Proceedings of the 2016 Winter Simulation Conference
T. M. K. Roeder, P. I. Frazier, R. Szechtman, E. Zhou, T. Huschka, and S. E. Chick, eds.

APPROXIMATE BAYESIAN INFERENCE AS A FORM OF STOCHASTIC APPROXIMATION: A NEW CONSISTENCY THEORY WITH APPLICATIONS

Ye Chen

Department of Mathematics
University of Maryland
College Park, MD 20742, USA

Ilya O. Ryzhov

Robert H. Smith School of Business
University of Maryland
College Park, MD 20742, USA

ABSTRACT

Approximate Bayesian inference is a powerful methodology for constructing computationally efficient statistical learning mechanisms in problems where incomplete information is collected sequentially. Approximate Bayesian models have been developed and applied in a variety of different domains; however, this work has thus far been primarily computational, and convergence or consistency results for approximate Bayesian estimators are largely unavailable. We develop a new consistency theory for these learning schemes by interpreting them as stochastic approximation (SA) algorithms with additional "bias" terms. We prove the convergence of a general SA algorithm of this type, and apply this result to demonstrate, for the first time, the consistency of several approximate Bayesian methods from the recent literature.

1 INTRODUCTION

We consider a class of sequential learning problems where incomplete or censored information is used to maintain and update beliefs about one or more unknown population parameters. Such problems arise in numerous application areas; three specific recently-studied applications include:

- In financial markets, market-makers are used to increase liquidity and promote trading after a shock event. The market-maker's optimal trading strategy seeks to learn the unknown value of assets based on traders' willingness to buy and sell (Das and Magdon-Ismail 2009).
- Suppose that buyer valuations of a product are drawn independently from an unknown distribution that depends on the price. The seller wishes to learn this distribution, but cannot observe valuations directly. Instead, they must be inferred based on sales (Qu, Ryzhov, and Fu 2013).
- Online gaming services attempt to match players of similar skill levels to promote competitive play. However, "skill" is a modeling construct that cannot be observed directly, but rather must be estimated from win/loss outcomes (Dangauthier et al. 2007).

Sequential learning is particularly important when it is coupled with multi-stage optimization; for example, in dynamic pricing, the seller earns revenue from a sequence of pricing decisions, which can be improved in an adaptive manner using the incoming sales information. In the simulation community, learning is widely studied in the context of ranking and selection (see the tutorials by Hong and Nelson 2009 or Chau et al. 2014), in which a decision-maker seeks to discover the best of a finite set of alternatives by adaptively allocating a simulation budget. When decisions are made sequentially, Bayesian statistical models can be useful (Chick 2006, Chen et al. 2015) for quantifying the decision-maker's remaining uncertainty about the unknown values at every time stage; this uncertainty can then be traded off against the estimated values of different decisions to determine an optimal balance between exploration and exploitation. See Powell and Ryzhov (2012) for an introduction to the main algorithmic ideas in this literature.

978-1-5090-4486-3/16/$31.00 ©2016 IEEE

For the most part, the existing work on Bayesian optimal learning prefers to use simple statistical models in which unbiased observations of the unknown values are available (e.g., from simulation models). The most common framework (see, e.g., Gupta and Miescke 1996) assumes that the simulation output is normally distributed and centered around the unknown quantity to be learned; a simple normal prior can then be used to model our beliefs about the unknown mean. The property of *conjugacy* (DeGroot 1970) ensures that our posterior beliefs continue to be normally distributed after any number of observations. Consequently, the beliefs can always be completely characterized by a small set of parameters (e.g., means and variances for normal distributions), which can be updated recursively in closed form.

In traditional Bayesian statistics, where inference is performed once on a given sample, conjugacy may be less important and approximate Bayesian estimation (Sunnåker et al. 2013) may be based on Markov chain Monte Carlo procedures (Plagnol and Tavaré 2004), which are computationally expensive but known to converge (Asmussen and Glynn 2011). However, conjugacy becomes extremely valuable when samples are observed sequentially and information collection is guided adaptively by the decision-maker. In such cases, the ability to compactly represent beliefs using a small set of parameters enables the development of optimization algorithms that take those parameters as inputs and return recommended decisions. These algorithms may be required to run very quickly; for example, in dynamic pricing, it may be necessary to calculate and post a price as soon as the next buyer logs on to the seller's website. Conjugate learning models greatly simplify the design of such procedures.

However, in many applications, including those listed above, there is no natural choice of prior distribution that is conjugate with the observations. This issue has led to recent interest in *approximate* Bayesian methods, which force conjugacy by creating an artificial posterior distribution from the same family as the prior (e.g., normal), and choosing the parameters of that distribution to "optimally approximate" (in some sense) the exact posterior, which is typically much less tractable. Computational strategies for these methods include moment-matching (Zhang and Song 2015), density filtering (Qu et al. 2015), and variational approximations (Jaakkola and Jordan 2000). However, although these approaches have repeatedly demonstrated significant practical benefits, they remain largely unamenable to the usual forms of theoretical analysis. Ryzhov (2015) discusses some of the theoretical challenges involved; here, we can simply state that even the statistical consistency of approximate Bayesian estimators has heretofore been a largely open problem. In fact, it is not even clear whether we should expect them to be consistent, as each stage of sampling introduces a new approximation and thus more error into the Bayesian model.

