Scaling of exam results – policy summary

May 2015

This document is to provide a brief outline of how and why we scale exam marks in the School of Mathematical Sciences.

The need for scaling

Scaling is needed to ensure fairness across modules, with the intention that marks obtained in modules of a given level represents the same level of achievement. The marks obtained in mathematics exams tend to be both more spread out and more unpredictable than in other subjects, mainly because they tend to have lots of small questions. The examiner misjudging one or two of these questions can make the paper overall too difficult or too easy, and so we occasionally scale marks to rectify this.

Procedures

Once we have the exam marks available, we look at summary data – the average mark and the proportions of students obtaining each grade – to see whether a module looks as though it might be out of line. If we think that there may be a reason for scaling, then we consult the examiner and also look at any other information available, such as comments from the candidates or incidents occurring in the exam. (Note, however, that just because some candidates say an exam was hard we don’t necessarily agree with them!) We also examine scripts close to grade boundaries (especially the pass/fail borderline) to see whether the level of knowledge and understanding demonstrated by the candidate really corresponds to what we expect from a student obtaining a particular grade.

If we think that scaling is appropriate, then we try to judge the effect that the problems with the exam have caused, and scale appropriately. We may judge that something has had more of an effect on weaker candidates than on stronger ones, and so scale some marks more than others.

Note that we do not scale normatively, i.e. in order to obtain specific proportions of candidates with particular grades; we use statistics to indicate that scaling may be necessary, and then consider academic factors. We don’t like having a third of the candidates failing a particular exam, but if there are no academic reasons to suggest that those candidates should have passed, then we have to live with it.

The scaling formula

The scaling method we use is to adjust grade boundaries: we identify new grade boundaries that we think more accurately reflect achievement, and then transform the marks to map the new grade boundaries to the usual college boundaries. For example, we may decide that the pass/fail mark should be set at 38 rather than 40 because the scripts scoring 38 and higher deserve to pass, while those scoring 37 and below don’t; we would then multiply every mark between 0 and 38 by 40/38, and make corresponding adjustments to marks over 38.