

## **2011 Pass Ratio Outliers**

By Daniel Roth, Examination Committee Chairperson

The 2011 sittings for the CAS upper level exams resulted in pass ratios that varied to a much greater extent than has generally been experienced. Given that the intention of the committee's pass mark setting process is to establish a pass mark which reflects a consistent standard of performance for the achievement of a passing score, the implication of wide variation in the pass ratios is either that the committee is widely varying its standards or that overall levels of exam performance by the candidate population varies more widely than had been previously assumed<sup>1</sup>.

Some critics have suggested that the variations in pass ratios are somehow connected to a CAS or Exam Committee goal of controlling the number of candidates qualifying for membership in the CAS. The Exam Committee wishes to make clear that such assertions, even if made only in jest, are not only unfounded, but also offensive, as they call into question the integrity and professionalism of the members of the committee, all of whom are credentialed members of the society bound by the performance standards enumerated in the CAS Code of Conduct. It is sincerely hoped that this communication will result in the elimination of such comments in the future, especially from those who aspire to join our society and enhance its reputation.

Given the high level of competence of actuarial exam candidates and extensive availability of study resources (e.g., employer-provided study time, clearly identifiable and readily available source material, commercially available study manuals and seminars, etc.), both the CAS and the candidate population find it surprising that there would be wide variation in the performance of the candidate population between sittings. Therefore, the Exam Committee believes it valuable to share the evidence it has seen which highlights the variation observed in recent sittings. If such evidence were absent, the implication would be that the pass mark setting process might not be as effective as desired at establishing consistent passing standards, and the CAS should investigate adjustments to the process in order to achieve its objectives.

The Exam Committee considers the following exams' pass ratios to reflect "outliers" relative to historical results: Spring Exams 5, 5A, 5B and 7; and Fall Exams 6C, 6US and 8. The discussion below reviews the distribution of raw scores by exam against similar distributions from comparable exams.

### **Distribution of Aggregate Scores**

Attachment 1 provides the distribution of raw total exam scores for the referenced exams, along with distributions of relevant comparable exams (2010 Exams 7C, 7US and 8 are provided for comparison to 2011 Exams 6C, 6US and 9, respectively). The committee notes the following indications from this data:

---

<sup>1</sup> This is not to say that current variation in candidate performance is different than in the past. The historical pass mark setting process was based on an assumption that candidate performance was relatively constant, so the relatively lower level of variation in historical pass ratios gives no indication of historical variance in performance.

Exams 5, 5A, 5B

Given that these exams tested the same questions under the same time allotments, the wide disparity in scores provides the strongest indication that there exists an underlying wide dispersion of candidate performance. Granted part of this performance difference is due to the disparity in preparation requirements (partial exam takers had the same amount of calendar days to prepare half the material), the disparity was still wider than the committee expected.

Within the Exam 5 population, it is noted that Section B performance is worse than Section A, which could indicate either a disparity in performance (candidates were stronger on the A material) or a disparity in time allotment between the sections. The committee was aware of this data during its deliberations and considered it in the determination of the final pass marks for each of the exams.

The committee also noted that the overall pass ratio for the combined set of exams (5A, 5B, and 5) was 44% (455 passers / 1042 candidates), which further suggests that even when an exam pass ratio is near the “expected” 40%, there is an underlying wide distribution in candidate population performance.

Exam 7

With the mode of the population scoring only about 50%, it should not be surprising that the pass ratio was very far from expected.

Exams 8, 9

The committee notes that the proportion of candidates scoring 65% or fewer of the available points was 9 percentage points (pp) higher in 2011, which would indicate, prior to any discussion of the pass mark, that the pass ratio would likely be lower than the previous year’s 37% by about this amount, which would land it at less than 30% regardless. As communicated previously (see [How CAS Exam Pass Marks Are Set](#)), the final determination of the pass mark is dependent on the content of the questions of the individual exam as it relates to the standards defined by the committee for a minimally qualified candidate (MQC). With respect to these standards, the 2011 exam was determined to be relatively easier than 2010, and so the pass mark was higher in an effort to achieve a common standard, thus resulting in a disparity in pass ratios between 2010 and 2011 greater than these 9 pp.

