£s. 84 and 30 farmers; assuming that he plants x acres of potatoes and y acres of tomatoes, write down three inequalities that must be satisfied by x and y.

CHAPTER 13

Statistical Education in South African Schools

J.M. JURITZ UNIVERSITY OF CAPE TOWN, SOUTH AFRICA

Each of the four provinces of South Africa has its own education authority and syllabuses differ from province to province. All pupils who complete twelve years of schooling write matriculation examinations either set by the province, or the Joint Matriculation Board. Changes in the structure of the matriculation examination are being made at present, and it is now possible to take subjects in one of two grades, lower or higher. This has made it possible to include extra topics in the higher grade syllabus. Differentiation between the subjects starts at Standard 8 when pupils are about 15 years old.

Very little statistics is taught in schools at present and what is does not form part of the official syllabus. Some teachers do teach children about histograms and bar charts, and simple averages in Standard 5 – the highest class of the junior school. These are taught as an extension of arithmetic and, in more enterprising schools, the children collect data as class projects.

At present the mathematics syllabuses are being revised and there is considerable interest in including statistics and probability in the new syllabus. The South African Statistical Association has been consulted and has recommended that statistics and probability replace some of the geometry in the present syllabus. No officially recognised new syllabus has yet been produced. There is no agreement among statisticians as to exactly what should be included in the syllabus either.

The Mathematical Association of South Africa, whose members are mostly mathematics teachers, have a working party (The South African Mathematics Project), which is revising the entire mathematics syllabus from standards 5 to 10. They have achieved more than the official committees, and although their syllabuses have no official standing with the education departments, their proposals will be given serious consideration by them.

This proposed new syllabus includes some statistics and probability in each of the standards from 5 to 10 (for children from about 12 to 17 years). Also a statistics/probability option is included in the higher grade standard 10 syllabus. The content in the proposed syllabus is based largely on what is done in England at present. Basically it includes histograms, frequency polygons, mean, median, mode, interquartile ranges, standard deviation, regression, and standardisation of scores using the normal distribution. Methods of drawing samples, and discussion of use and abuse of statistics are also included. Mutually exclusive and independent events, addition and multiplication rules, the binomial distribution and uniform distributions is the main content of the probability section. Experimentation with dice etc. and collection of data by the pupils is encouraged.

As remarked earlier, there is disagreement amongst staisticians as to what should be included; the foregoing presents the author's personal opinion of the syllabus. A copy of the probability and statistics material in the South African Mathematics Project proposed syllabus is presented in the Appendix. It would appear that the content is sufficient, if not a little too much, bearing in mind that statistics and probability is only a small part of the total mathematics course. Inclusion of experimentation both in the probability and in the collection of data is a welcome aspect; it should be even more firmly emphasised.

In the hands of an experienced teacher such a course should work. However there is a danger of too much emphasis being placed on techniques – i.e. on drawing a histogram 'nicely' or calculating a standard deviation from grouped data. These topics are easily examinable and easily teachable, but they give rather a dull view of statistics.

There is need for more experimental work to replace technique, and some intuitive ideas of statistical inference– more emphasis on the abuse of statistics with real-life examples. It would make the topic less easy to examine but more useful.

Such changes would make the subject much more difficult to teach. However, there is good rapport between the teachers and the statisticians and further work on revision and training sessions is likely to take place at least in the Cape. One can feel optimistic that a satisfactory syllabus will be finalised and the enthusiasm of the teachers is most encouraging.

APPENDIX *South African Mathematics Project, Proposed New Mathematics Syllabuses: Statistics and Probability*

Standard 5, 7th year of schooling, Age group up to 12

11. *Statistics*

    Representation of statistical data by bar graphs.
    Arithmetic mean.
    (This section to be treated within the context of practical problems.)

Standard 6, 8th year of schooling, Age group up to 13

15. *Statistics*

    Methods of depicting data, including bar graphs and pie charts.
    Criticisms of methods used in newspapers etc. . . .
    Calculation of mean from simple data.
    The types of abuses in publications which pupils should look for at this stage are: distorted scales, lack of scales, suppression of origin, use of linear scales in pictogram where, in fact, areas are involved, and similar deliberate or fortuitous attempts to mislead.

