Gaussian Record Holders

Mik Zlatin

March 13, 2022

Intro

Is it just me, or are athletes getting better these days?[1]

Let's focus on long-jump. The world record for long-jump distance has been steadily increasing the first Olympics in ancient Greece.

But does this mean that athletes are getting better, or is it simply due to more and more people attempting the long-jump over time?

We will attempt to answer this question as follows. Suppose that the distance a person can jump is a Gaussian random variable with mean μ and standard deviation σ. If random people of varying skill levels attempt the long-jump one at a time, how long do we expect to wait before the first world record is broken?

You might complain that no one cares about the first world record, and that over time, it should become harder and harder to break the previous record. However, we will just focus on the first record to be broken. I think that once we find the answer, it should be clear why we do this.

Puzzle

Formally, the problem is as follows. You repeatedly draw numbers from a normal distribution G with mean μ and standard deviation σ. A draw is a world record when it is the highest so far. Obviously the first draw will technically be the highest so far, but this doesn't count.

Let X be the random variable whose value is the number of draws (after the first one) it takes to break the first record. What is $E[X]$?

Solution: There are several valid approaches to solving this question. Here is one:

Let's condition on the value of the first draw x. Then X is simply the number of draws until we sample a value greater than x. Since each sample is independent, this is a geometric random

[1] I heard the idea for this puzzle from Ben Edelman

variable with parameter $p = 1 - F(x)$, where $F(x)$ is the cdf of a normal random variable $G \sim N(\mu, \sigma)$. In this case, the expected value of X is simply $1/p = \frac{1}{1-F(x)}$.

Since we conditioned on the value of x, we now integrate over all possibilities for x. So,

$$E[X] = \int_{-\infty}^{\infty} f(x) \cdot \frac{dx}{1 - F(x)}$$

where $f(x)$ is the pdf of G. How can we compute this integral?

Recall that $\frac{d}{dx}F(x) = f(x)$. We can perform a change of variables to solve this integreal. Let $u = F(x)$, so $du = f(x)dx$.

We have (remembering to change the limits of integration):

$$E[X] = \int_0^1 \frac{du}{1-u} = -\ln(|1-u|)\Big|_0^1$$

This integral diverges to ∞.

Remark. It's somewhat surprising that the expectation is infinite. According to our model, this means that we would have to wait an infinite amount of time before seeing another world record, assuming the skill levels of athletes don't change over time. There are clearly important differences between our model and the real world, but it is still interesting.

Remark. Where in this calculation did we use the fact that G was a Gaussian random variable? We didn't! The result holds for any distribution G.