Proceedings of the 32nd Conference
on Decision and Control
San Antonio, Texas • December 1993



# OPTIMAL EVALUATION POLICIES FOR WORKFORCE: A BAYESIAN STOCHASTIC MODEL

Emmanuel Fernández-Gaucherand †; Sanjay Jain ‡;
Hau L. Lee ¶; Ambar G. Rao‡; and M.R. Rao *.

## I. INTRODUCTION

We model the situation where the productivity of members of a group, such as a salesforce, is periodically evaluated; those whose performance is sub-par are dismissed and replaced by new members. Individual productivity is modeled as a random variable, the distribution of which is a function of an unknown parameter. This parameter varies across the members of the group and is specified by a prior distribution. In this manner, the heterogeneity in the group is explicitly accounted for. We model the situation as a partially observable (Bayesian) stochastic control problem, and use dynamic programming techniques and the appropriate optimality equations to obtain solutions. We prove the existence of an optimal policy in the general case. Further, for the case when the sales process can be characterized by a Beta-Binomial or a Gamma-Poisson distribution, we show that the optimal policy is of the threshold type at each evaluation period, depending only on the accumulated performance up to a given period. The work that we present here is motivated by that in [CLR]. A complete version of this paper can be obtained from the first author.

## II. MODEL FORMULATION

We assume that each salesperson is characterized by a ***productivity parameter*** $\theta_t$ at time $t$. The sales level achieved by a salesperson is assumed to be either a discrete or a continuous random variable. If the sales level is discrete then let $f_t(y \mid \theta_t)$ represent the probability mass function that the salesperson with productivity level $\theta_t$ will achieve sales of $y$ units in time period $t$. If the sales level is continuous $f_t(y \mid \theta_t)$ represents the analagous probability density function. For simplicity, we assume that the sales environment and the salesperson's productivity are static, i.e.,

$$f_t(y \mid \theta_t) = f(y \mid \theta_t), \quad \theta_t = \theta, \qquad \forall t.$$

We assume that the expected sales in any given period are uniformly bounded, i.e., for $Y_t$ the (random) variable denoting sales in period $t$ we assume that $\mathbf{E}[Y_t \mid \theta] \leq M$, for some constant $M > 0$, for all $\theta \in \Theta$. The parameter $\theta$ is *unknown* to the firm. However, the firm has a prior distribution with density function $G_0(\theta)$ on the value of $\theta$. The firm observes the sales achieved by the salesperson; all information available to the firm about $\theta$ is contained in the history of sales achieved by the salesperson. Using these observations, the firm updates $G_0(\theta)$ at the beginning of each period to obtain an estimate of the salesperson's productivity. Based on this estimate the firm decides either to retain or replace the salesperson.

† Systems & Industrial Engineering Department, The University of Arizona,
Tucson, AZ 85721. Email: emmanuel@sie.arizona.edu. Research partially supported by The Engineering Foundation under grant RI-A-93-10, in part by the National Science Foundation under grant NSF-INT 9201430, and in part by a grant from the AT&T Foundation.

‡ Department of Marketing, The University of Arizona, Tucson, AZ 85721.

¶ Department of Industrial Engineering and Engineering Management, Stanford University, Stanford, CA 94305.

* Statistics and Operations Research Area, Stern School of Business, New York University, New York, NY 10006. On leave at IIM Bangalore, INDIA. Research partially supported by the National Science Foundation under grant NSF-DDM 9001705.

When the salesperson is replaced or retires, the firm incurs a cost $C$. When a firm replaces a salesperson with a new salesperson, the new salesperson's productivity parameter is again assumed to have a prior density function $G_0(\theta)$. We assume that each salesperson has a maximum finite lifetime $N$ within the organization, and that the firm has an infinite lifetime.

We model the decision problem as a **partially observable Bayesian stochastic control problem**. The (partially observable) state variable is defined to be the pair $(\theta_t, T_t)$, where $T_t$ is the time the salesperson has been in service and $\theta_t$ is the value of the productivity parameter at time $t$. The variable $T_t$ takes values in the the range $\{0, 1, \ldots, N\}$ while $\theta_t$ takes values in some parameter space $\Theta$, which depends on the choice of distribution $f_t(\cdot \mid \cdot)$.

