-2012.pdf>