

1 Question: What information does temperature sensitivity capture?

Consider the following plant-watching experiment. Let the constant s denote the total amount of temperature needed for a plant to bloom (i.e. leaf out). Suppose the plant begins collecting temperature on day 0. Let the random variable Y denote the day of leafout. Then the average temperature absorbed by the plant on all days between 0 and Y is the random variable $X = \frac{s}{Y}$. We will primarily be interested in the inverse relationship: $Y = \frac{s}{X}$.

We assume the experiment is to be performed n times—perhaps with n different plants, each experiencing different daily temperatures. The data are the n leafout days, $Y = \{Y_1, \dots, Y_n\}$, and the n mean temperatures $X = \{X_1, \dots, X_n\}$. We assume all n plants have the same temperature sum s .

The question is what information is given by the regression slope $\beta = \frac{\text{Cov}(X, Y)}{\text{Var}(X)}$, called the ‘temperature sensitivity’ or ‘sensitivity’. In particular, whether it gives any information about s . Let μ and σ^2 denote the mean and variance of X . i.e. $\mu = E[X]$ and $\sigma^2 = E[(X - \mu)^2]$.

2 Temperature sensitivity can be decomposed into three factors: temperature sum, temperature percent change, temperature precision

First note, the regression coefficient can be simplified:

$$\beta = \frac{\text{Cov}(X, Y)}{\sigma^2} = \frac{\text{Cov}(X, \frac{s}{X})}{\sigma^2} = s \frac{\text{Cov}(X, \frac{1}{X})}{\sigma^2} = s \frac{E[X \frac{1}{X}] - E[X]E[\frac{1}{X}]}{\sigma^2} = s \frac{1 - \mu E[\frac{1}{X}]}{\sigma^2} = s \frac{E[1 - \frac{\mu}{X}]}{\sigma^2} = s \frac{E\left[\frac{X - \mu}{X}\right]}{\sigma^2}$$

The numerator, $E\left[\frac{X - \mu}{X}\right]$ is the average percent change if X were to be replaced or summarized by its mean, μ . This statistic is not unusual; ‘average percent change’ and ‘average absolute percent change’ are common alternatives to the measures ‘bias’ and ‘standard deviation’. (n.b. numerator is different than percent bias, $\frac{E[X - \mu]}{\mu} = E\left[\frac{X - \mu}{\mu}\right]$)

Thus, the regression coefficient, temperature sensitivity, can be decomposed into three parts three factors: temperature sum, temperature percent change, temperature precision (i.e. one over temperature variance).

If two sets of experiments are performed, say in two different time periods, the first with temperature sum s_1 and data X_1, Y_1 and the second with temperature sum s_2 and data Y_1, Y_2 , then comparing β_1 and β_2 yields difference

$$\frac{\beta_1}{\beta_2} = \frac{s_1}{s_2} \frac{E\left[\frac{X_1 - \mu_1}{X_1}\right]}{E\left[\frac{X_2 - \mu_2}{X_2}\right]} \frac{\sigma_2^2}{\sigma_1^2}$$

It follows that change in the temperature sensitivity may result from the change in the temperature sum, the change in the ‘average percent change’, or the change in the temperature variance—or any combination of the three. In fact, temperature sums may increase, i.e. $\frac{s_1}{s_2} > 1$, but overall sensitivity declines, either because $\frac{E\left[\frac{X_1 - \mu_1}{X_1}\right]}{E\left[\frac{X_2 - \mu_2}{X_2}\right]} < 1$ or $\frac{\sigma_2^2}{\sigma_1^2} < 1$.

3 Further simplification under normality

So far we have not made any assumption about the distribution of X . However, it follows from the Markov chain central limit theorem that the distribution of X is approximately normal. Some simplification is then possible.

Note that the numerator of $\beta = s \frac{E\left[\frac{X - \mu}{X}\right]}{\sigma^2}$, $E\left[\frac{X - \mu}{X}\right]$, can be approximated by Taylor expansion around μ :

$$E\left[\frac{X - \mu}{X}\right] = \sum_{k=1}^{\infty} (-1)^{k+1} \mu^{-k} E[(X - \mu)^k]$$

Since $X \sim \text{normal}(\mu, \sigma)$, $E[(X - \mu)^k] = 0$ for k odd and $\sigma^k (k-1)!!$ for k even (n.b. the double ! is a factorial only including every other term, e.g. $5!! = 5 \times 3 \times 1 = 15$) leading to the simplification.

$$E\left[\frac{X - \mu}{X}\right] = -\sum_{k=1}^{\infty} \left(\frac{\sigma}{\mu}\right)^{2k} (2k-1)!! = -\left(\frac{\sigma}{\mu}\right)^2 - 3\left(\frac{\sigma}{\mu}\right)^4 - 15\left(\frac{\sigma}{\mu}\right)^6 - 105\left(\frac{\sigma}{\mu}\right)^8 - \dots$$

and

$$\beta = s \frac{E\left[\frac{X - \mu}{X}\right]}{\sigma^2} = -\frac{s}{\mu^2} + O(\sigma^4/\mu^4)$$

I think it's probably safe to approximate $\beta = -\frac{s}{\mu^2}$. For example, Washington DC, spring-time temperature is 55 degrees with a standard deviation of 2.3 degrees. Assuming plants collect temperature over springtime, $\beta = 0.0003323282s$ while the first order approximation is $0.0003305785s$. (The error is even smaller if you only take the last twenty-five years; the mean is 57, and the sd is 1.9.)

Returning to the comparison of two sets of experiments, we have for $k = 1$,

$$\frac{\beta_1}{\beta_2} = -\frac{s_1 \mu_2^2}{s_2 \mu_1^2}$$

For $k = 2$,

$$\frac{\beta_1}{\beta_2} = -\frac{s_1 \mu_2^2}{s_2 \mu_1^2} \frac{1 + 3 \frac{\sigma_1^2}{\mu_1^2}}{1 + 3 \frac{\sigma_2^2}{\mu_2^2}}$$

etc.