An offprint from

The Mathematical Gazette

Modelling learning in youth archery

Volume 104
Number 561
November 2020

MATHEMATICAL ASSOCIATION
supporting mathematics in education

THE MATHEMATICAL ASSOCIATION £33.00

Modelling learning in youth archery

TYLER SKORCZEWSKI

The National Archery in the Schools Program (NASP) began in Kentucky, USA in 2002 and has rapidly expanded to thousands of students around the United States. The program teaches archery in physical education classes and organises tournaments for student archers in elementary school and high school. The program goals include improving student motivation, attention, behaviour, attendance and focus, as well as introducing students to an outdoor skill with the hope that this may increase attention to wildlife conservation efforts in the future.

In the NASP tournaments everyone from elementary school AND high school uses the same equipment. All students shoot the Matthews Genesis bow with Easton arrows. These bows have a variable draw length and allow modest draw weight adjustments so that the same equipment can be adapted to everyone. Students are scored on 30 arrows, each earning 0 (missed target) to 10 points (bullseye) points. Three rounds, called 'ends,' consisting of 5 arrows each are shot at 10 metres, and three more rounds are shot at 15 metres. Thus for each tournament an archer earns a score from 0 to 300. These data are recorded, along with the date of the tournament, school, gender and grade variables, at thousands of tournaments for tens of thousands of students each year by the NASP.

Coaches and archers naturally want to use this information to their benefit. We have data showing tournament scores by date from the 2015-2016 Washington Elementary School team located in Mount Vernon, Iowa. We restrict ourselves to single season data to avoid issues with off-season improvement or declines. In this article we use a mix of ideas from elementary calculus and statistics to develop a model of archer learning that addresses two common questions in youth archery.

1. How can an archery coach use data to make decisions about which students might be recruited to join the tournament team?
2. What are appropriate goals for tournament scores for a new student archer?

Learning in archery

Both of the questions above require an understanding of what it means to improve or learn in archery. We define an archer's performance, $P(t)$, to be his or her score at a tournament on day t, and let $t = 0$ denote the date of the first scored tournament. We can think of $P(t)$ as a learning curve. Research exploring mathematical functions for learning curves has existed for many years, at least as far back as 1919 [1, 2]. Common functions in past research include power law functions, S-shaped curves, and exponential curves [3, 4, 5, 6]. To begin building our learning curve we assume that, given an infinite amount of time, anyone can learn to shoot perfect scores.

While this seems optimistic given the paucity of perfect scores recorded, there is evidence in other fields that show that given enough practice people can move from average to very high performance levels [3, 7]. We also assume that over a single season performance increases continuously at a rate proportional to the distance from a perfect score. Put another way, archers learn faster if there is more left to learn. We call this constant of proportionality the learning rate.

Putting these assumptions together we have an elementary differential equation

$$\frac{dP}{dt} = r(300 - P). \tag{1}$$

In this equation r is the learning rate parameter and has units of reciprocal days. This differential equation is first order, linear and separable. Together with the initial condition $P(0) = P_0$, we can integrate with respect to t to find the solution

$$P(t) = 300 - (300 - P_0)e^{-rt}. \tag{2}$$

We see our learning curve in (2) has an exponential form. In [4] Newell notes that exponential curves appear particularly suited for learning curves dealing with gross motor movements, and the archers in our study must learn several motor functions ranging from feet, arm and hand placements to the motion of the fingers upon release of the arrow.

Determining r

We still need to determine the value of the learning rate parameter r before our model can become useful. To do this we could use nonlinear fitting techniques to fit the model to data, but we do not need such advanced techniques here. Instead we can exploit the properties of logarithms and exponentials. Note that

$$\begin{aligned} P &= 300 - (300 - P_0)e^{-rt} \\ \Rightarrow \ln(300 - P) &= \ln\left((300 - P_0)e^{-rt}\right) \\ &= \ln(300 - P_0) - rt. \end{aligned}$$

Thus if our model is correct and we make a plot with t on the horizontal axis and $\ln(300 - P)$ on the vertical axis we expect to see a linear relationship. Then if we fit a line to these data the slope is $-r$. For an example student archer this procedure is shown in Figure 1 and the resulting model fit in Figure 2. We see that for our example archer here the fitted curve appears to model the general pattern of the data well. Note that if $r < 0$ then the archer is not learning and actually getting worse.

