To Round or to Truncate?
That is the Question

by Michael T. Kane, Ph.D.

As of September 2003, the Office of Management and Budget (OMB) was projecting a budget deficit of $455 billion for 2003. OMB could report projections to the nearest million dollars or thousand dollars, or even to the nearest penny, based on complex economic models, but it chooses to round its projections to the nearest billion dollars. Given that errors of tens of billions of dollars are not uncommon in such projections (in February, the projected deficit for 2003 was $304 billion), reporting the projection at a finer level (e.g., to the nearest million dollars) would not be more informative than reporting it to the nearest billion.

Test scores are much more precise than economic forecasts, but the need to avoid exaggerated claims to precision is important in all kinds of measurement. The computer programs used to generate test scores (e.g., the programs used to equate MBE scores, or to scale essay scores to the MBE) generate results with a very large number of digits after the decimal point; therefore their practical application requires the adoption of some rule for limiting the number of digits.

The most common way to eliminate extra digits in computations is to round the result. In rounding a score to an integer, the score goes to the next higher integer if the decimal part is greater than or equal to .5 and goes to the next lower integer if the decimal part is less than .5. If two scores, 134.421. . . and 134.792. . ., were rounded to integers, the first would become 134 and the second would become 135.[1] Similarly, if these two scores were rounded to the first decimal place, they would be 134.4 and 134.8.

Truncation is an alternative method for eliminating extra digits. In truncating a score, the extra digits are simply dropped. So, if 134.421. . . and 134.792. . . were truncated to integers, they would both become 134. If they were truncated to the first decimal place, they would become 134.4 and 134.7, respectively.

Truncated scores are always less than or equal to the corresponding rounded scores. If the first digit to be eliminated is less than a 5, truncating will yield the same result as rounding. If the first digit to be eliminated is greater than or equal to a 5, the rounded score will be slightly higher than the corresponding truncated score. So, from a candidate's point of view, rounding the test scores would be preferred to truncating the scores.

Whichever method is chosen, the issue is when and how to round or truncate scores. Given that we can round or truncate both MBE and essay scores to integers, to one decimal place, or to several decimal places, the number of possible rules is very large. To keep the discussion reasonably simple, I will focus on three relatively simple rules, any of which might be used by a jurisdiction:

- ***Rule 1***: For each examinee, round both the scaled MBE score and the scaled essay score to integers and then combine them to get the final score.[2]
- ***Rule 2***: Round both the MBE and the essay scores to one decimal place, combine the two scores, and round the result to a whole number to get the final score.
- ***Rule 3***: Round both the MBE and the essay scores to one decimal place, combine the two scores, and truncate the result to a whole number to get the final score.

Note that the rules involving rounding to two or more decimal places have essentially the same general properties as those with rounding to one decimal place.

Assuming that the passing score is an integer, a comparison of a final score with one or more decimal places to the passing score is equivalent to truncating the final score to an integer and then comparing the truncated score to the passing score. For example, assuming that the passing score is 270, a final score of 169.9 and a truncated final score of 169 are both less than 270. That is, if we report the MBE and essay scores to one or more decimal places, and compare their sum to the passing score, we get the same pass/fail decision that we would get if we truncated the final score to an integer and compared that to the passing score.

IMPORTANCE OF THE ROUNDING / TRUNCATING RULE

For most candidates, the rule used to decrease the number of digits does not make any practical difference, but for a few candidates (with final scores very close to the passing score), the choice of a rounding / truncating rule may affect their pass / fail status and thereby make a big difference. So the choice is an important policy decision.

Assume that a jurisdiction has a passing standard on the total examination (MBE plus essay) of 270, and that the essay part (which may include MEEs, state essays, or MPTs) is scaled to the MBE. The potential impact of the choice of rounding rule can be illustrated by considering three hypothetical examples.[3] For the first example, consider a candidate who gets an unrounded MBE scaled score of 134.41. . . and an unrounded essay scaled score of 135.34. . . . Using Rule 1, we round the two scores to the nearest integer and add them; as a result, we get 269 (134 + 135), and the candidate fails. Using Rule 2 or Rule 3, we would round the two scores to the first decimal place and add them, yielding 269.7 (134.4 + 135.3). Under Rule 2, we would round this result to 270 yielding a passing score, and under Rule 3, we would truncate 269.7 to 269, a failing score. In this case, the candidate would be better off if Rule 2 had been adopted, rather than Rule 1 or Rule 3.

Alternately, consider another candidate in the same jurisdiction, with scores of 134.62. . . and 134.72. . . . Under Rule 1, we get 270 (135 + 135), a passing score. Under Rule 2, we get 269.3 (134.6 + 134.7), which would round to 269, yielding a failing score. Under Rule 3, 269.3 is truncated to 269, which is also failing. In this case, the candidate would be better off if Rule 1 had been adopted and the MBE and essay scores were both rounded to integers before being added.

Finally, consider another candidate with scaled scores of 134.86. . . and 134.79. . . . Under Rule 1, we get 270 (135 + 135), and the candidate passes. Under Rule 2, we get 269.7 (134.9 + 134.8), which would round to 270, and the candidate passes. Under Rule 3, 269.7 would be truncated to 269, and the candidate would fail. In this case, the candidate passes under Rules 1 and 2, and fails under Rule 3.