Exams 6US, 7US

The comparison of performance between the 2010 and 2011 sittings shows nearly 20 pp more candidates scoring at or below 65% of the available points. The variation in pass ratios is larger than this amount, but it is clear that even if the committee were to adjust its pass mark arbitrarily to pass more candidates, it would not find a reasonably strong standard that would result in a pass ratio above 30%. Also noteworthy is that even though there were almost 400 more candidates sitting in 2011, the number of individuals scoring more than 75% of the available points dropped on a nominal basis relative to 2010.

Exams 6C, 7C

This grouping shows the distribution of candidate scores to be generally comparable, but even here the committee notes that the portion of candidates scoring 45% or less doubled in 2011, with the 11 pp difference essentially equal to the difference in pass ratios between the two years.

Additional Comments on Raw Score Comparisons

The committee notes that length of exam will have an influence on these results. Clearly variations in exam length make it difficult to compare the results between groupings. However, the committee is not aware of significant differences in length within these groupings that would suggest the items noted above are not properly comparable. For example, the 2010 Exam 7US/2011 Exam 6US group consists of similar 100-point, 4-hour exams.

Based on the data here, the committee believes there is sufficient evidence to support the assumption that candidate population performance can vary significantly between sittings. The committee will continue to seek improvements in applying its passing standards to each exam to help ensure consistent, quality determination of the right pass mark, but the pass/fail result for each and every candidate will remain a function solely of that candidate's performance against the standard, whether the resulting pass ratio is 100% or 0%.

**Review of Individual Item Scores**

Another assumption in the grading process is the reasonableness of the committee's MQC standard relative to each of the questions on the exam. As a general rule, if the candidate population generally scores where the committee would expect, this assumption is not questioned in great depth, though it always goes through at least a two-step validation process (pre- and post-exam assessment). However, for any exam item where candidate performance is materially different than expectation, the General Officers and Committee Chair discuss in detail with the Part Chair and the graders: the item, the *a priori* grading rubric, the Pass Mark panel's *a priori* MQC standard, the grader's adjustments to the grading rubric through the grading process, and the grader's opinion regarding the MQC standard in light of their final grading rubric. The purpose of the discussion is to determine the extent to which the committee may have faltered in any of: question design, grading, or assessment of difficulty.

Unfortunately, due to the multi-layered nature of this process, and the considerable amount of judgment that is applied, it is difficult to produce data which provides a level of clarity of the reasonability of this process comparable to the aggregate score review provided above.

This difficulty is illustrated by one of the items the committee looked at very closely during the recent exam sitting, Exam 8, Question 12. This item was viewed by the Pass Mark Panel, and verified by the graders, to be a clear, relatively easy question with a liberal grading rubric. As such, the MQC standard for the item was set quite high, 1.25 out of 1.50 points. The candidate population, including the candidates who scored very high on the exam overall, only averaged about 1.00 point. Would 1.00 have been a better MQC standard than

1.25? It certainly is reasonable to suggest such. However, the final determination is based on the reasonableness of expectation against the passing standard. In this case, the decision was that 1.25 is the right value.

In other cases on the same exam, for example, Question 18, candidate population score was high, with a mean of 2.65 on this 3.00 point question, yet the committee concluded that the fact that this was significantly higher than the MQC standard of 2.25 did not indicate that the MQC standard underestimates either the level of difficulty or the required level of performance.

We wish we could provide a level of detail that would convince our strongest critics of the reasonableness of the approach and outcome in determining the individual and aggregate pass marks, but a full report of the committee's MQC deliberations is beyond the scope of our external communication. It is hoped that this small insight on these couple of items is at least somewhat helpful toward this goal.

### Looking Ahead

The comments below are intended to enhance guidance previously provided by the committee in *Future Fellows* and other venues to help improve candidates' success in demonstrating mastery of the Learning Objectives on their next exams.