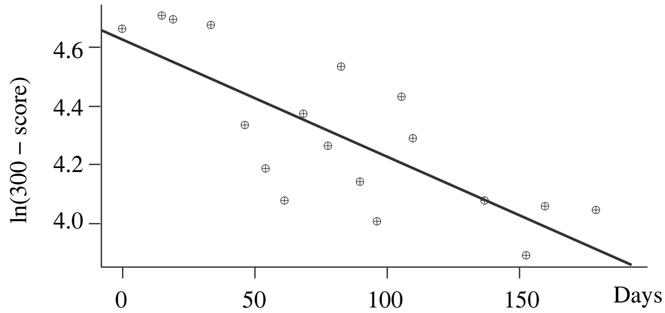

FIGURE 1: The line of best fit ln $(330 - p) = 4.632 - 0.004t$. 0.004, i.e. $-$(coefficient of t), is the learning rate.

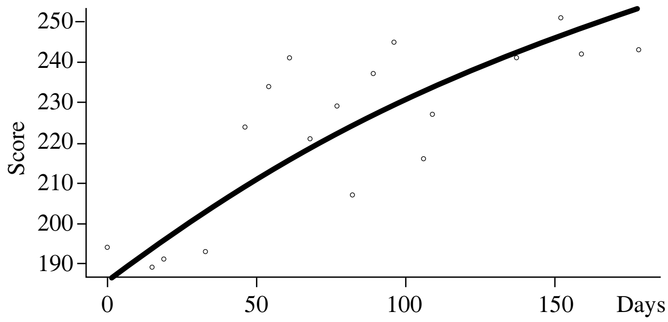

FIGURE 2: Example archer data plotted with fitted model. We see the model does a good job of capturing the general behaviour of the data.

Recruitment scenario

Now that we have a model for how student archers learn and improve, we can use it to address the two questions posed in the introduction. Consider the following scenario: a coach teaches an archery unit in a physical education course early in the fall. At the end of this unit, students shoot a mock tournament and score themselves. The coach wants to recruit some of these students to the tournament team. It is easy to say the coach should focus on the students with the top scores, but what do we mean by 'top'? Let's assume that for this particular scenario, we are focusing on elementary school students, the existing tournament team already contains many high scoring archers, and the coach only wants to recruit those students who have a greater than 50% chance of shooting a score of 250 at an end-of-season tournament that takes place in 140 days time.

An example of such a team is the 2015-2016 Washington Elementary School team from Mount Vernon, Iowa. This team won the 2016 Iowa state tournament and contains 46 talented archers. A relative frequency histogram of learning rates for this team using the analysis described above is shown in Figure 3. It is important to calculate r for individuals and look at a distribution, instead of averaging scores across a group and calculating a single r to represent the team learning rate, because the latter leads to a misleading representation of learning through an artificial smoothing of the data [8]. We see the r values are normally distributed with sample mean $\bar{r} = 0.004 \text{ days}^{-1}$ and standard deviation $s_r = 0.003 \text{ days}^{-1}$. We use this data to help the coach determine whom to recruit.

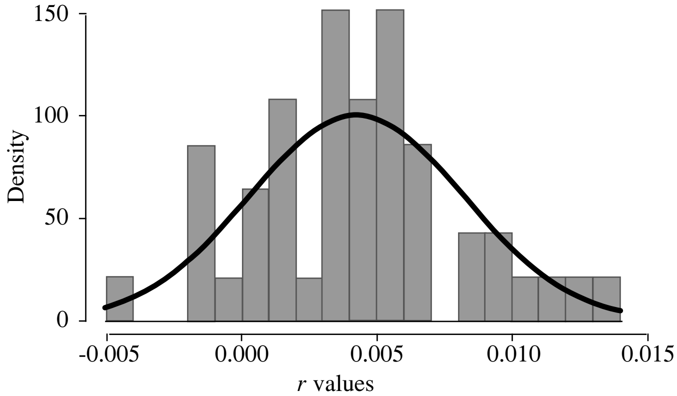

FIGURE 3: Relative frequency histogram of calculated r values and associated normal distribution curve for the 2015-2016 Washington Elementary archery team ($n = 46$)

Monte Carlo approach

Monte Carlo analysis is a good tool to use to answer this type of question, where we have some statistical data and we want to calculate probabilities [9, 10]. To perform a Monte Carlo analysis here, we begin by picking an initial score, P_0. Then we randomly choose an r value from the normal distribution and calculate the predicted score 140 days later using (2). If we randomly choose an r value and predict a score 10000 times, the fraction of predicted scores above 250 approximates the probability of shooting a score greater than 250 at the tournament. Doing this for every possible initial score from 0 to 300 lets us build the graph shown in Figure 4. Analysing this graph we see that the probability of shooting a score greater that 250 at the state tournament is greater than 50% for students who shoot better than 211 in the physical education course. These are the students the coach should recruit.

