NUMBAS User Meeting 2022

Longer computational question setting and marking using NUMBAS

Chris Pearson & Nigel Penna

School of Engineering

Newcastle University

Christopher.pearson@ncl.ac.uk

Newcastle University

School of Engineering

Why did we use NUMBAS?

- NUMBAS used in geospatial and civil engineering for past 3 years
- Replaced paper assessments
- New assessments due to Covid-19
- Unique questions to reduce plagiarism
- Automated marking to reduce marking time
- Automated marking improves individual feedback
- Follow through marking / Adaptive marking easy to implement

Successes

- Having unique questions / data reduced the most blatant plagiarism (pure copying) but did not necessarily result in independent working on the assessment
- Significantly reduced marking load, although question design and NUMBAS implementation carries a large overhead
- More consistent marking
- Easy to adapt for remote exams
- NUMBAS is preferred by students over paper based assessment

Challenges

- Adaptive Marking / Follow through marks
  - Should students be awarded follow through marks?
  - Awards marks for correct method
  - Does not “penalise the same error twice”
  - Awards marks for answers that are fundamentally incorrect

- Question Design
  - Paper based assessment questions vs e-assessment questions
  - Hard to award ‘method marks’
  - Splitting up longer answer questions into multiple smaller steps, but how many?

Adaptive Marking

- Coursework assessment completed by 57 students in Oct 2021

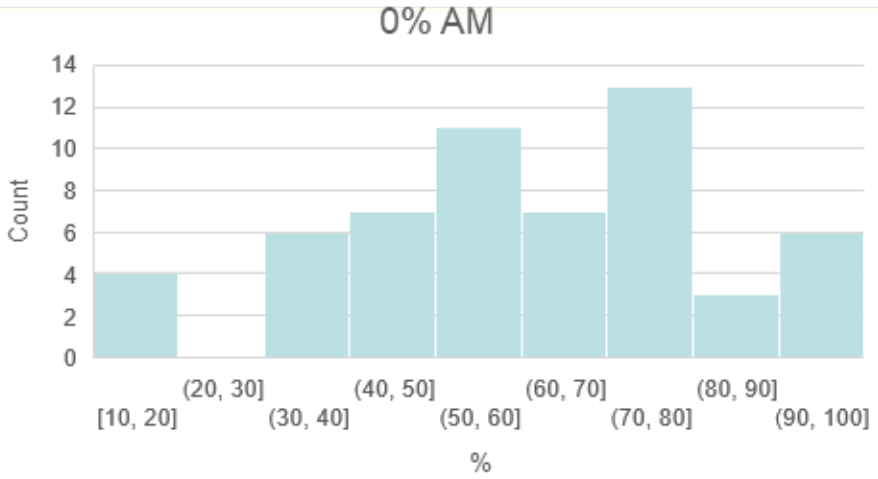
- Q1
  - Table layout
  - 30 answers
  - All follow on from each other

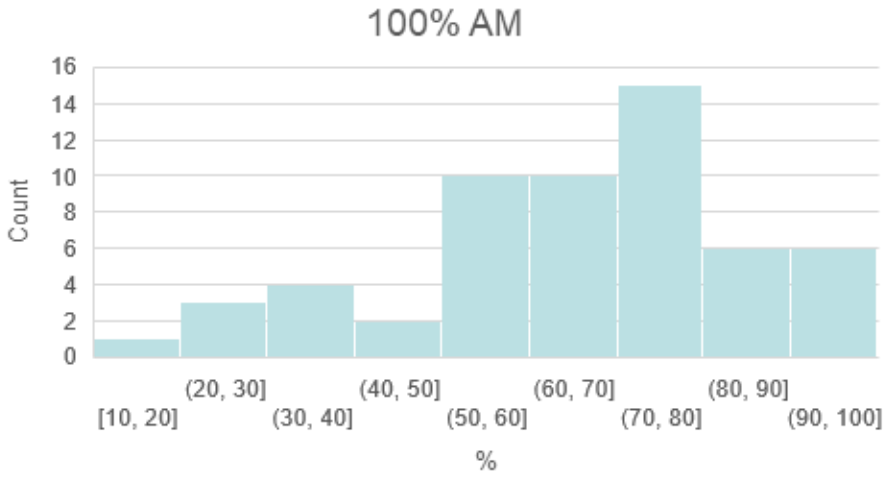
- Q2
  - 4 parts (3 follow on and one MCQ)
  - Traditional question statement with one box