In this paper, we summarize a new analysis by Chen and Ryzhov (2016) that sheds light on these issues for the first time. To provide context, we first describe in Section 2 a simple example where approximate Bayesian inference is used to create approximately normal posteriors based on censored binary observations. We observe that the updating equations derived for this learning scheme can be interpreted as a form of stochastic approximation (Robbins and Siegmund 1985, Kushner and Yin 2003, Borkar 2008), and more specifically bears resemblance to online gradient descent procedures (Bottou 1998) with the addition of a "bias" term. In Section 3, we propose a modified Robbins-Monro algorithm with a similar bias term, and prove its convergence. In Section 4, we explain how these results can be applied to prove the asymptotic consistency, not only of the example procedure from Section 2, but also of several other learning schemes from the recent literature developed for *all* of the above-listed applications. Section 5 concludes and points out directions for future work. We believe that our work is the first to offer general theoretical machinery for consistency analysis of approximate Bayesian schemes; our results complement the existing algorithmic literature and lend support to approximate Bayesian inference as a rigorous statistical learning methodology.

2 MOTIVATING EXAMPLE: LEARNING FROM CENSORED BINARY OBSERVATIONS

Before we begin our theoretical analysis, we present a simple motivating example where approximate Bayesian inference is used to construct a computationally tractable estimator of an unknown population mean in a setting where only censored binary observations of that mean are available. This setting provides context for our main insight that approximate Bayesian inference can be viewed as a form of stochastic approximation.

Let θ be an unknown value and denote by $(Y_n)_{n=1}^{\infty}$ a sequence of i.i.d. samples from the common distribution $\mathcal{N}\left(\theta,\lambda^2\right)$. The variance λ^2 is assumed to be known for simplicity. We impose the Bayesian model $\theta \sim \mathcal{N}\left(\mu_0,\sigma_0^2\right)$ where the parameters μ_0,σ_0 represent our prior beliefs about θ.

If the samples $Y_1,Y_2,...$ can be observed directly, the posterior distribution of θ given $Y_1,...,Y_n$ is known to be normal for any n (DeGroot 1970). In that case, our beliefs about θ can always be compactly characterized by a pair (μ_n,σ_n), which can be easily updated recursively after each successive sample. The consistency of the estimator μ_n follows trivially, as the updating equation for the mean is essentially a form of recursive sample averaging.

Suppose, however, that we cannot observe $Y_1,Y_2,...$ directly. Instead, we can only observe a sequence $(B_n)_{n=1}^{\infty}$ of censored observations defined by

$$B_{n+1} = 1_{\{Y_{n+1}<b_n\}},$$

where the sequence $(b_n)_{n=0}^{\infty}$ is deterministic and known; for example, b_n could be a dosage decision for a drug, with B_{n+1} indicating the presence or absence of adverse effects. It is readily evident that the posterior density $P(\theta \in dx\,|\,B_1)$ is no longer normal, even after just one observation. As we collect more samples, the posterior will become a more complicated mixture, making it difficult to compactly represent and update our beliefs.

We will now apply approximate Bayesian inference (Ryzhov 2015) to create a computationally efficient learning mechanism for this problem. We will "force" the posterior to be conjugate by creating an artificial density from the desired distributional family (in this case, a normal density) and choosing the parameters of that density to "optimally approximate" the exact, non-normal posterior. We will then discard that posterior and proceed under the assumption that our beliefs are accurately modeled by the artificial normal distribution. In this way we regain the ability to represent our beliefs using just two parameters (mean and variance), but presumably incur statistical error from the approximation.

In this example, we create the posterior using the well-known method of moment-matching, also known as expectation propagation (Minka 2001). Suppose that our initial Bayesian assumption $\theta \sim \mathcal{N}\left(\mu_0,\sigma_0^2\right)$ holds, and that B_1 is given. Let $\tilde{\theta}$ be a random variable following a normal distribution with mean μ_1 and variance σ_1^2, where these parameters are chosen to satisfy the equations

$$\begin{aligned}
\int_{\mathbb{R}} xP\left(\tilde{\theta} \in dx\right) &= \int_{\mathbb{R}} xP(\theta \in dx\,|\,B_1),\\
\int_{\mathbb{R}} x^2P\left(\tilde{\theta} \in dx\right) &= \int_{\mathbb{R}} x^2P(\theta \in dx\,|\,B_1).
\end{aligned}$$

In other words, the first two moments of $\tilde{\theta}$ are equal to those of the non-normal posterior distribution. We then move to the next stage of sampling and repeat the process with the assumption that $\theta \sim \mathcal{N}\left(\mu_1,\sigma_1^2\right)$ and the next observation B_2.

