Surveys in Mathematics and its Applications

ISSN 1842-6298 (electronic), 1843-7265 (print)
Volume 12 (2017), 65 – 70

ACQUIRED KNOWLEDGE AS A STOCHASTIC PROCESS

George Stoica and Barry Strack

Abstract. The paper argues that stochastic and deterministic knowledge complement and improve upon each other. We introduce a stochastic model for acquired knowledge and prove that the numerical data provided fits quite well the estimated outcomes of the model.

1 Purpose and data

Between 2005 and 2015, a number of 420 students registered and finalized an upper-level specialized undergraduate program, under the first author's coordination. Among them, 312 students had the correct and complete pre-requisites for that course, whereas the other 108 had only partial exposure to the pre-requisites, due to objective reasons, such as: they were coming from another university, obtained equivalents for pre-requisites having slightly different content, certain topics have not been taught by lack of time (bad weather cancellations), took the pre-requisites long time ago, when the curriculum was less charged than today, etc. However, the performance of all 420 students has been evaluated using the same criteria: a midterm worth 30% of the final grade, a final exam worth 50% and assignments worth 20%. The overall failure rate at the midterm was 30.48%, and at the final exam of 14.76%. The assignments were, in average, in the range of 80%, and this seems to have contributed to reducing the failure rate between midterm and final exam.

Motivated by the second author's interest in the possibility of representing this model in probabilistic terms, in this paper we shall describe a stochastic model that incorporates and explains quite faithfully this type of acquired knowledge, based on a deterministic part (coming from taking the correct pre-requisites) and a random or stochastic part (coming from partial exposure to the pre-requisites). Inter alia, the

2010 Mathematics Subject Classification: 91E10; 91E40; 60J60; 62F10.
Keywords: Knowledge; Stochastic model; Parameter estimation.
There is a Corrigendum for this paper published volume 13 (2018)

model will demonstrate that, in spite of poor odds, students with incomplete prerequisites obtained better results than expected, by working harder on the midterm and assignments, and thus correctly explaining the small failure rates at that course.

The pioneering ideas from [5] on deterministic models of acquired knowledge have been developed in [1] and [6], and our model can be seen as an extension, with a stochastic counterpart, of the recent findings in [3]. The paper is organized as follows. In Section 2 we describe the mathematical model, obtain the main theoretical results of the paper, and in Section 3 we estimate the parameters involved in the model, compare with the numerical data, and interpret the results. In the Appendix we provide the proofs of the main results.

2 The model and main theoretical results

According to [1], at each time $t > 0$, the level of acquired knowledge is given by the integral (or sum, if time is discrete) of the quantities of learnt material prior to time t. We shall work with continuous time t, and the term "integral" -the key point in this paper- will be explained shortly.

The individual's goal is to reach a level of knowledge $b > 0$, as a result of the following learning model. On a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$, let W_t denote the level of acquired knowledge and π_t the quantity of learnt material at time $t \geq 0$. Obviously $\pi_0 = 0$ and we assume that the strategy of learning, i.e., the collection of all quantities of learnt material, denoted by $\Pi = \{\pi_t\}_{t\geq 0}$ is an $\{\mathcal{F}_t\}_{t\geq 0}$-progressively measurable stochastic process, satisfying $\displaystyle\int_0^t \pi_s^2 ds < \infty$ P-almost surely for all $t \geq 0$. We also assume that the individual has an initial level of knowledge $W_0 = w_0 \geq 0$, whereas the level of acquired knowledge W_t at time $t > 0$ follows the dynamics of a diffusion process, i.e.,

$$W_t = w_0 + \varepsilon \int_0^t \pi_s ds + (1-\varepsilon)\int_0^t \pi_s dB_s. \tag{2.1}$$

Here $B = \{B_t\}_{t\geq 0}$ is a standard Brownian motion process on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$; the first integral in (2.1) is Lebesgue-type (or deterministic), whereas the second one is Itô-type (a.k.a. stochastic integral, cf. [4]). For a fixed $0 < \varepsilon \leq 1$, the values ε and $1-\varepsilon$ weigh the deterministic, respectively random (or stochastic), quantities of learnt material in the model. For $\varepsilon = 1$ we obtain a purely deterministic model, whereas for $0 < \varepsilon < 1$ we are dealing with mixed deterministic-stochastic model.

The individual has a limited time t_d to acquire knowledge; we assume that the random variable t_d follows an exponential distribution with mean $1/\lambda$. The individual seeks to maximize the probability that the level of knowledge is at least b by the time t_d, i.e., $W_{t_d} \geq b$, by optimizing over all strategies of learning Π. To avoid "bad strategies of learning", we stop the learning process when the level of knowledge reaches 0 before time t_d. Define $t_0 = \inf\{t \geq 0 : W_t \leq 0\}$, and define the maximum probability of reaching the level of knowledge b by

$$p(w) := \sup P^{w_0}\{W_{t_d \wedge t_0}(\omega) \geq b\}, \tag{2.2}$$

in which P^{w_0} denotes conditional probability given $W_0 = w_0 \geq 0$, and sup is calculated upon all strategies of learning Π.