Note that Rule 3 is more stringent than Rule 2 in the sense that all candidates who fail under Rule 2 must fail under Rule 3, but some candidates who pass under Rule 2 may fail under Rule 3. Using similar but slightly more complicated reasoning, it can be shown that Rule 3 is also more stringent than Rule 1.

The relative stringency of Rules 1 and 2 is more difficult to determine. As illustrated above, one candidate can pass under Rule 1 and fail under Rule 2, and another candidate can fail under Rule 1 and pass under Rule 2. However, making some reasonable assumptions about the distributions of scores, it is possible to show that Rules 1 and 2 are equally stringent on average.

All three of the rules stated above (as well as many others) are fair in the sense that all candidates are subject to the same criteria. Yet, as indicated above, the choice of rule can make a difference in the final pass/fail decision for some candidates. Ultimately, the choice among the different rules is a policy decision, not a mathematical one.

Significant digits in science

What criteria should be used in choosing a rule? One potential source of guidance in thinking about this issue is found in the guidelines for reporting scientific data. These guidelines are usually discussed under the labels “significant figures” or “significant digits.”

Scientists recognize that reporting results with more decimal places is not necessarily more accurate than reporting the results of the same measurement with fewer decimal places. In fact, reporting more decimal places can reduce the overall accuracy with which the results are interpreted by suggesting that they are more precise than they actually are.

In scientific analyses, scores are almost always rounded rather than truncated. The guidelines for reporting scientific results generally suggest that the last digit reported (i.e., the last “significant digit”) should be meaningful in the sense that we can have reasonable confidence in its accuracy. For intermediate calculations, additional digits are carried to avoid the accumulation of small rounding errors. Also, it is not unusual to include an extra digit or two if the reported score is to be used in a subsequent calculation. The significant-digits guidelines are intended to avoid exaggerated claims to precision in results that are publicly reported and used in making decisions.

Given the statistical properties of MBE and essay scores, the guidelines for significant figures would suggest that MBE and essay scores should be rounded to integers, or possibly to one decimal place.[4]

In situations like the typical scenario for bar examinations, in which MBE and essay scores are combined before a final decision is made, the rules for significant digits also require that the numbers to be added together have the same number of digits after the decimal point. So, if one of the two scores is rounded to the nearest integer, the other should also be rounded to an integer. If one score is reported with one decimal digit, the other should also be reported with one decimal digit.

The bottom line is that retaining two or more digits after the decimal point does not increase the accuracy of MBE, essay, or combined scores. On the other hand, retaining those digits can be justified if it serves some practical purpose and does not lead to the over-interpretation of insignificant differences.

Practical issues

There are some practical advantages to carrying scaled scores for the MBE to one or two decimal places. Probably the strongest of these practical advantages results from the fact that when MBE scaled scores are rounded to integers, as they are now reported by ACT, some pairs of consecutive raw scores will be converted to the same scaled score (e.g., raw scores of 113 and 114 might both be reported as scaled scores of 131) because of rounding. Candidates with the higher raw score (e.g., 114) may feel that they received no credit for their last question answered correctly. If the scaled scores were reported to at least one decimal place, a candidate with a higher raw score would always get a higher scaled score.

A second practical advantage of computing scaled scores to one decimal place is that the scaled scores would be less easily confused with raw scores. The MBE raw and scaled scores are both reported as integers on a scale from 0 to 200. If the scaled scores were reported to one decimal place, rather than as integers, they would be less likely to be confused with raw scores.

Reporting MBE and essay scores to one decimal place also has some disadvantages. In particular, as suggested by the reasoning behind the guidelines for reporting significant figures, it may focus attention on small differences (tenths of a point) that are of questionable significance. This is more of an issue with two or more decimal places than with one.

Recommendations

As noted above, there is no single best rule, and the choice of rule is a policy decision. However, there are some general suggestions that can help to simplify the choice.

First, whatever rule is adopted should be applied consistently to all candidates on any test date, and across test dates, until a decision is made to change the rule. Given that the same rule is applied to all candidates, all of the rules discussed here, plus many others, are fair.

Second, given the statistical properties of MBE and essay scaled scores, the guidelines for significant figures suggest that MBE and essay scaled scores should be reported as integers, but would allow for the use of one decimal place.

Third, reporting MBE scaled scores to one decimal place has the practical advantages of (1) eliminating cases in which two adjacent raw scores are converted to the same scaled score, and (2) making confusion between raw and scaled scores less likely.

Fourth, it is appropriate to employ the same number of digits for all scores that are being combined. So if the MBE scores are rounded to integers before being combined with the essay scores, the essay scores should also be rounded to integers. If the MBE scores are rounded to the first decimal place, the essay scores should also be rounded to the first decimal place.

Fifth, rules that involve truncation are generally more stringent than comparable rules involving rounding. This is not necessarily good or bad. It is a policy choice. Rules that compare final scores with one or more decimal places to the passing score are equivalent to truncation rules.

