Using Multivariate Data as a Focus for Multiple Curriculum Perspectives

Jim Ridgway, James Nicholson and Sean McCusker
Durham University, UK

www.durham.ac.uk/smart.centre

Structure
- Professional statistics and school statistics
- Ambitions and Barriers
- MV data analysis for 13 year-olds
- Towards powerful weak methods

Statistics and School Statistics
- Statistics
  - data and problem driven
  - modelling targeted towards particular problem areas (biometrics, demography, and econometrics)
  - creating models to fit interesting problems NOT fitting interesting problems into standard models
    » from (da Silva 2006)

Statistics in School
- From Teaching statistics despite its applications (Ridgway et al 2007)
  - In England, ‘statistics’ = ‘technical mastery’
  - Uni or bivariate data ONLY
  - Analyses toy data via standard (1920s) models
  - Little use of computers
  - Quite unlike ‘real’ statistical practice
  - i.e. difficult, dull, and pretty useless

Barriers
- Access to technology
- Beliefs about what is ‘hard’
- (maths) teacher concerns about ‘messy data’

SMART Centre
- BIG ambition - to promote statistical literacy at school and in adults
- Engage everyone on the process of ‘reasoning with evidence’
  - To support the democratic process
  - For better political decision making
  - For better personal decision making
SMART Centre

- Creating novel interactive displays
  - Generic shells as freeware
  - Uploading interesting data sets
- Theory
  - What is worth knowing?
  - How does it develop?
  - How useful are 'weak' methods?
- Research
  - Studying user interactions
  - Studying user understandings
  - Empirical studies on difficulty
- Collaborating with data providers on better data displays
  - Raising adult literacy via Web 2.0 activities
- Curriculum development - Embedding MV data into curriculum materials

Our Ambitions

- Working with teachers with very limited mathematical skills (and little confidence)
  - use realistic data to aid understanding
  - engage students - to aid understanding and personal behaviour
  - (promoting statistical literacy across the curriculum)

Curriculum Development and Research

- Wellcome Trust project with the Geographical Association
  - Reasoning with BioMedical Evidence – Understanding Risk
- CCEA project
  - Data focussed cross-curricular materials for new curriculum for 11 – 14 year olds
- Becta project
  - What are the barriers to reasoning with multivariate data presented via ICT, and how can they be overcome?

Contexts and Data Sets

- Alcohol
- Poverty
- Drugs
- Obesity
- Pensions and savings
- Sexually transmitted diseases
- Tobacco
- Voting systems

Pupil Responses

- Yes, between 11 & 15 the level of drinking goes up and there is a big difference. At the ages of 12-13 the girls start to catch up with the guys.
- Boys drank more than the girls when they were younger but as they got older the girls drank much more than boys. There is a big difference between 11 & 15 year olds. Because girls are trying to act more grown up.

Informal Feedback

- Pupils
  - High levels of engagement observed
  - Prefer ‘discovery’ over ‘preaching’ about sex, drugs and alcohol
- Teachers
  - High levels of pupil engagement
  - Much more focussed discussions on ‘difficult’ topics
  - Much better written work than before BUT still big gaps between oral and written explanations
Contexts and Data Sets
- Alcohol
- Poverty
- Drugs
- Obesity
- Pensions and savings
- Sexually transmitted diseases
- Tobacco
- Voting systems

Conceptual and Research Issues
- What is the 'right list' for literacy?
- What key ideas are needed to understand different data sets?
- For any data set
  - What are the vices and virtues of applying quantitative methods?
  - What are the vices and virtues of applying qualitative methods?
- Defining and describing attainment
- How do ideas develop?
- Student (and teacher) misconceptions are?
  - diagnostic actions should be?

Towards Statistical Literacy
Describe and explore before you explain and model
- Critique the quality of the data
  - Check that the effect size is a bit bigger than the likely error of measurement
- Focus on effect size not significance level
- Identify variables that have the strongest effects
- Look at absolute levels — are they big enough to be worth worrying about?
- Look for non-linear relationships
  - Explore the effects over different values of each variable
- Look for changes over time
- Look for interactions, and think about ‘data surfaces’
- Think about possible confounding variables
- Disaggregate data, are the patterns the same?
- Is it worth quantifying?
  - Can you justify your strong assumptions?
- Look for the ‘dog that didn’t bark’
- Be cautious of claims about causality — especially in observational data

Data references
  - http://www.ic.nhs.uk/datasets
- Pension annuity rates available from:
  - http://www.fsa.gov.uk/tables/

Using Multivariate Data as a Focus for Multiple Curriculum Perspectives
Jim Ridgway, James Nicholson and Sean McCusker
Durham University, UK

www.durham.ac.uk/smart.centre