16. *Probability*

    Intuitive introduction to experimental probability and theoretical probability.
    Expected frequencies.
    Probability of an event not occurring.
    Pupils already have an intuitive grasp of and enjoy work on probability at this stage of their development. What is started here will be developed in a spiral approach up to Matric level, where probability and the stochastic processes of statistics are intimately bound together.
    Some actual experimentation by the pupils is extremely important here, for example such experiments as are contained in Geoff Giles's 'Probability Kit No 1' (Oliver & Boyd, 1977).
    Explicit or implicit use of set notions is useful here.
    Example: *If we toss two coins a set of equally likely possible outcomes can be designated by*

    $$S = \{HH, HT, TH, TT\}$$

    What is the (theoretical) probability of the event 'two heads' occurring? If A is the set of 'favourable' outcomes, then

    $A = \{HH\}$ *and the probability is defined as*

    $$p(A) = \frac{n(A)}{n(S)} = \frac{1}{4}$$

    Also $p\,(\text{not } A) = 1 - p(A) = \frac{3}{4}$.

    The treatment envisaged is as given in Scottish Maths Group Book 3 (pp. 185–193).
    This section should occupy no more than 5 lessons.

Standard 7, 9th year of schooling, Age group up to 14

17. *Statistics*

Ungrouped data: short methods for calculating the mean; measures of spread: range and interquartile range (no standard deviation).
Frequency tables constructed with the aid of tally marks.
Mode and mean from a frequency table.
Frequency tables for grouped data (no histograms); modal class and calculation of mean.
(Note: As far as possible data should be collected by the pupils themselves. Calculators should not be allowed in this section.)
Frequency tables: The marks in a class test were called out in alphabetical order of the pupils' names as follows:

7 6 4 8 3 5 5 6 1 4 4 7 7 5 9 5 4 6 3
6 2 4 5 3 5 6 5 6 4 5 5 6 3 4 5 8 7 7

This information can be more clearly shown in a frequency table:

Mark		*Frequency*
1	/	*1*
2	/	*1*
3	////	*4*
4	~~////~~ //	*7*
5	~~////~~ ~~////~~	*10*
6	~~////~~ ///	*8*
7	~~////~~	*5*
8	//	*2*
9	/	*1*
10		*0*
		Total 38

or grouped into Class Intervals (different data):

Mark		*Frequency*
20–29	//	*2*
30–39	/	*1*
40–49	~~////~~	*5*
50–59	~~////~~ ~~////~~ ~~////~~ ~~////~~	*20*
60–69	~~////~~ /	*6*
70–79	////	*4*
80–89	//	*2*
		Total 40

The mode *(most frequent score) is 5.*
The mean *(average) is 5.15.*

Modal class *is 50–59.*

18. *Probability*

The addition and multiplication rules.
Suppose we throw a red and a white die. The set of all possible outcomes is shown:

		White die					
		1	*2*	*3*	*4*	*5*	*6*
Red die	1	*(1;1)*	*(1;2)*	*(1;3)*	*(1;4)*	*(1;5)*	*(1;6)*
	2	*(2;1)*	*(2;2)*	*(2;3)*	*(2;4)*	*(2;5)*	*(2;6)*
	3	*(3;1)*	*(3;2)*	*(3;3)*	*(3;4)*	*(3;5)*	*(3;6)*
	4	*(4;1)*	*(4;2)*	*(4;3)*	*(4;4)*	*(4;5)*	*(4;6)*
	5	*(5;1)*	*(5;2)*	*(5;3)*	*(5;4)*	*(5;5)*	*(5;6)*
	6	*(6;1)*	*(6;2)*	*(6;3)*	*(6;4)*	*(6;5)*	*(6;6)*

A total score of 5 can be obtained in 4 ways: (4; 1), (3; 2) etc. hence p(5) = 4/36. Also, p(10) = 3/36.
From table, p(6 or 10) = 7/36 (either a 5 or a 10 will do) and we note that p(5 or 10) = p(5) + p(10).
Hence if A and B are mutually exclusive outcomes, p(A or B) = p(A) + p(B).
Similarly, we can show that if A and B are independent of each other, then p(A and B) = p(A) × p(B), e.g. we expect (4; 1) to occur $\frac{1}{6}$ *of* $\frac{1}{6}$ *of the time.*

$$\therefore p(4;1) = p(4) \times p(1) = \frac{1}{6} \times \frac{1}{6} = \frac{1}{36}.$$

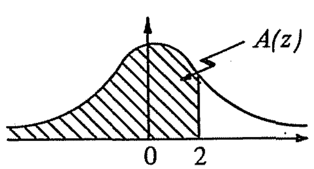
The treatment envisaged for this section is as in Scottish Mathematic Group, Book 3, pp. 196–201.