With some algebra, it can be shown that the moment-matching equations in the $(n+1)$st stage admit a closed-form solution given by

$$\mu_{n+1} = \mu_n - \sigma_n^2\left(B_{n+1}\frac{1}{\sqrt{\lambda^2+\sigma_n^2}}\frac{\phi(p_n)}{\Phi(p_n)} - (1-B_{n+1})\frac{1}{\sqrt{\lambda^2+\sigma_n^2}}\frac{\phi(p_n)}{1-\Phi(p_n)}\right), \tag{1}$$

$$\begin{aligned}
\sigma_{n+1}^2 = \sigma_n^2\Bigg(1 &- B_{n+1}\frac{\sigma_n^2}{\lambda^2+\sigma_n^2}\frac{p_n\phi(p_n)\Phi(p_n)+\phi^2(p_n)}{\Phi^2(p_n)}\\
&-(1-B_{n+1})\frac{\sigma_n^2}{\lambda^2+\sigma_n^2}\frac{\phi^2(p_n)-p_n\phi(p_n)(1-\Phi(p_n))}{(1-\Phi(p_n))^2}\Bigg),
\end{aligned} \tag{2}$$

where ϕ, Φ are the standard normal pdf and cdf, and

$$p_n = \frac{b_n - \mu_n}{\sqrt{\lambda^2 + \sigma_n^2}}.$$

The new belief parameters $(\mu_{n+1}, \sigma_{n+1})$ can be efficiently computed in a recursive manner from the old parameters (μ_n, σ_n) and the next observation B_{n+1}.

It is not obvious whether the sequence (μ_n) of posterior means will be able to estimate θ accurately. In fact there is some reason to expect that this will *not* happen: first, the censored observations (B_n) themselves now carry much less information than the complete observations (Y_n), and second, the approximate Bayesian update introduces statistical error that may be compounded over time (since we always advance to the next stage under the normality assumption). It is thus somewhat surprising that (μ_n) is, in fact, statistically consistent, i.e., it is guaranteed to recover θ w.p. 1.

We will describe a rigorous framework for proving this result in the next section. First, however, we provide additional intuition by pointing out that (1) can be viewed as a Robbins-Monro stochastic approximation (SA) procedure of the form

$$\mu_{n+1} = \mu_n - \alpha_n G_n(B_{n+1}, \mu_n), \tag{3}$$

with the posterior variance σ_n^2 serving as the stepsize α_n. SA has previously been applied to frequentist statistical estimation under the name "online gradient descent" or OGD (Bottou 1998). This version of SA assumes that the unknown mean θ is fixed and optimizes the log-likelihood of the observed samples. In the context of our example, the OGD algorithm is given by (3) with

$$G_n(B_{n+1}, \mu_n) = B_{n+1} \frac{1}{\lambda} \frac{\phi(q_n)}{\Phi(q_n)} - (1 - B_{n+1}) \frac{1}{\lambda} \frac{\phi(q_n)}{1 - \Phi(q_n)}, \tag{4}$$

where $q_n = \frac{b_n - \mu_n}{\lambda}$ and the only assumptions made on the stepsize α_n are the conditions

$$\sum_{n=0}^{\infty} \alpha_n = \infty, \qquad \sum_{n=0}^{\infty} \alpha_n^2 < \infty, \tag{5}$$

which are usually required for convergence in SA theory (Robbins and Monro 1951).

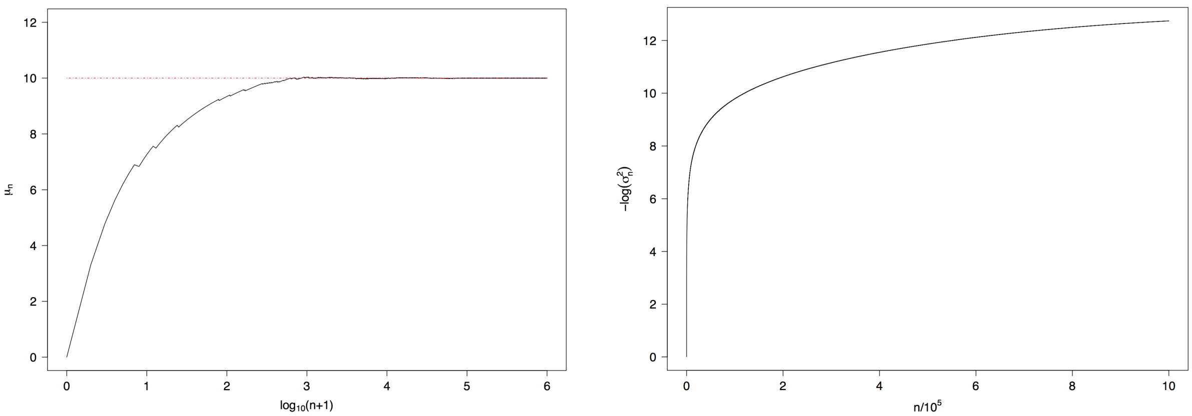
It is easy to see that (4) is the gradient of the marginal *frequentist* log-likelihood function of B_{n+1} evaluated at the current iterate μ_n. The approximate Bayesian update (1) also greatly resembles the OGD algorithm, with the posterior variance σ_n^2 playing two roles: first, it augments the noise λ^2 in the definition of G_{n+1}, and second, it replaces the stepsize. Thus, if σ_n^2 satisfies (5), and if the moment-matching update does not deviate "too far" from the OGD update, we may also expect (1) to converge. This approach is formalized in the next section.

We end this example with a numerical illustration. Figure 1(a) shows the sequence μ_n produced by (1)-(2) over 10^6 iterations. We set $\lambda^2 = 1.5$, $\mu_0 = 0$, $\sigma_0^2 = 1$, and the sequence $b_n = 8 + 0.000003 \cdot n$. The true value of the parameter is set to $\theta = 10$. Convergence is observed after just 1000 iterations. Figure 1(b) shows the trajectory of the approximate posterior variance; we see that σ_n^2 can be viewed as a kind of adaptive stepsize, whose declining behaviour slows down in later iterations in order to place sufficient weight on new information.