We are looking to obtain an explicit expression for $p(\omega)$, and find the optimal strategy of learning, say $\Pi^* = \{\pi^*\}$, where $\pi^* = \pi^*_t$ is the optimal quantity of learnt material at the time t when the level of knowledge reaches a level $w \in (0, b)$. The next result answers these requirements.

Theorem 1. *Assume the above assumptions and hypotheses, and let* $0 < \varepsilon < 1$.

(*i*) *The maximum probability of reaching the level of knowledge* b *is given by*

$$p(w) = (w/b)^{\lambda/[\lambda+\varepsilon^2/2(1-\varepsilon)^2]} \text{ for } 0 \leq w \leq b. \tag{2.3}$$

(*ii*) *At the time when the level of knowledge reaches a level* $w \in (0, b)$*, the optimal quantity of learnt material is given by*

$$\pi^*(w) = w[2\lambda + \varepsilon^2/(1-\varepsilon)^2]/\varepsilon. \tag{2.4}$$

Remarks 2. The probability p in equation (2.3) is a power function of ω, and increases as ω becomes closer to the a priori level of knowledge b, meaning that: the larger the quantity of learnt material, the more probable is to acquire a higher level of knowledge.

For fixed $0 < \varepsilon < 1$, the exponent of the power in equation (2.3) increases with λ, hence the probability of reaching the level of knowledge b decreases with λ, meaning that: as the individual becomes more likely to finish learning sooner than later, reaching the level of knowledge b becomes less likely.

For fixed $\lambda > 0$, the exponent of the power in equation (2.3) decreases with ε, hence the probability of reaching the level of knowledge b increases with ε, meaning that: as the model becomes "more deterministic" (that is, for larger values of ε), the probability of reaching the level of knowledge b increases.

As ε approaches 0 (that is, the model becomes more and more stochastic), the optimal quantity of learnt material in equation (2.4) increases, meaning that: in a stochastic environment, the individual needs to work more towards achieving the same goal as in the deterministic model.

Corollary 3. *Under the assumptions and hypotheses in Theorem 1, corresponding to the optimal strategy of learning* Π^**, the optimal level of acquired knowledge, say* W^*_t*, follows the dynamics of a Geometric Brownian motion process, i.e.,*

$$W^*_t = w + \frac{\varepsilon^4}{(1-\varepsilon)^2(\varepsilon^2 + 2\lambda(1-\varepsilon)^2)} \int_0^t W^*_s ds + \frac{\varepsilon^2 + 2\lambda(1-\varepsilon)^2}{\lambda\varepsilon} \int_0^t W^*_s dB_s, \text{ for } t > 0. \tag{2.5}$$

Remark 4. From formula (2.5) it follows that $W_t^* > 0$ P-almost surely for all $t \geq 0$, provided that $W_0^* = W_0 = w \in (0, b)$. As ε approaches 1, the optimal level of acquired knowledge W_t^* has a dominant deterministic component, and when ε approaches 0, the stochastic component becomes dominant (in accordance with the behavior of the level of acquired knowledge W_t).

3 Data Analysis and Interpretations

We shall use a hat ˆ to denote the estimated parameters of the model. Let us start with the midterm. The time to acquire knowledge follows an exponential distribution with parameter $1/\lambda$, whose maximum likelihood estimator is given by $\hat{\lambda} = 2$. The parameter ε of the stochastic process within the model (2.1) is estimated directly: $\hat{\varepsilon} = 312/420 \approx 0.742857$. Because the midterm worth 30% of the final grade, we have $\omega = b/3$. Using formula (2.3) above, we obtain that the maximum probability of reaching the level of knowledge required by the midterm equals

$$\hat{p}_{\text{midterm}} = (0.33)^{2/[2+\hat{\varepsilon}^2/2(1-\hat{\varepsilon})^2]} = 72.17\%,$$

which gives just a slight underestimate (27.83%) of the midterm failure rate (30.48%).

In the case of the final exam, we have $\hat{\lambda} = 1$, $\omega = b/2$, whereas the other estimated values do not change. We obtain that the maximum probability of reaching the level of knowledge required by the final exam equals

$$\hat{p}_{\text{final exam}} = (0.5)^{1/[1+\hat{\varepsilon}^2/2(1-\hat{\varepsilon})^2]} = 86.41\%,$$

which gives just a slight underestimate (13.59%) of the final exam failure rate (14.76%).

As for assignments, we have $\hat{\lambda} = 1$, $\omega = b/5$, hence

$$\hat{p}_{\text{assignments}} = (0.4)^{1/[1+\hat{\varepsilon}^2/2(1-\hat{\varepsilon})^2]}) = 78.3\%,$$

a slight underestimate of the assignments range (80%).

