

CONFIRMATION BIAS IN SYSTEMS ENGINEERING: A PEDAGOGICAL EXAMPLE

Griselda Acosta

Doctoral Student, e-mail: gvacosta@miners.utep.edu

Eric D. Smith

Ph.D. (Systems & Industrial Engineering), Associate Professor, e-mail: esmith2@utep.edu

Vladik Kreinovich

Ph.D. (Phys.-Math.), Professor, e-mail: vladik@utep.edu

University of Texas at El Paso, El Paso, Texas 79968, USA

Abstract. One of the biases potentially affecting systems engineers is the confirmation bias, when instead of selecting the best hypothesis based on the data, people stick to the previously-selected hypothesis until it is disproved. In this paper, on a simple example, we show how important it is to take care of this bias: namely, that because of this bias, we need twice as many experiments to switch to a better hypothesis.

Keywords: confirmation bias, systems engineering, hypothesis testing.

1. Formulation of the Problem

Confirmation bias. It is known that our intuitive reasoning shows a lot of unexpected biases; see, e.g., [2]. One of such biases is a *confirmation bias*, when, instead of selecting the best hypothesis based on the data, people stick to the previously-selected hypothesis until it is disproved. This bias is ubiquitous in systems engineering; see, e.g., [1, 4, 5, 8].

How important is it to take the confirmation bias into account? Taking care of the confirmation bias requires some extra effort; see, e.g., [3, 7–9] and references therein. A natural question is: is the resulting improvement worth this extra effort? How better the result will we get?

In this paper, on a simple example, we show that the result is drastically better: namely, that if we properly take this bias into account, then we will need half as many experiments to switch to a more adequate hypothesis.

2. Analysis of the Problem

Description of the simple example. Let us consider the simplest possible case when we have a parameter a that may be 0 and may be non-zero, and we directly observe this parameter. We will also make the usual assumption that the observation inaccuracy is normally distributed, with 0 mean and known standard deviation σ .

In this case, what we observe are the values x_1, \dots, x_n which are related to the actual (unknown) value a by a relation $x_i = a + \varepsilon_i$ ($i = 1, \dots, n$), where ε_i are independent normally distributed random variables with 0 means and standard deviation σ .

Two approaches. In the ideal approach, we select one of the two models – the null-hypothesis $a = 0$ or the alternative hypothesis $a \neq 0$ – by using the usual Akaike Information Criterion (AIC); see, e.g., [6].

In the confirmation-bias approach, we estimate the value a based on the observations x_1, \dots, x_n , and we select the alternative hypothesis only if the resulting estimate is statistically significantly different from 0 – i.e., e.g., that the 95% confidence interval for the value a does not contain 0.

What if we use AIC. In the AIC, we select a model for which the difference $AIC \stackrel{\text{def}}{=} 2k - 2 \ln(\widehat{L})$ is the smallest, where k is the number of parameters in a model and \widehat{L} is the largest value of the likelihood function L corresponding to this model.

The null-model $a = 0$ has no parameters at all, so for this model, we have $k = 0$. For n independent measurement results, the likelihood function is equal to the product of the values

$$\frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot \exp\left(-\frac{x_i^2}{2\sigma^2}\right)$$

of the Gaussian probability density function corresponding to these measurement results x_i . Thus,

$$L = \prod_{i=1}^n \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot \exp\left(-\frac{x_i^2}{2\sigma^2}\right)$$

and so, for this model,

$$AIC_0 = -2 \ln(L) = 2n \cdot \ln(\sqrt{2\pi} \cdot \sigma) + \frac{1}{\sigma^2} \cdot \sum_{i=1}^n x_i^2.$$

We assume that $x_i = a + \varepsilon_i$, where the mean value of ε_i is 0 and the standard deviation is σ . Thus, the expected value of x_i^2 is equal to $a^2 + \sigma^2$. For large values n , due to the Law of Large Numbers (see, e.g., [6]), the average $\frac{1}{n} \cdot \sum_{i=1}^n x_i^2$ is approximately equal to the expected value $E[x_i^2] = a^2 + \sigma^2$. Thus, $\sum_{i=1}^n x_i^2 \approx n \cdot (a^2 + \sigma^2)$ and hence,

$$AIC_0 = 2n \cdot \ln(\sqrt{2\pi} \cdot \sigma) + \frac{1}{\sigma^2} \cdot n \cdot (a^2 + \sigma^2). \tag{1}$$

The alternative model $a \neq 0$ has one parameter a , so here $k = 1$. The corresponding likelihood function is then equal to

$$L = \prod_{i=1}^n \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot \exp\left(-\frac{(x_i - \widehat{a})^2}{2\sigma^2}\right).$$

We select the parameter a that maximizes the value of this likelihood function. Maximal likelihood is the usual way of estimating the parameters, which in this case leads to $\hat{a} = \frac{1}{n} \cdot \sum_{i=1}^n x_i$. For large n , this estimate is close to the actual value a , so we have

$$\hat{L} = \prod_{i=1}^n \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot \exp\left(-\frac{(x_i - a)^2}{2\sigma^2}\right).$$

