CM1103 Coursework 2014 – Assessing Fairness of the Scoring System

C1219157

In order to assess overall fairness, we will investigate different aspects of the system to see how the results change in respect to a competitor's score.

Competitor Consistency

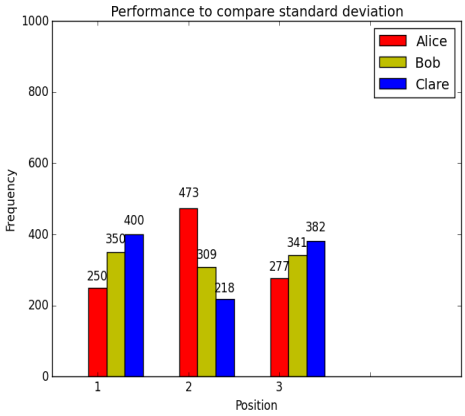

Figure 1

In order to test how consistency affects overall performance, I ran a large sample of races (enough for 1000 series) between 3 sailors who held the same mean performance but with different deviations (0, 5 and 10 for Alice, Bob and Clare respectively) and compared them against each other (Figure 1). I found that as a sailor's performance became more inconsistent the results would begin to properly reflect that inconsistency. For example, Clare has a deviation of 10 and, while she scores first place more often than Alice (who has a deviation of 0), she performs worse than both the other sailors at almost the same rate. While Alice never deviates, she ends up with a healthier balance of beating out 3rd place and reaching 1st place, but Bob (who has half the deviation of Clare) has an even healthier balance than Alice, reaching 1st place much more often than Alice does. What I can draw from these results is that being 100% consistent is relatively safe in terms of performance but a little inconsistency can definitely work to a sailor's advantage (when compared to other sailors of equal skill).

Consistency vs Skill

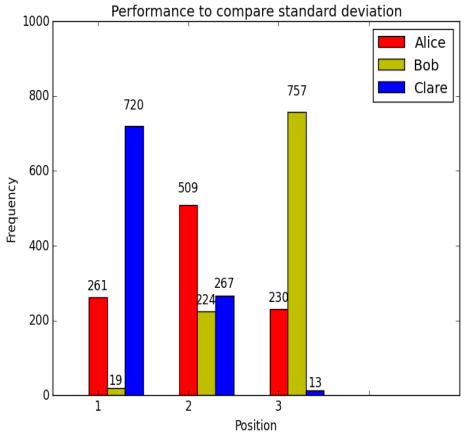

Figure 2

To compare consistency to overall skill, I ran the races again with modified skills. Alice, Bob and Clare had skill values of 100, 95 and 105 respectively while their deviations changed between 0, 5 and 10. The difference in results is staggering. Changing the skill value by even a small amount has a dramatic effect no matter what the circumstance. In Figure 2, the sailors have been tested while Bob and Clare have a deviation of 5 while Alice has a deviation of 10. That bigger deviation should help at least a small amount in getting Alice to keep up with Clare's greater skill level, but Clare still completely overwhelms the competition in a large sample. Conversely, Bob loses out to both other sailors by an even larger majority despite Alice's larger deviation.

Out of consistency and skill, it seems that consistency plays a non-insignificant but relatively minor part in comparison to the raw skill level of a sailor. If a sailor has a higher skill level than another, an inconsistency can only go so far to help catch them. That being said, while weighing up the importance of both aspects I believe that the importance of skill is in a good place with the system as it stands right now.

Effects from Total Participants

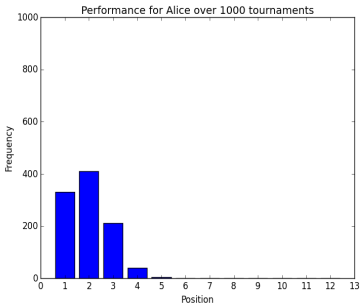

Figure 3

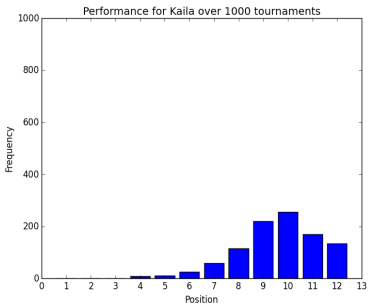

Figure 4

I ran the tests again with 12 sailors (all with varying skill levels and deviations) and the results were what I was expecting. Even when the number of participants increases, the performances of the sailors are correctly reflected. For example, Alice, Bob and Clare have almost the same statistics as the first test, and Alice's results are very similar (Figure 3), with some adjustment because of Dennis, the 4th sailor (who has a much larger deviation of 15 with a skill level of 100, meaning that he was able to get good positioning semi-regularly). Another example is Kaila (Figure 4) who also reflects her positioning in the lower rankings to an expected degree judging from her deviation and skill.