According to formula (2.4), the optimal percentages of learnt material for midterm, final exam and assignments are given, respectively, by

$$\pi^*_{\text{midterm}} = [4 + \hat{\varepsilon}^2/(1-\hat{\varepsilon})^2]/3\hat{\varepsilon} = 54.8\%;$$

$$\pi^*_{\text{final exam}} = [2 + \hat{\varepsilon}^2/(1-\hat{\varepsilon})^2]/2\hat{\varepsilon} = 70.8\%;$$

$$\pi^*_{\text{assignments}} = [2 + \hat{\varepsilon}^2/(1-\hat{\varepsilon})^2]/5\hat{\varepsilon} = 27.9\%.$$

The excess of 53.5% (= 54.8% + 70.8% + 27.9% − 100%) comes from overlapping and consolidation of knowledge from assignments, midterm and final exam (these latter categories are not independent!).

Finally, from formula (2.5), in the case of the final exam, we obtain the following estimates for the coefficients therein:

$$\frac{\hat{\varepsilon}^4}{(1-\hat{\varepsilon})^2[\hat{\varepsilon}^2+2(1-\hat{\varepsilon})^2]} = 5.64 \text{ and } \frac{\hat{\varepsilon}^2+2(1-\hat{\varepsilon})^2}{\hat{\varepsilon}} = 0.55,$$

showing that the optimal level of knowledge is in the ratio $5.64 : 0.55 \approx 10 : 1$ between deterministic and random components.

The results obtained in this section confirm that, in spite of poor odds, students with incomplete pre-requisites obtained better results than expected, by working harder on the midterm and assignments, and thus improving the final grade. Thus, our model correctly explains the small failure rates at that course. Note that the errors committed in this section are within the acceptable limits, see the guidelines in [2].

4 Appendix

For the proof of Theorem 1, one needs to produce a classical (smooth) strong solution of a boundary value problem. More precisely, for $\alpha \in \mathbb{R}$, let us define the differential operator L^α by its action on test functions f, as follows:

$$L^\alpha f = \varepsilon\alpha f' + \frac{1}{2}(1-\varepsilon)^2\alpha^2 f'' - \lambda(f - \mathbf{1}_{\{w\geq b\}}).$$

According to [4], Chapter 6, if we find a non-decreasing, concave function $f = f(w)$ of class C^2 on $[0, b]$ satisfying the boundary value problem

$$\max_{\alpha\in\mathbb{R}} L^\alpha f(w) = 0; f(0) = 0, f(b) = 1, \tag{4.1}$$

then f is the maximum probability of reaching the level of acquired knowledge b, and the optimal quantity of learnt material is given by

$$\pi_t^* = -\frac{\varepsilon f'(W_t^*)}{(1-\varepsilon)^2 f''(W_t^*)},$$

for all $t \in [0, t_d \wedge t_0)$, in which W_t^* is the optimal level of knowledge.

One can see that the function p in formula (2.3) satisfies the above requirements for the boundary value problem (4.1). Indeed, $p(0) = 0, p(b) = 1$ and, for $0 < \omega < b$, we have

$$L^\alpha p(\omega) = -K\left(\alpha - \frac{2[\lambda + \varepsilon^2/(1-\varepsilon)^2]}{\varepsilon}\omega\right)^2, \tag{4.2}$$

where $K > 0$ depends on $\varepsilon, \omega, \lambda$ and b. The maximum value of the right-hand side of (4.2) equals 0, and is achieved for $\alpha = 2[\lambda + \varepsilon^2/(1-\varepsilon)^2]\omega/\varepsilon$.

Corollary 3 follows by applying Itô's formula (cf. [4], Chapter 5) in (2.1) and using π^* from formula (2.4).

Acknowledgment. The authors would like to thank an anonymous Referee and the Area Editor for valuable comments and suggestions that improved the final form of the manuscript.

References

[1] E. Dayan, B.B. Averbeck, B.J. Richmond and L.G. Cohen, *Stochastic reinforcement benefits skill acquisition*. Learning and Memory **21** (2014), 140-142.

[2] N. Gilbert and K.G. Troitzsch, *Simulation for the Social Scientist*, Open University Press, 2005.

[3] M. Iancu and M. Kohlhase, *Math literate knowledge management via induced material*. Lecture Notes in Computer Science **9150** (2015), 187-202. Zbl 06512418.

[4] F.C. Klebaner, *Introduction to Stochastic Calculus with Applications*. Imperial College Press, 2005. MR2160228. Zbl 1077.60001.

[5] A.S. Reber, *Implicit learning and tacit knowledge*. Journal of Experimental Psychology: General **118** (1989), 219-235.

[6] T.C. Silva, L. Zhao, *Detecting and preventing error propagation via competitive learning*. Neural Networks **41** (2013), 70-84. Zbl 1296.68144.

George Stoica
Horizon Health Network, Research Services,
Saint John Regional Hospital,
400 University Ave., Saint John NB, E2L 4L2, Canada.
e-mail: George.Stoica@HorizonNB.ca

Barry Strack
Horizon Health Network, Research Services,
Saint John Regional Hospital,
400 University Ave., Saint John NB, E2L 4L2, Canada.
e-mail: barry.strack@horizonnb.ca

License