By assumption, we have the following state transition equations:

$$\theta_t = \theta, \qquad \forall t; \tag{1}$$

$$T_{t+1} = \begin{cases} T_t + 1, & \text{if } T_t < N \text{ and the salesperson is not fired.} \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

Note that the state variable $T_t$ is reset to 0 every time a salesperson is fired or retires. The decision alternatives available to the firm at the beginning of a time period are: a) fire the salesperson and replace him with a new salesperson; or b) retain the salesperson. Let $U_t$ to be the decision taken by the firm at the beginning of time period $t$.

$$U_t = \begin{cases} 1, & \text{if the salesperson is replaced ;} \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

Let the contribution margin per unit of sales be $m$. Then the expected one-stage *reward* or *profit* obtained by the firm is

$$r(\hat{\theta}_t, U_t) = \begin{cases} m \int_0^\infty y f(y \mid \hat{\theta}_t)\, dy, & \text{if } U_t = 0; \\ -C + m \int_\Theta \int_0^\infty y f(y \mid \theta) G_0(\theta)\, dy d\theta, & \text{if } U_t = 1; \end{cases} \tag{4}$$

where $\hat{\theta}_t$ represents the estimated value of $\theta$ for the present salesperson †. Note that since $\theta$ is not observable, the true expected reward cannot be computed. Therefore an estimate $\hat{\theta}_t$ of the parameter value is obtained using all the available data and used to compute $r(\hat{\theta}_t, U_t)$, the estimate of the expected reward. Denote the sales history of the salesperson by $Y^n := (Y_1, Y_2, \ldots, Y_n)$. Given that our state variable includes $T_t$, we can restrict attention to ***stationary admissible policies***, defined as maps $\pi : (Y^{T_t}, T_t) \mapsto \{0, 1\}$; which prescribe the actions to be taken by the firm at different times, given the available information.

Assume that the firm uses a policy $\pi \in \Pi$, where $\Pi$ is the set of admissible policies. Let $R^\pi$ and $L^\pi$ be the expected sales achieved during a lifetime and the expected lifetime respectively for a salesperson randomly drawn from the population. The firm's objective is assumed to be that of maximizing average productivity, i.e.,

$$\max_{\pi \in \Pi} \left\{ \frac{mR^\pi - C}{L^\pi} \right\}. \tag{5}$$

Equation (1) to (5) describes our model.

## III. BAYESIAN OPTIMAL DECISION PROCESS

The stochastic decision problem with incomplete (or partial) state information defined above can be converted into an equivalent problem with a completely observed state by replacing $\theta_t$ by a *sufficient statistic*, obtained using the observed sales achieved by the salesperson (see [AOK], [ABFGM]). Let $p(\theta \mid Y^n)$ denote the posterior probability distribution of $\theta$ given the observations $Y^n$. The sufficient statistic for $\theta$ at time $n$ is defined to be a vector $\beta_n$ such that $p(\theta \mid Y^n) = p(\theta \mid \beta_n)$, i.e., the posterior probability distribution of $\theta$ given $Y^n$ only depends on the vector $\beta_n$. Further, let $\beta_0$ be the parameter of the prior distribution for new salespersons $G_0(\theta)$. Therefore, the decision taken at the beginning of the *nth* period depends on past observations only as summarized in $\beta_n$, and consequently we can use $\beta_n$ to infer the expected reward in (4). Thus, an equivalent *completely observable* decision process can be formulated in terms of the $(\beta_{T_t}, T_t)$, which is taken as the new state variable. Note that when a salesperson is replaced, the state $X_{T_t} := (\beta_{T_t}, T_t)$ *regenerates* to the value $(\beta_0, 0)$. In the sequel, we study such an equivalent formulation, giving complete specifications of $\{\beta_n\}$ for two different cases. In this context (equivalent) admissible policies are maps $\pi : X_{T_t} \mapsto \{0,1\}$, we will continue to denote the set of all such equivalent admissible policies by $\mathbf{\Pi}$.

**Theorem 1:** For a given policy $\pi \in \mathbf{\Pi}$, the average profits over the lifetime of the firm equal the average productivity of the salesforce, i.e.,

$$\lim_{K\to\infty} \frac{1}{K+1}\mathbf{E}^{\pi}_{X_0}\left\{\sum_{t=0}^{K} r(\beta_{T_t}, U_t)\right\} = \frac{mR^{\pi} - C}{L^{\pi}} \qquad (6)$$

where $\mathbf{E}^{\pi}_{X_0}\{\cdot\}$ is the expectation operator induced by $\pi \in \mathbf{\Pi}$ with initial state $X_0$.

Given the above, we will pose the problem as a controlled Markov process with state $X_{T_t}$, and a long-run expected average reward criterion [ABFGM].