3 A GENERAL CONVERGENT STOCHASTIC APPROXIMATION ALGORITHM

In this section, we propose a general stochastic approximation algorithm of the form

$$x_{n+1} = x_n - \alpha_n (F_n(W_{n+1}, x_n) + \beta_n(W_{n+1}, x_n, \alpha_n)), \ n = 0, 1, ... \tag{6}$$



(a) Approximate posterior mean. (b) Approximate posterior variance (log-scale).

Figure 1: Empirical convergence of the approximate Bayesian estimator.

where $x_0 \in \mathbb{R}^m$ is an arbitrary m-vector, the step size $(\alpha_n)_{n=0}^{\infty}$ is a positive (deterministic or random) sequence satisfying (5) almost surely, $(W_n)_{n=1}^{\infty}$ is a sequence of random variables, $(F_n)_{n=0}^{\infty}$ is a sequence of real measurable functions mapping (w,x) into $\mathbb{R}^m$, and $(\beta_n)_{n=0}^{\infty}$ is another sequence of real measurable functions representing the "bias" of the SA update.

The algorithm in (6) closely resembles the SA procedures studied in Kushner and Yin (2003) and other references. However, our intended applications require us to introduce modifications into the procedure that necessitate a new convergence proof. The main difference between (6) and standard SA is the introduction of the bias β_n. Connecting (6) back to our example in Section 2, the SA update F_n would be identical to the OGD gradient G_n in (4), while the bias β_n would be set equal to the difference between the OGD gradient and the approximate Bayesian update in (1). Since the posterior variance σ_n^2 serves as the stepsize in (1), the bias β_n will thus depend on the random variable α_n. This dependence does not fit into the standard SA convergence conditions (e.g., those in Sec. 5.2 in Kushner and Yin 2003) and thus a new set of conditions is required. These are described below.

We define

$$\begin{aligned}\mathscr{F}_n &\triangleq \mathscr{B}\left(W_1,\ldots,W_n,x_1,\ldots,x_n,\alpha_1,\ldots,\alpha_n\right),\\ R_n\left(x_n\right) &\triangleq \mathbf{E}\left(F_n\left(W_{n+1},x_n\right)\mid\mathscr{F}_n\right),\end{aligned}$$

and impose several conditions as follows. The first condition ensures that the SA algorithm is searching for a unique root θ:

Assumption 1 For any n, the equation $R_n(x)=0$ has a unique root θ, which does not depend on n.

In the example from Section 2, this quantity corresponds to the unknown population mean, with the minor distinction that here θ is treated as a fixed value (as in frequentist statistics). The second condition resembles many standard SA convergence proofs, where the expected value of the stochastic gradient is assumed to point in the "right" direction:

Assumption 2 (Convexity) For $n=1,2,\ldots$ and any $\varepsilon>0$,

$$\inf_{\|x-\theta\|_2^2>\varepsilon,\, n\in\mathbb{N}} (x-\theta)^T R_n(x) > 0.$$

The third condition bounds the growth of the second moments of F_n and β_n:

Assumption 3 There exist positive constants C_1 and C_2 such that

$$\sup_{n\in\mathbb{N}} \mathbf{E}\left(\|F_n(W_{n+1},x)\|_2^2|\mathscr{F}_n\right) \leq C_1\left(1+\|x-\theta\|_2^2\right), \tag{7}$$

$$\sup_{n\in\mathbb{N}} \mathbf{E}\left(\|\beta_n\left(W_{n+1},x,\alpha_n\right)\|_2^2|\mathscr{F}_n\right)/\alpha_n^2 \leq C_2\left(1+\|x-\theta\|_2^2\right) \tag{8}$$

for all x.

Equation (7) controls the amount of noise in the SA update. Equation (8) ensures that the bias of the update (recall that this is analogous to the difference between the frequentist OGD and approximate Bayesian updates) is "small".

We now state our main results about algorithm (6); the full details can be found in Chen and Ryzhov (2016). Theorems 1 and 2 essentially state the same result in two ways, with the second version using an explicit projection operator to ensure the boundedness of the iterates (a widely-used approach in the application of SA convergence theory).

Theorem 1 Suppose that Assumptions 1-3 hold and (α_n) satisfies (5) almost surely. Let x_n be defined by (6). Then $x_n \to \theta$ almost surely.

Theorem 2 Suppose that Assumptions 1-3 hold and (α_n) satisfies (5) almost surely. Define

$$x_{n+1} = \Pi_H\left(x_n - \alpha_n\left(F_n\left(W_{n+1},x_n\right)+\beta_n\left(W_{n+1},x_n,\alpha_n\right)\right)\right),\ n=0,1,... \tag{9}$$

where $\Pi_H : \mathbb{R}^m \to H$ is a projection operator and H is a compact subset of $\mathbb{R}^m$ taken to be large enough such that $x_0,\theta \in H$. Then, $x_n \to \theta$ almost surely.

In the following section, we apply these results to prove the convergence of the example from Section 2 as well as several well-known approximate Bayesian schemes from different backgrounds.

4 APPLICATIONS OF THE CONVERGENCE THEORY

We now present several applications of our convergence analysis, based on recent algorithmic work in approximate Bayesian inference. Section 4.1 states the convergence of the approximate Bayesian scheme derived for the example in Section 2. Section 4.2 shows the consistency of a learning scheme developed for learning player skills in competitive online gaming. Section 4.3 applies our technique to show the consistency of a statistical procedure for learning the market value of an asset. In Section 4.4, we show similar results for the problem of learning buyer valuations in online auctions.