Keeping the Discarded Race

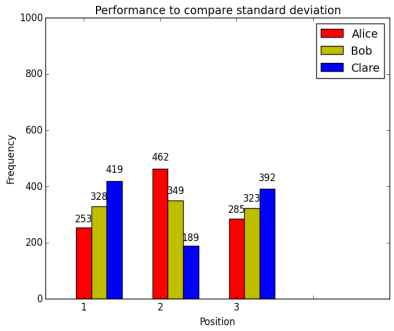

Figure 5

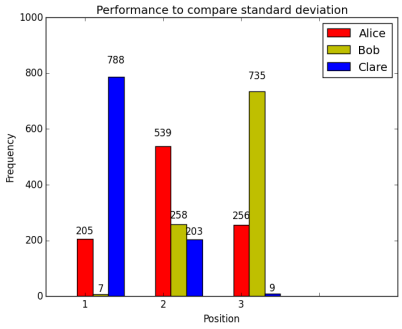

Figure 6

If we remove the code that discards the worst race performance per series, there is an interesting effect in the repeated tests. In the first test, where we compared consistency between 3 sailors, the results do not change by a very significant margin (Figure 5). The balance between the consistencies remains almost the same with a few wins going in either direction. The test for consistency vs. skill also returns very similar results (Figure 6).

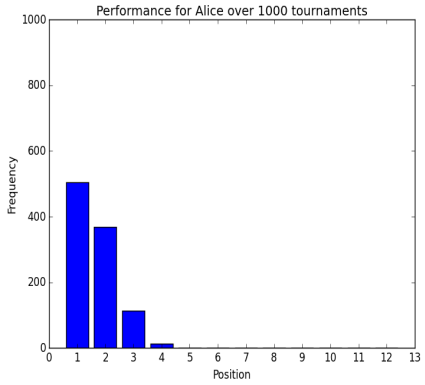

Figure 7

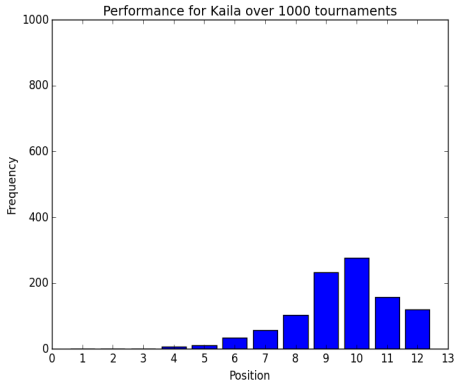

Figure 8

The main differences occur during tests with more sailors. The upper tier of sailors perform significantly better. Alice performs so strongly that she begins to take 1st place in the majority of her races (Figure 7) and is never unlucky enough to hit 5th. Kaila, on the other hand, returns almost identical results (Figure 8), as do the rest of the sailors in the lower half of the race. It seems that not removing the worst performance puts more weight in to sailors that were already performing well and reduces the effect of inconsistencies. For example, Bob and Clare should be taking 1st place away from Alice more often as a result of their deviations beating out Alice's static skill level, but that effect is far less visible here.

Conclusion

I have found this scoring system to be fair in its current implementation. In maintaining the balance between a sailor's consistency and their raw skill, both are accounted for in appropriate measures (skill having the largest impact with inconsistencies shifting the balance to a sensible degree so that an unskilled but wildly inconsistent sailor has to overcome significant odds). The system works just as well with a larger amount of participants. The removal of the worst performance actually works to prevent some imbalance in the system.

Overall, I am unable to envision any other improvements to the scoring system and am satisfied with the results it gives.