As noted earlier, there is no one best rule for reporting MBE and essay scores, and the choice of scoring rule involves some tradeoffs. In particular, the scientific guidelines for reporting data suggest that both MBE and essay scores should be rounded to integers but allow for the reporting of scores to one decimal place. On the other hand, practical considerations, particularly a desire to avoid having adjacent raw scores mapped into the same scaled score, support reporting MBE and essay scores to one decimal place.

The choice of scoring rule also depends on how the jurisdiction uses the different scores. To the extent that the MBE and essay scores are reported separately and used separately, the guidelines for significant figures could apply to them.[5] To the extent that it is only the total score that is reported, the significant figures guidelines would not apply to the MBE and essay scores separately, but only to the final score.

How might these scientific and practical guidelines apply to bar examination scores, and in particular to the three rules outlined earlier?

To the extent that MBE and essay scores are reported separately and used independently (e.g., are transferred from one jurisdiction to another or from one year to another), the guidelines for significant figures apply to each; these guidelines tend to suggest that Rule 1, in which both MBE and essay scores are rounded to integers and then combined to get the final score, would be appropriate. As noted above, this choice does involve the practical limitation that, in some cases, adjacent raw scores will be mapped into the same scaled score.

Rules 2 and 3, in which both MBE and essay scores are rounded to the first decimal place and then combined to yield a composite score with one decimal place, solve this problem; because the MBE scaled scores have an additional digit under these rules, adjacent raw scores are always assigned different scaled scores.

Under Rule 3, candidates can fail by slimmer margins than under Rule 2. Assuming a passing score of 270, candidates with final scores below 269.5 would fail under both rules, and candidates with composite scores above 270.0 would pass under both rules, but candidates with composite scores between 269.5 and 269.9 would pass under Rule 2 but fail under Rule 3. Therefore, Rule 2 has the practical advantage that no candidate fails by less than 0.6 points, which if rounded, would be one point. Under Rule 3, candidates can fail with a final score that is a tenth of a point below the passing score.

As indicated earlier, all three of these rules (as well as a number of other possible rules) are fair in the sense that they are all reasonable and can be applied consistently to all candidates. These three rules are all also relatively simple and straightforward; simplicity and transparency are highly desirable properties for high-stakes decision making. Rules 1 and 2 are equivalent in their stringency, and Rule 3 is slightly more stringent than the other two.

Given all of these considerations, Rule 2 seems to be a good compromise. It respects the significant-figures guidelines for reporting results, and, at the same time, avoids having adjacent raw scores being mapped into the same scaled score.

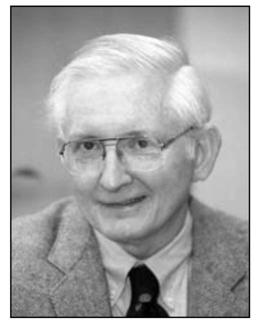
In applying a specific rule, a board is adopting a policy. As indicated above, although there are many rules that achieve the same general purpose and are fair in the sense that all decisions are made in the same way, different rules have different outcomes in some cases. There is no rule that will make everyone happy all of the time, and the best we can do is to choose a rule that is fair, that appears to be fair, and that generates reasonable outcomes.

ENDNOTES

1. As used here, the ellipsis marks, ". . ." are meant to indicate a very large number of additional decimal digits.
2. Weighting the scores can introduce additional complexity. For example, if a weighted average of integer MBE and essay scores (e.g., 131 and 134) is taken with weights of 0.6 and 0.4, respectively, we get,

   0.6 x MBE + 0.4 x essay = 78.6 + 53.6 = 132.2

   The combined score has an extra decimal place, and a decision will have to be made about whether to round or truncate this combined score (similar to the rounding / truncating decisions required for Rules 2 and 3).
3. In these and subsequent examples, I will assume that the MBE and essay scores are combined by adding them together with equal weights. If a different combining formula were used, the specific results would obviously change, but the same basic issues would arise.
4. Carrying an extra digit or perhaps even two extra digits for MBE and essay scores would not be inconsistent with the significant-digits guidelines for at least two reasons. First, the guidelines are basically rules of thumb, and allow for some flexibility as long as such flexibility does not lead to confusion. Second, the guidelines apply only to final results, and additional digits are routinely carried in the calculations leading to these final results. Whether the MBE and essay scores should be considered intermediate results or final results is not obvious in most cases. MBE and essay scores are often reported separately, and may be transferred separately from one jurisdiction to another or from one year to another, and thereby treated as independent results. However, the result that is of most importance is the final score used to make the bar admission decision, and therefore in most cases the MBE and essay scores can be considered intermediate results.
5. For technical reasons, it is generally not recommended that the MBE and essay tests be treated as separate hurdles. Requiring that candidates pass both the MBE and the essay tests tends to decrease the reliability of the decision-making process. From a psychometric point of view, it is better to combine the test scores and base the pass / fail decision on the combined score. However, even if the final combined score is used for decision making, the MBE and essay scores may be reported separately to candidates and other jurisdictions, and therefore be considered separately in the limited sense.



MICHAEL T. KANE, PH.D., is the Director of Research for the National Conference of Bar Examiners.