4.1 Learning An Unknown Mean From Censored Binary Observations

Recall, from Section 2, the problem of learning an unknown mean θ based on censored binary observations. Consider a version of this algorithm where the posterior variance σ_n^2 is updated using (2), and the posterior mean μ_n is updated using a modification of (1), given by

$$\mu_{n+1} = \Pi_H\left(\mu_n - \sigma_n^2\left(B_{n+1}\frac{1}{\sqrt{\lambda^2+\sigma_n^2}}\frac{\phi\left(p_n\right)}{\Phi\left(p_n\right)} - \left(1-B_{n+1}\right)\frac{1}{\sqrt{\lambda^2+\sigma_n^2}}\frac{\phi\left(p_n\right)}{1-\Phi\left(p_n\right)}\right)\right), \tag{10}$$

where $H=[-M,M]$ for large enough M satisfying $|\mu_0|<M$ and $|\theta|<M$. Thus,

$$\Pi_H\left(x\right) = x\cdot 1_{\{|x|\leq M\}} + M\cdot 1_{\{x>M\}} - M\cdot 1_{\{x<-M\}}.$$

Theorem 2 can then be applied to establish consistency. The projection operator serves to ensure the boundedness of the sequence (μ_n), which is used to prove that the stepsize σ_n^2 satisfies (5) almost surely.

Proposition 1 Suppose that μ_n and σ_n^2 are updated using (10) and (2), and suppose that the sequence $(b_n)_{n=0}^{\infty}$ is bounded. Then, $\mu_n \to \theta$ almost surely.

4.2 Learning Player Skills In Competitive Online Gaming

Herbrich, Minka, and Graepel (2006) and Dangauthier et al. (2007) describe a computational learning scheme, based on approximate Bayesian inference, that was implemented in Microsoft's Xbox Live online gaming service for learning player skills in competitive events. In this application, large numbers of players log on to the service and ask to play a game; the system then seeks to match players whose skill levels are likely to be similar, in order to promote fair play and create a more rewarding experience.

We give a streamlined summary of the model, assuming without loss of generality that there are only two players. Let $\theta^{(i)}$ represent the "skill" of player $i \in \{1,2\}$. Denote by $Y_n^{(i)}$ the "performance" of player i in the nth game, with the assumption that

$$Y_{n+1}^{(i)} \sim \mathcal{N}\left(\theta^{(i)}, \lambda^2\right),$$

where the variance λ^2 is known. We further impose the Bayesian modeling assumption

$$\theta^{(i)} \sim \mathcal{N}\left(\mu_0^{(i)}, \left(\sigma_0^{(i)}\right)^2\right).$$

Finally, we assume that all skills and performance values are mutually independent.

The game master does not have the ability to observe $Y_n^{(i)}$ directly, but rather must infer the unknown mean from the binary outcome of the competition, given by

$$B_{n+1}^{(i)} = 1_{\left\{Y_{n+1}^{(i)} < Y_{n+1}^{(j)}\right\}},$$

where j denotes the index of the opponent. In words, if player j wins the match against i, we interpret this as $Y_{n+1}^{(j)} > Y_{n+1}^{(i)}$. We suppose that draws cannot happen in the game in question.

As in Section 2, moment-matching can be applied to derive the approximate Bayesian updating equations

$$\begin{aligned}
\mu_{n+1}^{(i)} &= \mu_n^{(i)} - \left(\sigma_n^{(i)}\right)^2 \left(B_{n+1}^{(i)} \frac{1}{\sqrt{\left(\sigma_n^{(i)}\right)^2 + \left(\sigma_n^{(j)}\right)^2 + 2\lambda^2}} v\left(\frac{\mu_n^{(j)} - \mu_n^{(i)}}{\sqrt{\left(\sigma_n^{(i)}\right)^2 + \left(\sigma_n^{(j)}\right)^2 + 2\lambda^2}} \right) \right. \\
&\quad \left. - \left(1 - B_{n+1}^{(i)}\right) \frac{1}{\sqrt{\left(\sigma_n^{(i)}\right)^2 + \left(\sigma_n^{(j)}\right)^2 + 2\lambda^2}} v\left(\frac{\mu_n^{(i)} - \mu_n^{(j)}}{\sqrt{\left(\sigma_n^{(i)}\right)^2 + \left(\sigma_n^{(j)}\right)^2 + 2\lambda^2}} \right) \right), \qquad (11)
\end{aligned}$$

$$\begin{aligned}
\left(\sigma_{n+1}^{(i)}\right)^2 &= \left(\sigma_n^{(i)}\right)^2 \left(1 - B_{n+1}^{(i)} \frac{\left(\sigma_n^{(i)}\right)^2}{\left(\sigma_n^{(i)}\right)^2 + \left(\sigma_n^{(j)}\right)^2 + 2\lambda^2} w\left(\frac{\mu_n^{(j)} - \mu_n^{(i)}}{\sqrt{\left(\sigma_n^{(i)}\right)^2 + \left(\sigma_n^{(j)}\right)^2 + 2\lambda^2}} \right) \right. \\
&\quad \left. - \left(1 - B_{n+1}^{(i)}\right) \frac{\left(\sigma_n^{(i)}\right)^2}{\left(\sigma_n^{(i)}\right)^2 + \left(\sigma_n^{(j)}\right)^2 + 2\lambda^2} w\left(\frac{\mu_n^{(i)} - \mu_n^{(j)}}{\sqrt{\left(\sigma_n^{(i)}\right)^2 + \left(\sigma_n^{(j)}\right)^2 + 2\lambda^2}} \right) \right), \qquad (12)
\end{aligned}$$

where

$$\begin{aligned}
v(x) &= \frac{\phi(x)}{\Phi(x)}, \\
w(x) &= v(x)(v(x) + x).
\end{aligned}$$

As in Section 4.1, we can replace (11) by a projected version where $\left(\mu_n^{(i)}\right)$ is constrained to be within a suitably large interval.

