

*Original from J. Auerbach on 2 February 2020 ...*

I figured the first step was to write  $n$  as a function of  $S_a^b$ . (Dividing by  $(b-a)$  shouldn't make a difference.) Then you can calculate variances, covariances, the regression slope, etc. (Although it's not obvious to me how to do this in closed form.)

It's actually easier to write  $S_a^b$  as a polynomial in  $n$ .

We defined  $n$  to be the day such that the cumulative temperature hit  $\beta$ :

$$\beta \approx \sum_{i=a}^n X_i$$

Assuming  $n$  is large, we can replace this sum with its average

$$\sum_{i=a}^n X_i = \sum_{i=a}^n \mu * i$$

So,

$$\beta \approx \sum_{i=a}^n \mu * i = \mu * \sum_{i=a}^n i$$

and,

$$S_a^b = \mu * \sum_{i=a}^b i = \mu * \sum_{i=a}^n i + \mu * \sum_{i=n}^b i = \beta + \mu * \sum_{i=n}^b i$$

But  $\sum_{i=n}^b i$  is just the sum of the integers from  $n$  to  $b$  which is the sum of integers from 1 to  $b$  minus the sum of integers from 1 to  $n$

$$b * (b + 1) / 2 - n * (n + 1) / 2$$

So,

$$S_a^b = \beta + \mu * (b * (b + 1) / 2 - n * (n + 1) / 2)$$

Letting  $\alpha_1 = \beta + \mu * b * (b + 1) / 2$  and  $\alpha_2 = \mu / 2$

$$S_a^b = \alpha_1 - \alpha_2 * n * (n + 1)$$

which is clearly not a straight line, and OLS (either  $S_a^b$  on  $n$  or  $n$  on  $S_a^b$ ) is not a good idea.

*Original from J. Auerbach on 22 January 2020 ...*

Sorry, I'm confused by the additional notation. Is it correct to denote the time window  $[a, b]$  and  $S_a^b = \sum_{i=a}^b X_i$ ? The question is then how  $a$  and  $b$  relate to  $n$ , the first time,  $i$ , that  $S_a^i$  hits  $\beta$ , i.e.  $\sum_{i=a}^n X_i \approx \beta$ .

I originally assumed  $a = 0$  and  $b = n$ , in which case  $M_b = M_n \approx \beta / n$ .

If  $a, b$  are chosen independently of  $n$ , for example the entire year/spring, then the relationship would be similar:

Fix  $\beta$  so that:

$$\beta \approx \sum_{i=a}^n X_i \approx \mu \sum_{i=a}^n$$

or

$$\mu = \beta / \sum_{i=a}^n = \beta / ((n * (n + 1) - a * (a + 1)) / 2) \approx 2 * \beta / n * (n + 1)$$

*Edited by Lizzie ...*

Let:

$i$  = index the days,  $i = 0, 1, \dots, N$

$X_i$  = temperature on day  $i$ , assume  $X_0 = 0$

$\mu$  = average temperature on day  $i = 1$ ;  $X_i \sim \mathcal{N}(\mu * i, \sigma), i > 0$

$$S_i = \sum_{X_1}^{X_n} X_i, \text{ (for example, GDD)}$$

$M_i$  = cumulative mean,  $S_i/i$  and thus  $M_n = S_n/n$

$\beta$  = a threshold of interest,  $\beta > 0$ , (for example,  $F^*$  or required GDD)

$n$  = the first day such that  $\beta < S_n$ , (for example, day of budburst)

$l$  = the last date over which temperature is integrated

‘Sensitivity’ is generally measured by taking the slope—I am calling it  $m$  here—of a linear regression of  $n$  against  $M_l$ .

It is not surprising that  $n$  decreases with  $\mu$  (and thus  $M_l$ ) under the null model (this is equivalent to saying it is not surprising that plants leafout earlier with warmer spring temperatures, we all are fine with this). Researchers are surprised that the decrease of  $m$  as  $\mu$  increases and suggest this is evidence of declining sensitivities;  $n$  is “reacting” or “changing” to  $M_l$ .

So I say (I think, assuming the math still holds): Plotting  $n$  against  $M_l$  is similar to plotting  $n$  against  $S_n/n$ , which is approximately the same as plotting  $n$  against  $\beta/n$ , which is a hyperbola. The data is consistent with the null model; no need for  $n$  to react to  $M_l$ .

*Original from J. Auerbach on 15 January 2020, with corrections from 22 Jan 2020 ...*

Let:

$i$  = index the days,  $i = 0, 1, \dots, N$

$X_i$  = temperature on day  $i$ , assume  $X_0 = 0$

$\mu$  = average temperature on day  $i = 1$ ;  $X_i \sim \mathcal{N}(\mu * i, \sigma), i > 0$

$$S_i = \sum_{X_1}^{X_n} X_i$$

$M_i$  = cumulative mean,  $S_i/i$  and thus  $M_n = S_n/n$

$\beta$  = a threshold of interest,  $\beta > 0$

$n$  = the first day such that  $\beta < S_n$

Note: a time  $n$ , we have  $M_n = S_n/n \approx \beta/n$ .

Some say: It is NOT surprising that  $n$  decreases in  $\mu$  (and thus  $M_n$ ) under the null model. It IS surprising that the decrease of  $n$  in  $M_n$  slows under the null model. This may be evidence of declining sensitivities;  $n$  is “reacting” or ”changing” to  $M_n$ .

You say: Plotting  $n$  against  $M_n$  is the same as plotting  $n$  against  $S_n/n$ , which is approximately the same as plotting  $n$  against  $\beta/n$ , which is a hyperbola. The data is consistent with the null model; no need for  $n$  to react to  $M_n$ .

If this is the correct framing of the problem,  $X_i$  is a gaussian random walk with drift and  $n$  is a hitting time. I believe the mean and variance are difficult to work out in closed form since you would have to integrate  $S_n$  over  $n$ . However, the continuous time generalization can be worked out: a Brownian motion with instantaneous variance sigma and drift mu. The time,  $t$ , the process hits some threshold beta is distributed Inverse Gaussian Distribution, having mean  $\mu/\beta$  and variance  $\mu * \sigma/\beta^3$ . This seems like a more accurate description of the problem anyway since the relationship between temperature and time is not discrete.