2001年度日本オペレーションズ・リサーチ学会
春季研究発表会

A Hidden-Markov Model for Mean-Shift Detection of Fraction Defective in Production Process Control

02101865 鳥取大学 *木村光宏 KIMURA Mitsuhiro
01702425 鳥取大学 山田茂 YAMADA Shigeru

1 Introduction

This study treats a statistical method of estimating mean-shift for a fraction defective of population. One traditional method for this estimation problem has been known as a CUSUM (cumulative sum) method[1, 2], and the CUSUM provides the method of estimating the occurrence of mean-shift from the observed data. We assume that the process has two states, one is good (fraction defective is low) and the other bad (high), and starts at the good state with probability one. We are interested in judging when the state has moved to the bad state by analyzing the observed data with a hidden-Markov model.

2 CUSUM method

In this section, we briefly explain the CUSUM method for estimating the mean-shift occurrence of population fraction defective (PFD). The CUSUM method has been developed for doing this and used in several practical situations. Especially, the CUSUM method is very suitable for online-monitoring of production processes if mean-shift of PFD occurs as one-way shift.

First, we assume the conditions of this shift-detection problem and the notation as follows:

1. A production process has two states as *good* and *bad* in the sense of PFD.
2. Let P_0 and P_1 be the PFDs in *good* and *bad* processes, respectively ($0 < P_0 < P_1$).
3. We denote each product from the production process by using the symbols 0 and 1. The symbol 0 means that the process produced a good product and the other symbol an unacceptable one. Thus, we obtain the sequence of these symbols (e.g. $\{0001001110...\}$) as a result of one observation of the production process.

2.1 Estimation procedure of CUSUM

The CUSUM method can statistically indicate the occurrence of the mean-shift of PFD. If the condition $P_0 < P_1$ can be assumed in the production process, the estimation procedure is given as follows:

1. Calculate y_j (usually called *score*) for the j-th symbol of the observation sequence ($j = 1, 2, \ldots$) as

$$y_j = \begin{cases} -P_0 \times \alpha & \text{for symbol } 0 \\ (1-P_0) \times \alpha & \text{for symbol } 1 \end{cases}, \qquad (1)$$

where P_0 is assumed to be given and α is a positive constant.

2. Evaluate the following equation for each symbol's index n:

$$S_n - \min_{1 \le i \le n} S_i \ge \gamma, \qquad (2)$$

where S_i is the cumulative sum of scores, $S_i = \sum_{j=1}^{i} y_j$, and γ is a constant value which can be given empirically.

3. The minimum n which satisfies Eq. (2) gives the (time) point of the shift-occurrence of PFD.

3 Hidden-Markov Modeling

We propose hidden-Markov modeling for estimating the occurrence of mean-shift of PFD in this study. Hidden-Markov models (HMMs) have been mainly developed in the research area of speech recognition. A HMM is a doubly stochastic process with an underlying process that is not observable, but can only be observed through another set of stochastic process that produce the sequence of observed symbols (Rabinar et al. [3]).

3.1 Model description

We assume the following assumptions:

1. The process has two states (denoted as 1 and 2), these represent *good* and *bad* conditions of production process, respectively.
2. Initial state is the state 0 with probability 1.
3. PFDs of the state 0 and 1 are λ and μ, respectively ($0 < \lambda < 1$, $0 < \mu < 1$), where λ is given but μ is unknown.
4. State transition probability from state 0 to 1 is P ($0 < P < 1$) (see, Fig. 1).

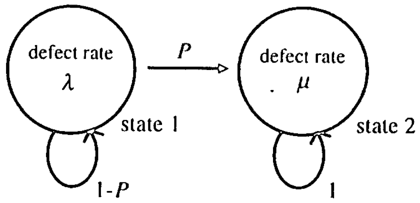

Figure 1: Structure of the model.

3.2 Estimation method

Under the assumptions above, it seems to be difficult to estimate the parameter P and μ based on the sequence of the observed symbols. However, we can apply the Baum-Welch re-estimation formulas (Rabinar et al. [3]) to the model for estimating the unknown model parameters P and μ, by analyzing the data sets.

Notation

M: the model considered in this study

n: shift-occurrence (time) point which would be estimated

T: length of the observed data

N: number of states ($N = 2$)

$O[l, m]$: sequence of the symbol observed from time l to m.

$o[i]$: i-th observed symbol

$q[t]$: state at time t

v_k: observable symbol (0 or 1) at state k ($k = 1, 2$)

a_{ij}: state transition probability between the states denoted as $Q = \begin{pmatrix} 1-P & P \\ 0 & 1 \end{pmatrix}$

$b_j[v_k]$: probability that v_k is observed at state j, that is, $b_1[0] = 1 - \lambda,\ b_1[1] = \lambda,\ b_2[0] = 1 - \mu,$ and $b_2[1] = \mu$

π: initial state distribution, i.e., $\pi = \{\pi_1, \pi_2\} = \{1, 0\}$

3.2.1 Forward-backward procedure

First, we evaluate the observation probability of $O[1, T]$ by the following procedure. The forward-path probability, $\alpha_i[t]$, is introduced here.