Staff Pos	BS	IS	FS	Rise	Fall	Reduced Level	Remarks
BM wall	0.700					21.599	BM on wall: 21.599 m
MC1					0.677		Manhole cover 1
CP1	1.246		1.049				Change point 1
MC2					1.590		Manhole cover 2
MC3		1.582					Manhole cover 3
MC4		1.846					Manhole cover 4
MC5						20.354	Manhole cover 5
CP2			2.493		0.351		Change Point 2
MC6				0.131			Manhole cover 6
Bridge roof		-2.927					Inverted staff on roof of bridge
MC7						19.975	Manhole cover 7
MC8		1.824					Manhole cover 8
BM church				0.446			BM on church: 20.019 m
Sum	3.340						
Difference (m)							

Adaptive Marking (AM)

	0% AM	25% AM	50% AM	75% AM	100% AM
Mean	60.02	61.44	62.86	64.28	65.70
Median	62.5	64	65.5	67	68.75
Std Dev	21.00	20.59	20.22	19.87	19.57
p value	0.13	0.10	0.06	0.03	0.01





Shapiro-Wilk statistical test for normal distribution

P value >0.04 = normal distribution at 95% confidence interval

We tend to use 25% adaptive marking as standard now

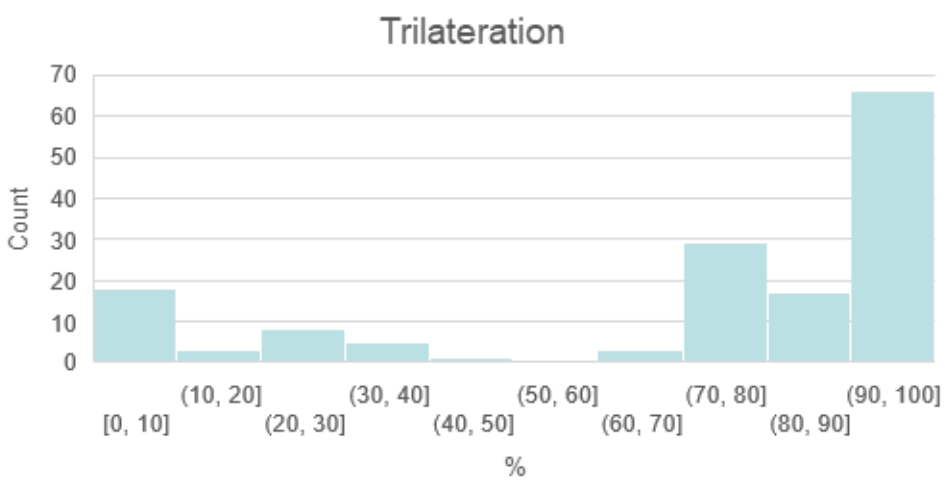
Question Design

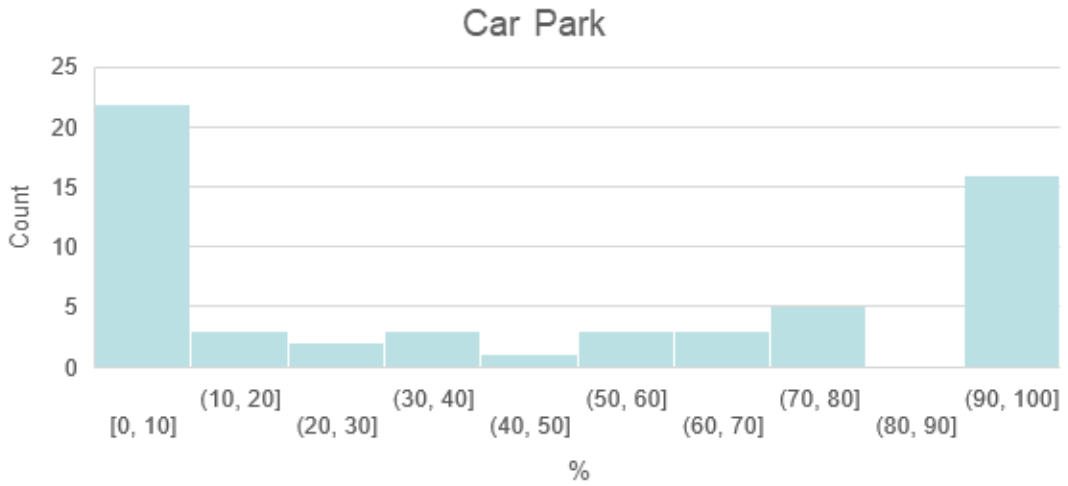
- Longer answer questions with multiple steps frequently required in engineering
- Minimal guidance required when hand marking due to method marking

- How to best design questions to allow students to pick up marks while still assessing if students have achieved learning outcome?
- Minimal guidance can result in ‘all or nothing’ style question
- Too much guidance can result in questions being too easy