Goal setting

The primary purpose of NASP is educational. However, archery tournaments are also competitive endeavours. Substantial college scholarship money is attached to some of these tournaments [11]. In addition, we live in a culture where athletic performance is highly prized, and this can cause students and parents considerable stress [12]. Effective goal-setting is a way to alleviate this stress and increase the positive developments that youth sports can provide [13].

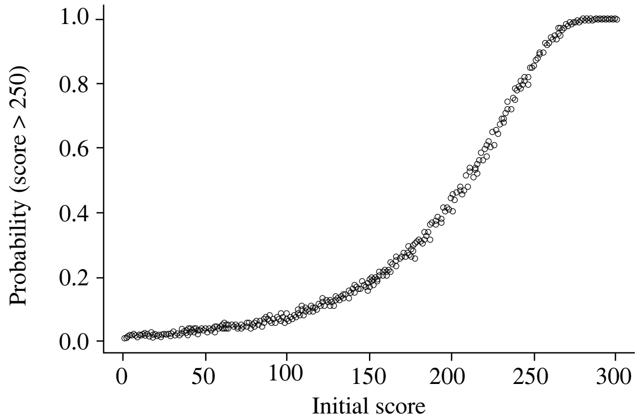

FIGURE 4: Results of a Monte Carlo analysis showing the probability of a Washington Elementary student archer scoring greater than 250 at a tournament 140 days after an initial tournament.

We can use our learning curve to help archers and coaches set realistic and attainable goals. Consider a scenario where a new student archer from Washington Elementary has shot a score of 200 at his or her very first meet ($P_0 = 200$). What are appropriate short (3 week), medium (6 week), and long (12 week) term goals for this archer over the season? We assume the student participates in the same manner as other Washington Elementary students with the same coaching and thus will realise a learning rate that can be sampled from the distribution shown in Figure 3. We let G be a goal score and $T = 21, 42, 84$ days. Substituting these values into our learning curve equation we have

$$G = 300 - (300 - P_0) e^{-r^* T}.$$

Here r^* is the learning rate required to achieve the goal score in T days. Letting $P_0 = 200$ and solving for r^* yields

$$r^* = \frac{\ln\left(\frac{300 - 200}{300 - G}\right)}{T}. \tag{3}$$

If we are sampling from the distribution shown in Figure 3 then the probability of achieving the goal score by T days is equivalent to area under the fitted normal distribution with $r > r^*$ (see Figure 5). These probabilities can be calculated numerically in most standard statistics software.

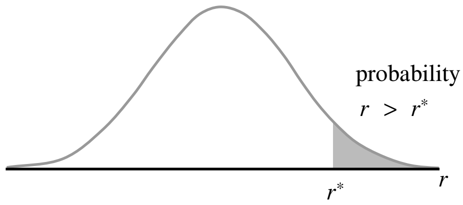

FIGURE 5

Figure 6 shows curves describing the probabilities of achieving various short, medium, and long term goal scores for our hypothetical student. Analysing these curves we see that this student has about a 1% chance of achieving this score in three weeks, about a 35% chance of achieving a score of 220 within six weeks, and about a 75% chance of achieving this score within twelve weeks. Coaches can use this information to relay to the student and his or her parents that a goal of 220 is unrealistic in the short term, ambitious in the medium term, and perhaps a bit modest in the long term.

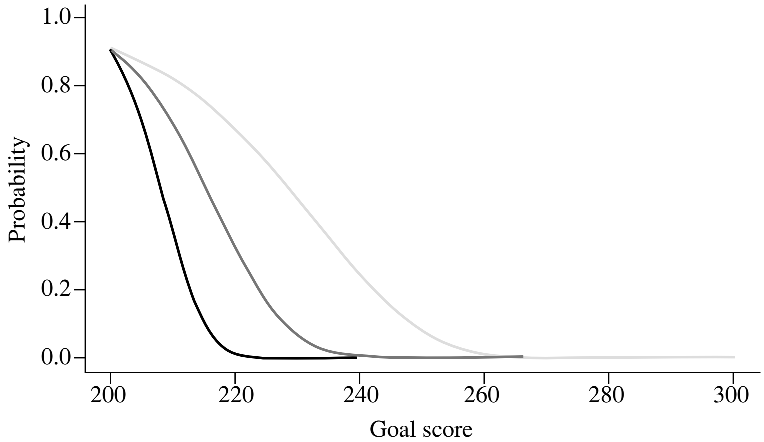

FIGURE 6: Curves showing the probabilities of achieving short (black), medium (dark grey) and long term (light grey) goal scores for an archer with an initial score of 200.

