Riegel's Figure 3 and Riegel's Formula

Andy Long

January 23, 2020

Abstract

We derive Riegel's formula for the case of world-class male runners from the regression results in Figure 3.

It's a little tricky, because the graphs are for log-log transformed data: we want a regression equation of the form

$$t = a\ d^{\ b} \tag{1}$$

where t is in minutes, and d is in kilometers.

1 Data

Riegel's equation is based on the apparent linearity of a plot of log(average speed) versus log(distance) of world class runners, and on the regression analysis he carried out based on that linearity. As he says in his article, "...the logarithms of time vs. distance for each sport were best-fitted to straight lines, using least-squares technique. The curve-fitting process produced an 'endurance equation,', which is a simple power function and requires knowledge of only two basic constants to describe time, speed, and distance over the endurance range for each sport.... The equation is of the form $t = ax^b$, where t = time, x = distance, and a and b are unique for each activity. The constant a is dependent on the units of measurement chosen and has no particular absolute significance...."

The speed values ("Average running speed", in units of meter/second) occur along the y-axis, although their occurrence along that axis is not uniformly spaced. This is because it is actually the **logs** of these values which are represented **uniformly** along that axis.

The distance value (given in kilometers) appear along the x-axis, and, again, are not uniformly spaced – rather their logs are.

The data look linear if plotted in coordinates $\ln(speed)$ and $\ln(distance)$ – so the labeling used is a little confusing.

There is another important issue we must resolve before proceeding: the issue of units. Speeds are given in terms of m/sec, while distances are given in km. In the end, Riegel gave the equation entirely in terms of km and min, so we need to

a. transform the speeds to km/min;

b. log-transform the speed and distance values;

c. estimate the equation of the line from two points;

Figure 1: Each of these lines is a best-fit line to a set of log-log transformed data. We focus on the top line, which represents data on world-class male runners at a wide variety of distances, and the best-fitting linear model which leads to Riegel's equation (Eq. (1)). This graphic is taken from Riegel's 1981 publication[1].

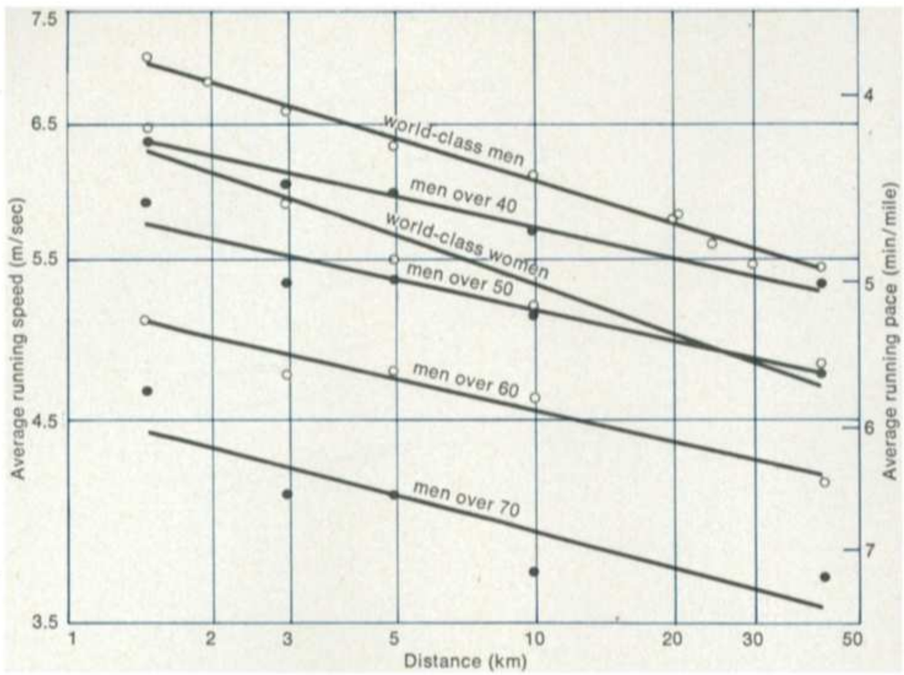

Figure 3. Runners provide the greatest amount of data for performance comparison. World records for various classes of runners highlight the parallelism between world-class men and women, and also among the men from 40 to 70 years of age. Parallel lines result from identical fatigue factors.

d. transform the equation so that it is in the form of Riegel's formula, Eq. (1). I.e., we solve for t.

Unit conversion and log transformation lead to us relabeling the axes according to the following values:

$Speed(m/s)$	$Speed(km/min)$	$ln(Speed)$	$Distance$	$ln(Distance)$
7.5	.45	$-.80$	30	3.40
6.5	.39	$-.94$	20	3.00
5.5	.33	-1.11	10	2.30
4.5	.27	-1.31	5	1.61
3.5	.21	-1.56	3	1.10
			2	.70
			1	0

We pick two points, then, from the line, using the new coordinates $(\ln(distance), \ln(speed))$: I spy, with my little eye, the points (.70,-.883) and (3.40,-1.093), which fall on the straight regression line (sort of). We're not going to get this exactly right! Riegel determined the equation of the line using linear regression from the data points shown – we're just guessing at the equation for the line from two points on the line.

These two points will determine our line: hopefully we've chosen pretty well.

2 Calculations

The two points yield a slope estimate of

$$m = \frac{-.883 - -1.093}{.70 - 3.40} = 0.0778$$

From these, we estimate the linear equation (using point-slope form) as

$$(\ln(speed) - (-1.093)) = -0.0778(\ln(distance) - 3.40)$$

Simplifying a little,

$$\ln(speed) = -0.0778 \ln(distance) + (0.0778 * 3.40 - 1.093)$$

or

$$\ln(speed) = -0.0778 \ln(distance) - 0.82848$$

Exponentiating both sides, we obtain

$$speed = distance^{-0.0778} e^{-0.82848}$$

But distance is d and speed is simply $\frac{d}{t}$; so we can write

$$\frac{d}{t} = d^{-0.0778} e^{-0.82848}$$

Solving for t, we obtain

$$t = d^{1.0778} e^{.82848}$$

or

$$t = 2.290 d^{1.0778}$$

Riegel (who actually had the data points themselves, rather than the regression line to work from), obtained a model of

$$t = 2.299 d^{1.07732}$$

Pretty close!

3 Conclusions

Riegel built his model based on empirical data – data on world-record times of world-class runners, across the spectrum of distances available.

Transforming his data using logarithms, he determined that the data fell along a line in log-log space. He used a procedure called "linear regression" to find the equation of this line, and then back-transformed the equation to obtain his final model

$$t = 2.299 d^{1.07732}$$

We were able to "reverse engineer" his equation, using his graph, and verify that this, indeed, is where it came from.

References

[1] P. S. Riegel. Athletic records and human endurance: A time-vs.-distance equation describing world-record performances may be used to compare the relative endurance capabilities of various groups of people. *American Scientist*, 69(3):285–290, 1981.