Question Design

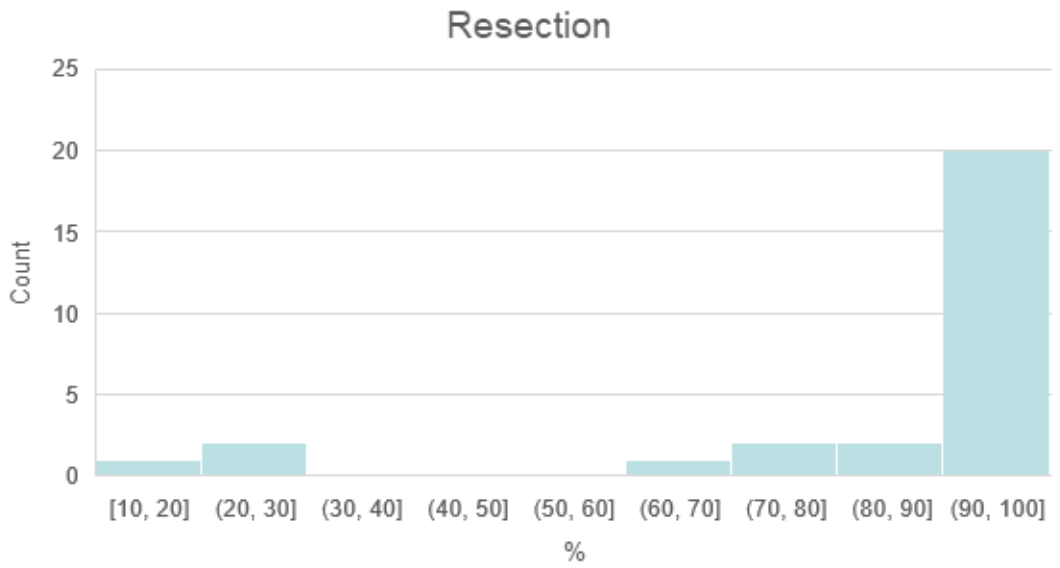
- Trilateration question given to 150 students
- Car Park question given to 58 students
- Main question asked for Easting and Northing of a point or points
- No intermediate steps given to students
- “Bi-modal” distribution





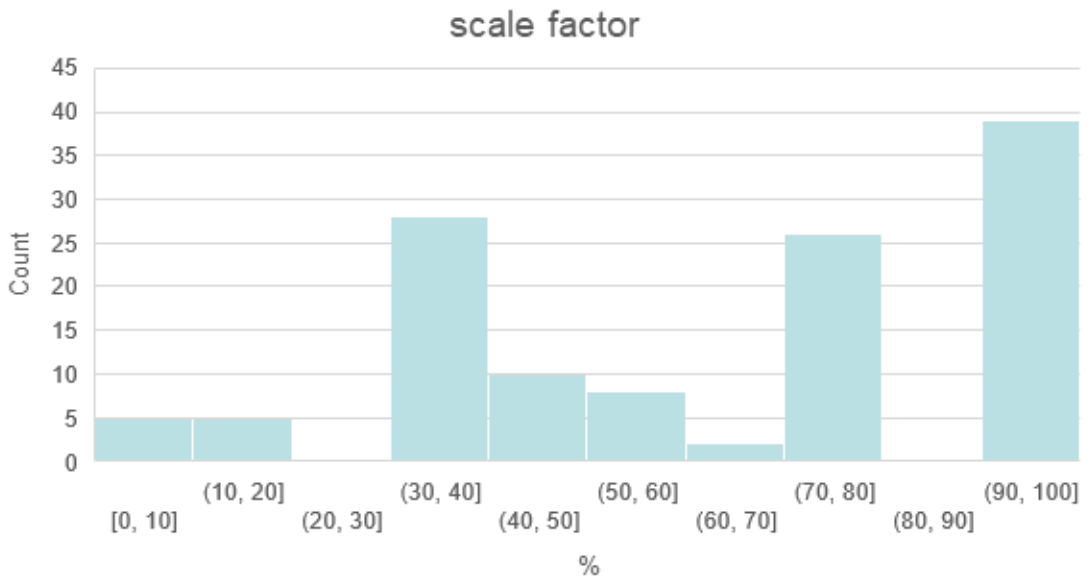
Question Design

- Resection question given to 28 students
- Final answers were again Easting and Northing of a Point
- 9 small steps required answers before getting to final answer
- Too much guidance on method resulted in skewed results
- Failed to test students understanding



Question Design

- Scale factors question given to 123 students
- Question split into 3 parts to find distance between two points
- Balance between ‘all or nothing’ one part questions and the ‘hand holding’ of questions with too many parts
- Better distribution of marks



Conclusions

- Full adaptive marking can lead to non normally distributed marks
- Is it appropriate to your subject area to award marks for incorrect answers?
- NUMBAS allows a percentage of follow through marks to be awarded

- Question design for e-assessment is very different to paper-based assessment for longer style questions
- Splitting longer questions into 3-4 parts appears to simulate students picking up some ‘method marks’ while not guiding them too much on appropriate methods