1. Review the Syllabus Carefully. The tests are designed to assess candidate's proficiency with the Learning Objectives. Be certain to understand what each Learning Objective is and to study the aspects of the source material that support proficiency in that direction. The Knowledge Statements provide guidance as to the context in which the exam will attempt to assess proficiency of the related Learning Objective. If external study manuals are used, candidates should ensure that their study is structured around the current year's *Syllabus* in case the manual is not. With respect specifically to Exams 6US and 6C, the committee strongly encourages candidates to ensure that their study leads them to a high level of proficiency in the *Syllabus* section titled "Professional Responsibilities of the Actuary in Financial Reporting" on each exam<sup>2</sup>, though, of course, not to the exclusion of other sections.
2. Align Answers to the Questions Asked. When taking the exam, read carefully the direction given in each item, with particular attention to any adverbs used, and the number of points assigned (see [The Importance of Adverbs](#) for more detail regarding adverb usage on exams). Use the verbs, adverbs and each item's point assignment to guide you to the level of detail required to achieve full credit on each exam item. Provide responses that are directly related to those verbs at the level of detail indicated by the adverbs.

Candidates should also be aware of the coming increase in the Bloom levels of questions (see [The Evolution of CAS Exam Questions](#) and [What is Bloom's Taxonomy?](#)). The dearth of historical questions at higher Bloom levels may make reliance on external study guides and

---

<sup>2</sup> These *Syllabus* sections are highlighted because of the consistent shortfall each year in candidate performance relative to the MQC standards on questions drawn from them.

past exam questions less effective than ground-up study of the Learning Objectives. Candidates with sufficient proficiency in the Learning Objectives should have what is needed to respond to higher Bloom level questions, but preparation may be helped by developing one's own questions that reach to these higher Bloom levels.

For example, consider a Learning Objective which says "Calculate...", which in the past had been tested by a question which says "Calculate..." In the future, an exam may have a question which says "Evaluate the reasonableness of the calculation in the following situation..." If a candidate is sufficiently proficient in the understanding of the calculation, s/he should be able to answer the new question as effectively as the old.

The background and guidance above is provided in the hope that future concerns with variances in pass ratios will be focused on why they're suddenly 20+pp above historical levels.

Distributions of Raw Exam Scores  
Selected Exams 2010-2011

		<= 65%	0%-45%	45%-55%	55%-65%	65%-75%	75%-80%	> 80%	TotCan	Num Pass
2011 Exam 5 (section A)	Number		94	74	109	177	89	110	653	
	Percent	42%	14%	11%	17%	27%	14%	17%		
2011 Exam 5A	Number		3	2	16	45	34	67	167	106
	Percent	13%	2%	1%	10%	27%	20%	40%		
2011 Exam 5 (section B)	Number		164	99	122	147	68	53	653	
	Percent	59%	25%	15%	19%	23%	10%	8%		
2011 Exam 5B	Number		5	14	24	64	47	68	222	171
	Percent	19%	2%	6%	11%	29%	21%	31%		
2011 Exam 5	Number		110	91	143	189	76	44	653	178
	Percent	53%	17%	14%	22%	29%	12%	7%		
2011 Exam 7	Number		1	5	1	2	0	0	9	1
	Percent	78%	11%	56%	11%	22%	0%	0%		
2011 Exam 8	Number		50	35	112	149	47	25	418	93
	Percent	47%	12%	8%	27%	36%	11%	6%		
2010 Exam 9	Number		60	54	160	236	127	75	712	262
	Percent	38%	8%	8%	22%	33%	18%	11%		
2011 Exam 6US	Number		189	180	353	244	30	6	1002	194
	Percent	72%	19%	18%	35%	24%	3%	1%		
2010 Exam 7US	Number		59	73	208	225	47	12	624	278
	Percent	54%	9%	12%	33%	36%	8%	2%		
2011 Exam 6C	Number		34	26	41	41	10	3	155	38
	Percent	65%	22%	17%	26%	26%	6%	2%		
2010 Exam 7C	Number		12	16	39	34	3	3	107	38
	Percent	63%	11%	15%	36%	32%	3%	3%		