Statistical Equity in Ohio Cross Country Competition

Matt Kretchmar, Zhe Wang, John Platt
Department of Mathematics and Computer Science
Denison University
Granville, OH 43023
Email: {kretchmar, wangz, platt_j1}@denison.edu

***Abstract*—The state of Ohio partitions its high schools into different athletic divisions. Divisions are apportioned according to school size in order to make competition between schools as fair as possible. Here we examine the recent historical record of Ohio high schools competing in the sport of cross country distance running to assess the success of the current divisional model in creating fair competition. We employ a chi-square statistical test to measure equity.**

I. Introduction

In the sport of Ohio high school cross country running, we hypothesize that schools with larger enrollments realize a disproportionately unfair competitive advantage over schools of smaller enrollment sizes. To validate this hypothesis, we have identified three different methods of assessing equity:

1) A Monte Carlo simulation of many seasons of Ohio high school competition, measuring the frequency of wins for teams of various sizes.
2) A statistical test of actual Ohio high school cross country competitive results.
3) An exact analytical probability calculation using order statistics to compute the probability of success for schools of different sizes.

The first approach is the subject of our previous work in this problem domain [10], [11]. In that work we built a model of high school running performance and used that model, with actual enrollments of various schools, to compute outcomes of simulated XC seasons. We found a strong bias favoring larger schools over schools with smaller enrollments. The second approach is the subject of this paper; we accumulate five years of recent racing data and perform a $\mathcal{X}^2$ statistical test. The third approach, using order statistics, is the subject of our next work. Note that in all three cases, we use a compare *observed results* against *expected results*. We carefully define and make the case for the expected results in Section II.

The Ohio High School Athletic Association (OHSAA) governs and oversees Ohio's high school interscholastic athletic competitions[5]. Part of their responsibility is to sort Ohio's 735 recognized high school programs into different divisions in order to create fair competition for the teams involved. In the sport of cross country (XC), OHSAA divides the schools into three different, equally-sized athletic divisions[6][1]. The intention of the division process is to create an equal playing field of three divisions with the same number of teams, thus ensuring no team is overly advantaged or disadvantaged by competing against a greater or lesser number of teams.

[1]The actual division of teams is computed separately for boys and girls, and is based on a slightly more complicated enrollment computation.

Figure 1 shows the distribution school population (enrollment numbers) of the 735 recognized high schools in the state of Ohio [2][7]. They vary from the smallest school, The Choffin Career Center, with 47 students to the largest, William Mason High School, with 3531 students. Notably, the distribution is not uniform; there are many more small schools and very few large schools.

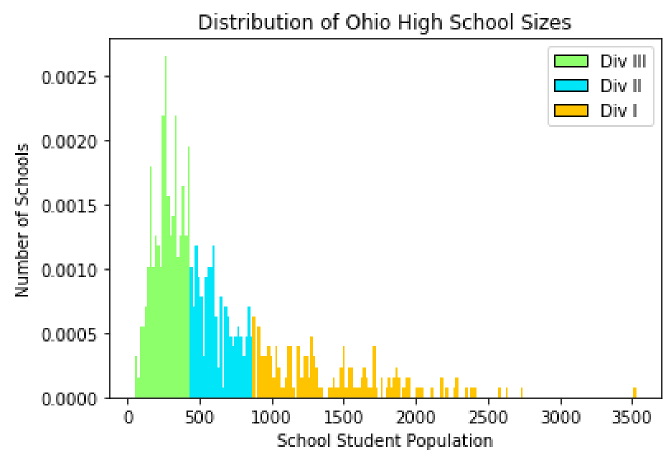

Fig. 1. School Enrollment Distribution

In this same figure, we see the three athletic divisions created for cross country. Division I, composed of the schools with the largest enrollment, is indicated with orange. Division II, medium sized schools, is in blue, while the smallest third of Ohio schools are placed in Division III shown in green. Because of the nature of the school enrollment distribution, we can already see the problem inherent in the OSHAA division process. Namely, the difference in enrollment sizes among Division I schools is much larger than it is for the other two divisions. This disparity creates the competitive inequity in the sport, especially within Division I.

[2]Only 501 of these schools fielded a cross country team in 2018.

II. Fair Distribution and Expected Results

In this section we state a case for a fair distribution of running performance across high schools of different enrollment sizes. These are the results we should *expect* to observe if High School XC competition in Ohio is equitable. We then detail the computation of expected results that we will use to feed our $\mathcal{X}^2$ statistical test.

A. What is Fair?

In order to perform the $\mathcal{X}^2$ test, we need a benchmark of expected performances against which to compare the actual results. This leads to a question that is more philosophical than computational: *What is fair?* Certainly precise equality cannot be expected. Some programs are big, and some are small. This is the nature of competition and, frankly, makes it more interesting. However, there should be some notion of equity. Specifically, smaller programs should feel like they are able to compete against their larger rivals with a non-zero chance of winning. Sports is driven by Hoosiers type stories [12]; hope is what gives sport is soul.