For this model, $x_i - a = \varepsilon_i$, thus,

$$\text{AIC}_1 = 2 - 2 \ln(\hat{L}) = 2 + 2n \cdot \ln(\sqrt{2\pi} \cdot \sigma) + \frac{1}{\sigma^2} \cdot \sum_{i=1}^n \varepsilon_i^2.$$

For large n , we have $\sum_{i=1}^n \varepsilon_i^2 \approx n \cdot \sigma^2$, hence

$$\text{AIC}_1 = 2 + 2n \cdot \ln(\sqrt{2\pi} \cdot \sigma) + \frac{1}{\sigma^2} \cdot n \cdot \sigma^2. \quad (2)$$

The second model is preferable if $\text{AIC}_1 < \text{AIC}_0$. By deleting common terms in these two values AIC_i , we conclude that the desired inequality reduces to $2 < \frac{n \cdot a^2}{\sigma^2}$, i.e., equivalently, to

$$n > \frac{2\sigma^2}{a^2}. \quad (3)$$

What if we use a confirmation-bias approach. In the confirmation-bias approach, we estimate a – and we have already mentioned that the optimal estimate is $a = \frac{1}{n} \cdot \sum_{i=1}^n x_i$. It is known (see, e.g., [6]) that the standard deviation of this estimate is equal to $\sigma_e = \frac{\sigma}{\sqrt{n}}$. Thus, the corresponding 95% confidence interval has the form $[a - 2\sigma_e, a + 2\sigma_e]$. The condition that this interval does not contain 0 is equivalent to $|a| > 2\sigma_e$, i.e., equivalently, to $a^2 > 4\sigma_e^2$. Substituting the above expression for σ_e into this inequality, we conclude that $a^2 > 4 \cdot \frac{\sigma^2}{n}$, i.e., equivalently, that

$$n > \frac{4\sigma^2}{a^2}. \quad (4)$$

Conclusion. By comparing the expressions (3) and (4) corresponding to the two approaches, we can indeed see that the confirmation-bias approach requires twice as many measurements than the approach in which we select the best model based on the data.

Thus indeed, avoiding confirmation bias can lead to a drastic improvement in our estimates and thus, in our decisions.

Acknowledgments

This work was supported in part by the National Science Foundation grants 1623190 (A Model of Change for Preparing a New Generation for Professional Practice in Computer Science) and HRD-1242122 (Cyber-ShARE Center of Excellence).

REFERENCES

1. Bohlman J. and Bahill A.T. Examples of mental mistakes made by systems engineers while creating tradeoff studies. *Studies in Engineering and Technology*, 2014, vol. 1, no. 1, pp. 22–43.
2. Kahneman D. *Thinking, Fast and Slow*. Farrar, Straus, and Giroux, New York, 2011.
3. Koriat R.M., Lichtenstein S., and Fischhoff B. Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory*, 1980, vol. 6, pp. 107–118.
4. Nickerson R.S. Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, 1998, vol. 2, pp. 175–220.
5. Petroski H. Framing hypotheses: A cautionary tale. *American Scientist*, 2003, vol. 91, no. 1, pp. 18–22.
6. Sheskin D.J. *Handbook of Parametric and Nonparametric Statistical Procedures*. Chapman and Hall/CRC, Boca Raton, Florida, 2011.
7. Slovic P. and Fischhoff B. On the psychology of experimental surprises. *Journal of Experimental Psychology: Human Perceptions and Performance*, 1977, vol. 3, pp. 544–551.
8. Smith E.D. and Bahill A.T. Attribute substitution in systems engineering. *Systems Engineering*, 2010, vol. 13, no. 2, pp. 130–148.
9. Wickelgren W.A. *How to Solve Problems: Elements of a Theory of Problems and Problem Solving*. Freeman, San Francisco, 1974.

СКЛОННОСТЬ К ПОДТВЕРЖДЕНИЮ В СИСТЕМНОЙ ИНЖЕНЕРИИ: ПЕДАГОГИЧЕСКИЙ ПРИМЕР

Г. Акоста

докторант, e-mail: gvacosta@miners.utep.edu

Э.Д. Смит

к.т.н., доцент, e-mail: esmith2@utep.edu

В. Крейнович

к.ф.-м.н., профессор, e-mail: vladik@utep.edu

Техасский университет в Эль Пасо, США

Аннотация. Одним из предубеждений, потенциально влияющих на системных инженеров, является склонность к подтверждению, когда вместо того, чтобы выбрать лучшую гипотезу на основе данных, люди придерживаются ранее выбранной гипотезы до тех пор, пока она не будет опровергнута. В данной работе на простом

примере мы показываем, как важно обращать внимание на это предубеждение — а именно, что из-за этого предубеждения нам нужно в два раза больше экспериментов, чтобы перейти к лучшей гипотезе.

Ключевые слова: склонность к подтверждению, системная инженерия, проверка гипотез.

Дата поступления в редакцию: 06.12.2019