Discussion and further directions

We show how knowledge learned in elementary calculus and statistics courses can be used to answer two questions that arise in youth archery. There are other questions enterprising students may wish to explore. For example given data for two archery teams, one could determine if the learning rate distributions are statistically different. Such information could be valuable for coaches wishing to determine which teaching practices are more effective. In this work we assume the learning rate can be measured with a single parameter that remains constant over an entire season, but archery requires several motor skills involving feet, arms, hands and fingers. Each of these skills may have its own associated learning rate. These ideas might be reflected in the data in Figure 2 where there appear to be two distinct trends, one from 45 to 65 days and another from 70 to 100 days. Each interval may represent a new skill being acquired. We could also allow the learning parameter r to vary with time or performance more generally. Incorporating these ideas into a more advanced model might yield interesting results. Both of the questions addressed in this article and these two proposed topics can be approached with mathematics learned in introductory courses. Moreover, we hope they show students that the mathematics learned in these introductory courses is applicable to real world situations and is not just prerequisite material for more advanced courses.

Acknowledgments

The author wishes to thank the reviewers for their helpful comments to improve this article and his daughter for inspiring the project.

References

1. L. L. Thurstone, The learning curve equation, *Psychological Monographs* **26** (1919) pp. 1-51.
2. M. F. Meyer, F. O. Eppright, The equation of the learning function, *The American Journal of Psychology* **34**(2) (1923) pp. 203-222, available at: http://www.jstor.org/stable/1413573
3. F. E. Ritter, L. J. Schooler, The learning curve, in *International encyclopedia of the social and behavioural sciences*. Amsterdam: Pergamon (2001) pp. 8602–8605.
4. K. M. Newell, Y. T. Liu, G. Mayer-Kress, Time scales in motor learning and development, *Psychological Review*, (2001) **108** pp. 57-82.
5. N. Leibowitz, B. Baum, G. Enden, A. Karniel, The exponential learning equation as a function of successful trials results in sigmoid performance, *Journal of Mathematical Psychology* **54**(3) (2010) pp. 338-340, available at: http: //www.sciencedirect.com/science/article/pii/S0022249610000179
6. J. M. J. Murre, S-shaped learning curves, *Psychonomic Bulletin & Review* **21**(2) (2014) pp. 344-356, available at: http://dx.doi.org/10.3758/s13423-013-0522-0
7. P. F. Delaney, L. M. Reder, J. J. Staszewski, F. E. Ritter, The strategy-specific nature of improvement: the power law applies by strategy within task, *Psychological Science* **9**(1) (1998) pp. 1-7.
8. C. R. Gallistel, S. Fairhurst, P. Balsam, The learning curve: Implications of a quantitative analysis, *Proceedings of the National Academy of Sciences of the United States of America* **101**(36) (2004) pp. 13124-13131 available at: http://www.pnas.org/content/101/36/13124.abstract
9. N. Metropolis, S. Ulam, The Monte Carlo Method, *Journal of the American Statistical Association* **44**(247) (1949) pp. 335-341.

10. K. D. P, B. Tim, T. Thomas, B. Z. I, Why the Monte Carlo method is so important today, *Wiley Interdisciplinary Reviews: Computational Statistics* **6**(6) (2014) pp. 386-392.

11. O. Love, Three of nation's top prep archers come from Eastern Iowa, *Cedar Rapids Gazette* (16 May 2016).

12. J. J. Coakley, Sports in society: issues and controversies, McGraw-Hill (2009).

13. S. Danish, T. Forneris, K. Hodge, I. Heke, Enhancing youth development through sport, *World Leisure Journal* **46**(3) (2004) pp. 38-49.

10.1017/mag.2020.97

TYLER SKORCZEWSKI
410 10th Avenue E.
Menomonie, WI, USA 54751
e-mail: *skorczewski@uwstout.edu*