We turn to Aristotle for an answer to our question about fairness. More than two millennia ago, Aristotle penned the *Nicomachean Ethics* [1] where he established the principle of *proportionality* as a standard of equity. Applying that to our situation, we arrive at the following principle of equity

> Ohio cross country distance running is equitable if the probability of success for each team is proportional to that school's population.

More formally, consider n schools with populations $p_1, p_2, ..., p_n$. Then the probability that team i achieves success is proportional to $\frac{p_i}{\sum_j p_j}$. Here is where we can see the attempt at fairness in the current divisional model. By apportioning schools to divisions based on school population, we hope that the difference between the smallest and largest school within a division does not create too much inequity.

"Success" can be measured in many different ways. Success can be a win in a dual meet (with two teams). Success can be measured by a team passing the first round of the playoffs (the District Meet). Success can be the number of top 20 finishes. Success can be the number of state titles each team wins. In each case, we can leverage the proportionality principle to compute an expectation for a particular team hitting a specific benchmark.

In this paper, we measure performance in post-season competition by counting each team's appearance in one of three categories: teams that stop in the first round of playoffs (District Meet), teams that make it to the second round of playoffs (Regional Meet) and teams that make it to the final competition (State Meet). We compute the expected and actual number of appearances at each level for the schools fielding cross country teams in the 2017 through 2021 seasons based on the proportionality principle.

B. Computing Expected Results

The computation of expected success rates involves a little bit of modeling and coding. During the season, each team competes in approximately 10 meets and invitationals. Teams' records in these in-season competitions do not affect their standing in the post season. For the post season, the metric we calculate, each team competes in one of twelve District races where there are usually 12-20 teams at a District competition. The top finishing teams (typically about 5 schools) move on to the second round; the top teams of the District races funnel into one of four Regional competitions featuring about 16 teams per Region. Again the top group (top 5) from each Region move on to the finals, the State Meet. The State meet features the top 20 teams of the state in one final running competition.

Fig. 2. Postseason XC Process

In practice, Districts and Regions are somewhat chosen geographically. Because the schools are not uniformly distributed geographically throughout the state, some District races and Regional races are slightly smaller or larger than the average; in these cases a greater or fewer number of teams may move on to the next level. The State meet always has 20 teams. Keep in mind that there are three different athletic divisions based on school size. This process (District-Regional-State) is replicated separately in all three divisions. Historically, there have been about 200 boys teams in each division and about 165 girls teams in each division over the past five years, though it does vary from year to year. It is also possible that any given team which is on the boundary between two divisions may move between them during a particular period of time, for example, being one of the largest schools in the smaller of the two divisions and then the next year becoming one of the smallest schools as they move up to the next division. School populations are constantly changing and thus team assignments to divisions are in regular flux.

As there are about 200 teams in a division, we cannot compute a probability distribution over all 200 factorial permutations of teams. We must instead use the proportionality principle to feed a simulation of actual post-season competition in order to discover which teams make it to Districts, Regionals and States. We will do this over many simulated seasons to find the expected probabilities of each team making it to a certain milestone in post-season competition. These probability calculations are then leveraged to find expected counts for our Chi-Squared test.

We simulate the competition of these teams during each of the five seasons from 2017 through 2021. For each meet (District, Regional or State), we assemble the list of teams competing in that meet along with their school's enrollment numbers. We then create a weighted random permutation of the teams in this meet by picking teams one at a time to insert into the permutation. The probability of team i being selected next for the permutation is $\frac{p_i}{\sum p_j}$ where p_i is school

i's population and the sum $\sum p_j$ is the sum of populations of all schools not yet selected in the permutation. We then identify the top group (top 5?) from the permutation to move on to the next round of competition. This process assures us that each team's chances of "success" is proportional to its population.

In order to lessen the random variation, we repeat this computation over 10,000 seasons to obtain each team's probability of making it into Districts, Regionals or States. In all, there are 30 separate experiments: 5 years $\times$ 2 genders $\times$ 3 divisions. We find that the variation from year to year is not statistically different, so we can average the results over the five years into one result. That leaves 6 experiments (2 genders $\times$ 3 divisions). Figure 3 is the result from Division I boys:

	Small	Middle	Big	
Districts	0.20	0.36	0.14	0.70
Regionals	0.04	0.11	0.06	0.21
States	0.01	0.04	0.04	0.09
	0.25	0.50	0.25	

Fig. 3. Boys Div I Expected Probability Distribution, 2017-2021

Of the schools competing in this division, we divide them into a Small group, a Middle group, and a Big group. The schools in the bottom quartile of enrollment are in the Small group. The schools with the top quartile of enrollment are in the Big group, leaving the middle 50% in the Middle group. We can see that approximately 70% of the schools stop at the District round, 21% stop at Regionals and only 9% make it to States. We can also see that in a "fair" distribution, larger schools do progress to Regionals and States at a higher rate than smaller schools.