Define

$$\begin{aligned} d_n &\triangleq \mu_n^{(1)} - \mu_n^{(2)}, \\ \delta &\triangleq \theta^{(1)} - \theta^{(2)}. \end{aligned}$$

In this setting, we do not observe sufficient information to learn $\theta^{(i)}$ exactly. However, we can learn δ, which is of primary interest to the game master.

Proposition 2 Suppose that $\mu_n^{(i)}$ is updated using a projected version of (11), while $\sigma_n^{(i)}$ is updated using (12). Then, $d_n \to \delta$ almost surely.

4.3 Learning The Market Value Of An Asset

Das and Magdon-Ismail (2009) presents a model for learning the unknown value θ of an asset after a market shock. In this setting, a market-maker is used to encourage trading and increase liquidity. The asset is traded sequentially, one unit at a time. The sequence $(Y_n)_{n=1}^{\infty}$ denotes the traders' perceptions of the unknown value, which are assumed to be independently drawn from the common distribution $\mathcal{N}\left(\theta, \lambda^2\right)$ with λ^2 known.

The market-maker's initial beliefs are represented by the prior distribution $\theta \sim \mathcal{N}\left(\mu_0, \sigma_0^2\right)$. Let $(b_n)_{n=0}^{\infty}$ and $(a_n)_{n=0}^{\infty}$ denote deterministic, known sequences of "bid prices" and "ask prices" set by the market-maker. If $Y_{n+1} < a_n$, the next trader buys one unit of the asset from the market-maker; if $a_n \leq Y_{n+1} \leq b_n$, the trader does not make any transaction; and, if $Y_{n+1} > b_n$, the trader sells one unit of the asset to the market-maker. In this way, the market-maker cannot observe Y_{n+1} directly, but only knows a range into which this value falls based on the trader's action. Let

$$B_{n+1}^{(1)} = 1_{\{Y_{n+1} < a_n\}}, \qquad B_{n+1}^{(2)} = 1_{\{a_n \leq Y_{n+1} \leq b_n\}}, \qquad B_{n+1}^{(3)} = 1_{\{Y_{n+1} > b_n\}}$$

represent the $(n+1)$st trader's actions (the three binary variables must sum to 1).

Das and Magdon-Ismail (2009) applied moment-matching to derive an approximate Bayesian learning model for this problem. We define

$$p_n = \frac{a_n - \mu_n}{\sqrt{\lambda^2 + \sigma_n^2}}, \qquad q_n = \frac{b_n - \mu_n}{\sqrt{\lambda^2 + \sigma_n^2}},$$

and apply the updating equations

$$\begin{aligned} \mu_{n+1} &= \mu_n - \sigma_n^2 \left(B_{n+1}^{(1)} \frac{1}{\sqrt{\lambda^2+\sigma_n^2}} \frac{\phi(p_n)}{\Phi(p_n)} + B_{n+1}^{(2)} \frac{1}{\sqrt{\lambda^2+\sigma_n^2}} \frac{\phi(q_n)-\phi(p_n)}{\Phi(q_n)-\Phi(p_n)} \right. \\ &\quad \left. - B_{n+1}^{(3)} \frac{1}{\sqrt{\lambda^2+\sigma_n^2}} \frac{\phi(q_n)}{1-\Phi(q_n)} \right), \qquad (13) \\ \sigma_{n+1}^2 &= \sigma_n^2 \left(1 - B_{n+1}^{(1)} \frac{\sigma_n^2}{\lambda^2+\sigma_n^2} \frac{p_n\phi(p_n)\Phi(p_n)+\phi^2(p_n)}{\Phi^2(p_n)} \right. \\ &\quad - B_{n+1}^{(2)} \frac{\sigma_n^2}{\lambda^2+\sigma_n^2} \frac{\left(q_n\phi(q_n)-p_n\phi(p_n)\right)\left(\Phi(q_n)-\Phi(p_n)\right)+\left(\phi(q_n)-\phi(p_n)\right)^2}{\left(\Phi(q_n)-\Phi(p_n)\right)^2} \\ &\quad \left. - B_{n+1}^{(3)} \frac{\sigma_n^2}{\lambda^2+\sigma_n^2} \frac{\phi^2(q_n)-q_n\phi(q_n)\left(1-\Phi(q_n)\right)}{\left(1-\Phi(q_n)\right)^2} \right). \qquad (14) \end{aligned}$$

In keeping with previous examples, we can use a projected version of (13). Consistency is then obtained from the following result.