Step 1: Calculate $\alpha_i[1]$ ($i = 1, 2$) by

$$\alpha_1[1] = \pi_1 b_1[o[1]] = b_1[o[1]], \quad (3)$$

$$\alpha_2[1] = \pi_2 b_2[o[1]] = 0. \quad (4)$$

Step 2: Evaluate the forward-path probability from $t = 1$ to $t = T - 1$ by

$$\alpha_1[t+1] = \left(\sum_{i=1}^{2} \alpha_i[t] a_{i1}\right) b_1[o[t+1]], \quad (5)$$

$$\alpha_2[t+1] = \left(\sum_{i=1}^{2} \alpha_i[t] a_{i2}\right) b_2[o[t+1]]. \quad (6)$$

As a result of the above procedures, we have

$$\Pr[O[1, T] | M] = \sum_{i=1}^{2} \alpha_i[T]. \quad (7)$$

We also obtain the backward-path probability, $\beta_i[t]$, as

Step 1: $\beta_i[T]$ is always 1 ($i = 1, 2$).

Step 2: Evaluate the backward-path probability $\beta_i[t]$ from $T - 1$ to 1 by

$$\beta_1[t] = \sum_{j=1}^{2} a_{1j} b_j[o[t+1]] \beta_j[t+1], \quad (8)$$

$$\beta_2[t] = \sum_{j=1}^{2} a_{2j} b_j[o[t+1]] \beta_j[t+1]. \quad (9)$$

3.2.2 Baum-Welch re-estimation formulas

Now we can estimate the unknown parameters included in the model, P and μ, by using $\alpha_i[t]$ and $\beta_j[t]$.

The conditional probability that the state transition from $q[t] = 1$ to $q[t+1] = 2$ occurs at time t under $O[1, T]$, is denoted by $\gamma_{ij}[t]$ as

$$\gamma_{ij}[t] = \Pr[q[t] = i, q[t+1] = j | O[1, T], M] = \frac{\alpha_i[t] a_{ij} b_j[o[t+1]] \beta_j[t+1]}{\Pr[O[1, T], M]}. \quad (10)$$

Also, we denote the probability Pr[State is i at time t in M] by $\gamma_i[t]$ as

$$\gamma_i[t] = \sum_{j=1}^{2} \gamma_{ij}[t] = \frac{\alpha_i[t] \sum_{j=1}^{2} a_{ij} b_j[o[t+1]] \beta_j[t+1]}{\Pr[O[1, T] | M]} = \frac{\alpha_i[t] \beta_i[t]}{\Pr[O[1, T] | M]}. \quad (11)$$

Thus, we can obtain the re-estimated values of state transition probability $\bar{a}_{ij}$ by

$$\bar{a}_{ij} = \frac{\sum_{t=1}^{T-1} \gamma_{ij}[t]}{\sum_{t=1}^{T-1} \gamma_i[t]} = \frac{\sum_{t=1}^{T-1} \alpha_i[t] a_{ij} b_j[o[t+1]] \beta_j[t+1]}{\sum_{t=1}^{T-1} \alpha_i[t] \beta_i[t]}, \quad (12)$$

We directly have $\bar{P}$ from Eq. (12) as

$$\bar{P} = \bar{a}_{12} = \frac{\sum_{t=1}^{T-1} \gamma_{12}[t]}{\sum_{t=1}^{T-1} \gamma_1[t]}. \quad (13)$$

By iterating with renewing the re-estimated values, we can obtain the estimated value of unknown parameter $\bar{n} = 1/\bar{P}$ asymptotically.

On the other hand, the re-estimated value of $\bar{b}_j[v_k]$ is given by

$$\bar{b}_j[0] = \frac{\sum_{t=1}^{T} (1 - o[t]) \gamma_j[t]}{\sum_{t=1}^{T} \gamma_j[t]}, \quad (14)$$

$$\bar{b}_j[1] = \frac{\sum_{t=1}^{T} o[t] \gamma_j[t]}{\sum_{t=1}^{T} \gamma_j[t]}. \quad (15)$$

Hence, we have the re-estimated value of the parameter μ, say $\bar{\mu}$, as

$$\bar{\mu} = \bar{b}_2[1] = \frac{\sum_{t=1}^{T} o[t] \gamma_2[t]}{\sum_{t=1}^{T} \gamma_2[t]}. \quad (16)$$

By using these formulas, we can also obtain the estimated value $\bar{\mu}$ as a result of this re-estimation procedure.

References

[1] Douglas, M. H. and David, H. O., *Cumulative Sum Charts and Charting for Quality Improvement*, Springer (1998).

[2] Montgomery, D. C., *Introduction to Statistical Quality Control*, John Wiley & Sons (1996).

[3] Rabinar, L. R. and Juang, B. H., An Introduction to Hidden Markov Models, *IEEE ASSP Magazine*, **3**, 1, pp. 4-11 (1986).