Standard 8, 10th year of schooling (Higher grade), Age group up to 15

14. *Statistics*

(New)
[Note: Calculators are not allowed in this section. As far as possible, the data treated should be collected by the pupils themselves.]
Standard techniques for drawing a random sample from a population.
Revision of measures of central tendency (mean, median, and mode) and of spread (range and interquartile range).
Histograms as graphs of frequency per stated class width. Frequency polygons. The area of a histogram and its use to give the fraction of a sample which falls between given limits, and to estimate the probability that a member of the parent population will fall between those limits.
Cumulative frequency table or curve, and simple deductions from it, including interquartile range.
Standard deviation as a better measure of spread. Calculation of means from grouped data (using an arbitrary origin and change of unit); standard deviation from ungrouped data (directly from the formula σ_x^2 = mean of $(x - \bar{x})^2$, with change of origin where convenient, but not change of unit).
Intuitive notion of the connection between relative frequency, probability, standard deviation, and the normal distribution. Use of simplified tables of the normal distribution to standardise scores.

A(z) for $z \in \{0; ½; 1; 1½; 2; 2½; 3; 4\}$



15. *Symmetry applied to graphs*
(Implicit in current Std 10 syllabus – now made *explicit* at Std 8 level.)
Consolidation and formalisation of Section 10 of the Std 7 syllabus: Pupils should know the eight symmetries of a square (ie reflection in the lines $x = 0$, $y = 0$, $y = \pm x$, and rotations about the origin through multiples of 90°) in the form

$$\text{Ref}_{y=x} : (a; b) \rightarrow (b; a) \quad \text{and so on.}$$

Some knowledge of the implications for graphs will be expected, eg (i) the symmetry of the parabola $y = ax^2$, (ii) the image of a graph of a linear function under reflection in the line $y = x$ is the graph of its inverse.
Brighter classes may wish to draw up group tables and/or explore the effects of translations, dilatations, and stretches parallel to the co-ordinate axes.

16. *Probability*
(New)
Revision and extension of Std 7 work. Tree diagrams. Use of Venn diagrams or Karnaugh maps in simple cases.

Standard 9 and 10, Higher grade, Age group up to 17

Explanatory Remarks

Time has not allowed us to give any serious consideration as yet to suitable Standard Grade syllabuses for Standards 8, 9 or 10. We do believe, however, that a syllabus for Standard Grade should not just be a watered-down subset of the Higher Grade syllabus, except possibly in Standard 8. What follows is for Higher Grade only. It is a translation (with minor modifications) of the revised draft prepared in January, 1977, as published in Afrikaans in *Spectrum* 15_4 (Dec., 1977).

The Compulsory Core consists of what we believe to be essential material for Higher Grade. The Options are so designed as to (i) accommodate much-needed material which is excluded at present, such as statistics, numerical methods, computing and calculus, and (ii) operate in such a way that a conservative teacher can follow a course that is not very different from the present Higher Grade syllabus.

5. *Probability and Statistics*
5.1 Revision of elementary probability: tree diagrams; use of Venn diagrams in simple cases.
5.2 Standard techniques for drawing a random sample from a population.
5.3 Revision of measures of central tendency (mean, median, mode) and spread (interquartile range and standard deviation).
Histograms and cumulative frequency ogives. Index numbers.
5.4 Scatter diagrams. Calculation of correlation coefficient (Pearson's) and equations of least squares regression lines, as an extension of the linear modelling done in 2.1. (No derivations of formulae; calculators with no more than one memory allowed.)

OPTIONS

6. *Calculus*
6.1 Limits (no formal definition).
6.2 Approximate determination of the area under a curve (using the trapezium rule, mid-ordinate rule etc.).
6.3 Differentation and integration of polynomials, including differentiation from first principles.
6.4 Applications, including areas, linear motion, maxima and minima and sketch graphs.

7. *Probability and Statistics*
7.1 Permutations and combinations; probabilities.
7.2 Idea of a discrete probability distribution. Calculation of mean and variance of a grouped frequency distribution and of a simple finite discrete probability distribution.
7.3 The uniform and binomial distribution.
7.4 The normal distribution as a tool applied in simple problems. Confidence intervals based on the normal distribution. (Pupils will not be required to carry out hypothesis tests as such, but should appreciate the meaning of the result of a test of significance.)