Proposition 3 Suppose that $(a_n)_{n=0}^{\infty}$ and $(b_n)_{n=0}^{\infty}$ are bounded, and that μ_n is updated using a projected version of (13), while σ_n^2 is updated using (14). Then, $\mu_n \to \theta$ almost surely.

4.4 Learning Buyer Valuations In Online Auctions

Chhabra and Das (2011) studies a dynamic pricing problem in the context of online digital goods auctions. In this application, the sequence $(Y_n)_{n=1}^{\infty}$ represents buyer valuations, which are assumed to be independently drawn from a common distribution. The seller sets a sequence $(q_n)_{n=0}^{\infty}$ of prices, and the nth price is accepted if $Y_{n+1} > q_n$, i.e., the buyer believes that the value of the item is greater than the price. Otherwise, the price is rejected and no revenue is earned.

Viewed as a function of the price q, the acceptance probability $\rho(q) = P(Y_{n+1} > q)$ is referred to as the "demand curve." A widely-used model in revenue management (Petruzzi and Dada 1999) assumes a linear relationship

$$\rho(q) = 1 - \gamma q.$$

In practice, the slope γ is unknown and must be learned. Suppose that the prices are normalized, that is, $q_n \in [0,1]$ for all n. We can then assume that $\gamma \in (0,1)$, which lends itself to a beta prior $\gamma \sim Beta(\alpha_0, \beta_0)$. Let I_{n+1} be a binary variable that equals 1 if the $(n+1)$st buyer accepts the price q_n, and zero otherwise.

Again, moment-matching can be applied to create an approximate Bayesian learning mechanism for this problem. We define

$$\begin{aligned}
\mu_n &= \frac{\alpha_n}{\alpha_n + \beta_n}, \\
\tau_n &= \alpha_n + \beta_n, \\
A_n &= \mu_n (1 - \mu_n), \\
B_n &= 2(1 - q_n) + (3 - 2q_n - 2\mu_n q_n + \mu_n)\tau_n + (1 - \mu_n q_n)^2 \tau_n^2, \\
C_n &= q_n \tau_n \mu_n (1 - q_n)(1 + \mu_n \tau_n), \\
D_n &= q_n \tau_n (1 - \mu_n)(1 + (1 - \mu_n)\tau_n),
\end{aligned}$$

and apply the updating equations

$$\alpha_{n+1} = \alpha_n - I_{n+1}\frac{C_n}{B_n} + (1 - I_{n+1}), \tag{15}$$

$$\beta_{n+1} = \beta_n + I_{n+1}\frac{D_n}{B_n}, \tag{16}$$

$$\tau_{n+1} = \alpha_{n+1} + \beta_{n+1}, \tag{17}$$

$$\mu_{n+1} = \mu_n - \frac{1}{\tau_n + 1}\left(I_{n+1}\frac{A_n q_n}{1 - q_n \mu_n} - (1 - I_{n+1})\frac{A_n}{\mu_n} \right). \tag{18}$$

Once again, a projection could be applied to (18) to bound the posterior mean away from 0 and 1. Consistency can then be obtained.

Proposition 4 Suppose that $\inf_n q_n > 0$ and $\sup_n q_n < 1$, and that μ_n is updated using a suitable projected version of (18), while (15)-(17) are used to update α_n, β_n and τ_n. Then, $\mu_n \to \gamma$ almost surely.

5 CONCLUSION

We have presented the first theoretical framework for consistency analysis of approximate Bayesian models for sequential learning. The main insight of our work is that many of these models can be viewed as a form of stochastic approximation; more specifically, they greatly resemble online gradient descent methods with the addition of a "bias" term representing the difference between the frequentist and Bayesian versions of OGD. We have proposed a convergent SA algorithm of this form and shown that it can be used to prove the consistency of four different approximate Bayesian schemes, most of which have been studied in previous work and proven themselves in practical applications.

Our ongoing work seeks to extend this analysis to multivariate problems, such as ranking and selection where correlated beliefs are used to model relationships between alternatives (Qu et al. 2015). Correlated beliefs have demonstrated substantial practical benefits in such settings; we believe that our consistency theory has potential to establish approximate Bayesian inference as a rigorous and powerful methodology for learning in these and other problems.

REFERENCES

Asmussen, S., and P. W. Glynn. 2011. "A new proof of convergence of MCMC via the ergodic theorem". *Statistics & Probability Letters* 81 (10): 1482–1485.

Borkar, V. S. 2008. *Stochastic Approximation*. Cambridge University Press.

Bottou, L. 1998. "Online learning and stochastic approximations". In *On-line Learning in Neural Networks*, edited by D. Saad, 9–42. Cambridge.

Chau, M., M. C. Fu, H. Qu, and I. O. Ryzhov. 2014. "Simulation optimization: a tutorial overview and recent developments in gradient-based methods". In *Proceedings of the 2014 Winter Simulation Conference*, edited by A. Tolk, S. Y. Diallo, I. O. Ryzhov, L. Yilmaz, S. Buckley, and J. A. Miller, 21–35.

Chen, C.-H., S. E. Chick, L. H. Lee, and N. A. Pujowidianto. 2015. "Ranking and selection: efficient simulation budget allocation". In *Handbook of Simulation Optimization*, edited by M. C. Fu, 45–80. Springer.

Chen, Y., and I. O. Ryzhov. 2016. "A new consistency theory for approximate Bayesian inference". *Working paper, University of Maryland*.