**Theorem 2:** There exists an optimal stationary policy which maximizes the average productivity of the salesperson.

In order to enable us to characterize the structure of the optimal policy, we assume that $G_0(\theta)$ is a conjugate distribution, i.e., the distribution on $\theta$ remains in the same family of distributions after it is updated using Bayes' Theorem. Thus, an m-dimensional vector $\beta_n = (\beta_n^1, \beta_n^2, \ldots, \beta_n^m)$ can be used to represent the sufficient statistic. Furthermore $\beta_n$, and thus $G_0(\theta)$, can be recursively updated by using the observed sales history $Y^n$ and Bayes' Theorem [AOK]. Standard results exist in the literature for the posterior distributions of many of these conjugate distributions (see [DGR]). For example, if $(\beta^1, \beta^2)$ represent the parameters of the Beta distribution and the sales process is Binomial with parameters $\overline{M}$ and $\theta$, then the posterior distribution only depends on the total sales and the value of $T_t$. Here $\overline{M}$ gives the maximum possible sales in a period, and $\theta$ the probability of closing a sale in each case. The formula for the updating in this case is given as follows.

**Case I: Beta/Binomial.**

$$\beta_n^1 = \beta_0^1 + Z_n; \qquad \beta_n^2 = \beta_0^2 + n\overline{M} - Z_n$$

Similarly, if the sales process is Poisson with parameter $\theta$, and the prior distribution is given by a Gamma distribution, with shape and the scale parameters $(\beta^1, \beta^2)$. The formula for the updating in this case is given as follows.

**Case II: Gamma-Poisson.**

$$\beta_n^1 = \beta_0^1 + Z_n; \qquad \beta_n^2 = \beta_0^2 + n$$

Thus, we see that the cumulative sales $(Z_n)$ and the periods that the salesperson has been in service $(n)$, serves as the sufficient statistic for both these cases.

## IV. OPTIMAL POLICY CHARACTERIZATION

We now consider the special cases when the sales are either Binomial or Poisson and the corresponding prior distribution is Beta or Gamma respectively. We show that the optimal policy is of a threshold type. First, we recall that the conditional density $G(\cdot \mid \phi)$ is said to increase in *likelihood ratios* as $\phi$ increases if $\frac{G(\cdot|\phi_1)}{G(\cdot|\phi_2)}$ increases for $\phi_1 > \phi_2$ (see [ROSS83, p. 146].

**Lemma 1:** For both the Beta-Binomial and the Gamma-Poisson case, the posterior density of $\theta$ increases in likelihood ratio as the cumulative sales $Z$ increases.

**Lemma 2:** For both the Beta-Binomial and the Gamma-Poisson case, we have that

$$\int_\Theta \sum_y y f(y \mid \hat{\theta}) G(\hat{\theta} \mid \beta_{T_t}) d\hat{\theta}$$

increases as the cumulative sales $Z$ increases. In other words, the expected production in a period increases as the cumulative sales achieved by a salesperson in $n$ periods increases.

**Lemma 3:** For both the Beta-Binomial and the Gamma-Poisson case, the posterior density of $\theta$ decreases in likelihood ratio as the number of periods $n$ a salesperson has been in service increases.

**Lemma 4:** For both the Beta-Binomial and the Gamma-Poisson case, we have that

$$\int_\Theta \sum_y y f(y \mid \hat{\theta}) G(\hat{\theta} \mid \beta_{T_t}) d\hat{\theta}$$

decreases in $n$. In other words, the expected production in a period decreases as the number of periods taken by a salesperson to achieve a given level of cumulative sales increases.

**Theorem 3:** For both the Beta-Binomial and the Gamma-Poisson, the optimal policy $\pi^*$ is:

$$\pi^*(Z_{T_t}) = U_t^* = \begin{cases} 0, & \text{(Retain) if } Z_{T_t} \geq c^*_{T_t}; \\ 1, & \text{(Replace) otherwise.} \end{cases}$$

where $c_i^*$ $i = 1, 2 \ldots N-1$ are constants such that $c_1^* \leq c_2^* \leq \cdots \leq c_{N-1}^*$

## REFERENCES


[AOK] Aoki, M., *Optimization of Stochastic Systems: Topics in Discrete Time Systems* (1967), Academic Press.

[ABFGM] Arapostathis, A., V. Borkar, E. Fernández-Gaucherand, M.K. Ghosh and S.I. Marcus, "Discrete-Time Controlled Markov Processes with Average Cost Criterion: A Survey," *SIAM Journal on Control & Optimization*, **31** (1993), 282-344.

[CLR] Carroll, V.P., H.L. Lee and A.G. Rao, "Implications of Salesforce Productivity Heterogeneity and Demotivation: A Navy Recruiter Case Study," *Management Science*, **32** (1986), 1371-1388.

[DGR] DeGroot, M.H., *Probability and Statistics* (1986), Addison-Wesley.

[ROSS70] Ross, S.M., *Applied Probability Models with Optimization Applications* (1970), Holden-Day, San Francisco.

[ROSS83] Ross, S.M., *Introduction to Stochastic Dynamic Programming* (1983) Academic Press.