This table enables us to compute the expected counts for the boys Division I competitions over the five year span. We multiply the total count by the probability cells in the expected table to produce the expected counts tables. We show only the first table here (Div I Boys) since they can all be computed with simple multiplication. The expected counts for boys Division I is shown in Figure 4. We have five other similar tables for the remaining experiments (boys div II and div III, girls div I, II and III) which are not shown.

	Small	Middle	Big
Districts	192.7	346.2	137.9
Regionals	39.1	102.7	58.7
States	9.8	45.0	46.0

Fig. 4. Boys Div I Expected Results, 2017-2021

III. Analysis

We now mine the database of Ohio XC results in the years 2017 through 2021 to produce the actual counts for each of the six experiments over a five year span. In any given year, all teams produce a District results, only about 30% of the teams also produce a Regional result, while the top 10% of teams log a State result. The six tables of actual counts are shown in Figures 5 through 10.

	Small	Middle	Big	
Districts	184	361	111	656
Regionals	47	96	80	223
States	12	35	52	99
	243	492	243	978

Fig. 5. Boys Div I Actual Results, 2017-2021

	Small	Middle	Big	
Districts	147	268	91	506
Regionals	47	116	63	226
States	13	34	52	101
	207	418	208	833

Fig. 6. Girls Div I Actual Results, 2017-2021

	Small	Middle	Big	
Districts	182	334	120	636
Regionals	48	118	72	238
States	11	40	49	100
	241	492	241	974

Fig. 7. Boys Div II Actual Results, 2017-2021

	Small	Middle	Big	
Districts	155	272	88	515
Regionals	46	122	68	236
States	10	36	55	101
	211	430	211	852

Fig. 8. Girls Div II Actual Results, 2017-2021

	Small	Middle	Big	
Districts	188	279	138	623
Regionals	33	132	67	232
States	16	53	32	101
	237	482	237	956

Fig. 9. Boys Div III Actual Results, 2017-2021

	Small	Middle	Big	
Districts	165	229	111	505
Regionals	31	119	62	212
States	7	64	30	101
	203	412	202	818

Fig. 10. Girls Div III Actual Results, 2017-2021

IV. Conclusion

Now that we have expected and actual counts for our six experiments, we can compute the $\mathcal{X}^2$ statistical test on each data set and compare the results against a significance level of 0.01.

In all six of these experiments, the p-value is well below the significance level of $\alpha = 0.05$ and thus we conclude that the observed historical records of XC results were not generated by a distribution based on the principle of proportionality. That is, the expected results based on the proportionality principle, are statistically significantly different than the observed results

Expr	**Div**	**Gender**	p-value
1	I	Boys	0.0006124
2	I	Girls	0.0195403
3	II	Boys	0.0011067
4	II	Girls	0.0000236
5	III	Boys	0.0007092
6	III	Girls	0.0016092

Fig. 11. $\mathcal{X}^2$ Results

from the last five years of Ohio XC competition.

This finding agrees with our previous work in using Monte Carlo simulations to produce long-term (10,000 years worth) observed results. In that work we too find that XC competition does not follow the principle of proportionality. We next turn our attention to using Order Statistics to exactly compute the probability of winning for teams of various sizes.

Acknowledgments

The authors would like to thank the Anderson Scholarship Fund and the William G. Bowen and Mary Ellen Bowen Endowed Fund for funding this work. Thanks to Granville Cross Country Coach Bart Smith and former coach Dave Agosta for providing data and insights as to Ohio's process of creating divisions. Thanks are noted to author John Platt for taking several weeks to record and compile the results of thousands of Ohio XC races into a usable database.

References

[1] Aristotle, *Nicomachean Ethics*. New York, NY: Macmillan Publishing Company, 1962.
[2] FloTrack: http://www.wfltrack.org/cross/resources/HowToScoreXC.pdf
[3] Health Line: https://www.healthline.com/health/average-mile-time
[4] Liu, Jun S. *Monte Carlo Strategies in Scientific Computing*. New York, NY: Springer-Verlag. 2001.
[5] Ohio High School Athletic Association: https://ohsaa.org.
[6] OHSAA XC: https://www.ohsaa.org/sports/cc
[7] Ohio School Enrollment Data: http://education.ohio.gov/Topics/Data/Frequently-Requested-Data/Enrollment-Data
[8] Outdoor Industry: https://outdoorindustry.org/press-release/2015-strava-insights-show-cycling-and-running-landscape-in-the-u-s/.
[9] Runner's World: https://www.runnersworld.com/news/g25333911/strava-annual-report-running-statistics/.
[10] Kretchmar, M. Analyzing Equity in High School Cross Country Competition. *Midwest Sports Analytics Meeting*. Central College, Pella, IA. November 23, 2019.
[11] Kretchmar, M. Measuring Equity in High School Cross Country Running. *Great Lakes Data Analytics Conference*. University of Wisconsin Stevens Point. October 2019.
[12] Indy Star. *The Milan Miracle* . August 27, 2013.