Chhabra, M., and S. Das. 2011. "Learning the Demand Curve in Posted-Price Digital Goods Auctions". In *Proceedings of the 10th International Conference on Autonomous Agents and Multi-Agent Systems*, 63–70.

Chick, S. E. 2006. "Subjective Probability and Bayesian Methodology". In *Handbooks of Operations Research and Management Science, vol. 13: Simulation*, edited by S. Henderson and B. Nelson, 225–258. North-Holland Publishing, Amsterdam.

Dangauthier, P., R. Herbrich, T. Minka, and T. Graepel. 2007. "TrueSkill Through Time: Revisiting the History of Chess". In *Advances in Neural Information Processing Systems*, edited by J. C. Platt, D. Koller, Y. Singer, and S. Roweis, Volume 20, 337–344.

Das, S., and M. Magdon-Ismail. 2009. "Adapting to a Market Shock: Optimal Sequential Market-Making". In *Advances in Neural Information Processing Systems*, edited by D. Koller, D. Schuurmans, Y. Bengio, and L. Bottou, Volume 21, 361–368.

DeGroot, M. H. 1970. *Optimal Statistical Decisions*. John Wiley and Sons.

Gupta, S., and K. Miescke. 1996. "Bayesian look ahead one-stage sampling allocations for selection of the best population". *Journal of Statistical Planning and Inference* 54 (2): 229–244.

Herbrich, R., T. Minka, and T. Graepel. 2006. "TrueSkill™: A Bayesian Skill Rating System". In *Advances in Neural Information Processing Systems*, edited by B. Schölkopf, J. C. Platt, and T. Hoffman, Volume 19, 569–576.

Hong, L. J., and B. L. Nelson. 2009. "A Brief Introduction To Optimization Via Simulation". In *Proceedings of the 2009 Winter Simulation Conference*, edited by M. Rosetti, R. Hill, B. Johansson, A. Dunkin, and R. Ingalls, 75–85.

Jaakkola, T. S., and M. I. Jordan. 2000. "Bayesian parameter estimation via variational methods". *Statistics and Computing* 10 (1): 25–37.

Kushner, H. J., and G. Yin. 2003. *Stochastic approximation and recursive algorithms and applications (2nd ed.)*. Springer.

Minka, T. P. 2001. "Expectation propagation for approximate Bayesian inference". In *Proceedings of the 17th conference on Uncertainty in Artificial Intelligence*, 362–369.

Petruzzi, N. C., and M. Dada. 1999. "Pricing and the newsvendor problem: A review with extensions". *Operations Research* 47 (2): 183–194.

Plagnol, V., and S. Tavaré. 2004. "Approximate Bayesian computation and MCMC". In *Monte Carlo and Quasi-Monte Carlo Methods*, edited by H. Niederreiter, 99–113. Springer.

Powell, W. B., and I. O. Ryzhov. 2012. *Optimal learning*. John Wiley and Sons.

Qu, H., I. O. Ryzhov, and M. C. Fu. 2013. "Learning logistic demand curves in business-to-business pricing". In *Proceedings of the 2013 Winter Simulation Conference*, edited by R. Pasupathy, S.-H. Kim, A. Tolk, R. Hill, and M. E. Kuhl, 29–40.

Qu, H., I. O. Ryzhov, M. C. Fu, and Z. Ding. 2015. "Sequential selection with unknown correlation structures". *Operations Research* 63 (4): 931–948.

Robbins, H., and S. Monro. 1951. "A stochastic approximation method". *The Annals of Mathematical Statistics* 22 (3): 400–407.

Robbins, H., and D. Siegmund. 1985. "A Convergence Theorem for Non Negative Almost Supermartingales and Some Applications". In *Herbert Robbins Selected Papers*, edited by T. L. Lai and D. Siegmund, 111–135. Springer.

Ryzhov, I. O. 2015. "Approximate Bayesian inference for simulation and optimization". In *Modeling and optimization: theory and applications*, edited by B. Defourny and T. Terlaky, 1–28. Springer.

Sunnåker, M., A. G. Busetto, E. Numminen, J. Corander, M. Foll, and C. Dessimoz. 2013. "Approximate Bayesian computation". *PLoS Computational Biology* 9 (1): e1002803.

Zhang, Q., and Y. Song. 2015. "Simulation selection for empirical model comparison". In *Proceedings of the 2015 Winter Simulation Conference*, edited by L. Yilmaz, W. K. V. Chan, I. Moon, T. M. K. Roeder, C. Macal, and M. D. Rossetti, 3777–3788.

AUTHOR BIOGRAPHIES

YE CHEN is a Ph.D. student of Statistics in the Department of Mathematics, University of Maryland. His research interests include statistical learning, stochastic approximation and approximate Bayesian inference. His email address is yechen@math.umd.edu.

ILYA O. RYZHOV is an Assistant Professor of Operations Management and Management Science in the Decision, Operations and Information Technologies department of the Robert H. School of Business at the University of Maryland. He received a Ph.D. in Operations Research and Financial Engineering from Princeton University in 2011. His research mostly focuses on simulation optimization and optimal learning, with applications in business analytics. He coauthored the book *Optimal Learning* (Wiley, 2012), received the Best Theoretical Paper Award at the 2012 Winter Simulation Conference, and was a finalist in the 2014 INFORMS Junior Faculty Forum paper competition. His email address is iryzhov@rhsmith.umd.edu.