시뮬레이션 실험설계에서 분산감소기법의 응용

Application of Variance Reduction Techniques in Designed Simulation Experiments*

권 치 명**
Kwon, Chimyung

Abstract

We develop a variance reduction technique in one simulation experiment whose purpose is to estimate the parameters of a first-order linear model. This method utilizes the control variates obtained during the course of simulation run under Schruben and Margolin's method (S-M method). The performance of this method is shown to be similar in estimating the main effects, and to be superior to S-M method in estimating the overall mean response in the hospital simulation experiment. For the general case, we consider that a proposed method may yield a better result than S-M method if selected control variates are highly correlated with the response at each design point.

1. INTRODUCTION

The variance reduction techniques of common random numbers and antithetic variates have been used successfully in simulation experiments that are designed to estimate an hypothesized metamodel of the mean response of interest and levels of the input factors set by the simulation analyst. (Discussions of these two variance reduction techniques are given by Bratley, Fox and Schrage [1] and Law and Kelton [5].) Schruben and Margolin showed that, for first-order metamodels, the method of common random numbers across all design points in the experiment yields superior estimates of the unknown parameters of main effects in the relationship to the method of conducting all simulation runs with independent, and randomly selected randon number streams. Furthermore, they exploited the random number assignment rule which uses a combination of common random numbers and antithetic random number streams across all design points in a simulation experiment for first-order relationship whose design matrix admits orthogonal blocking into two blocks. Their assignment rule is superior to the use of common random numbers alone, and that of independent random number streams when we estimate the parameters for the main effecfs. However, in estimating the overall mean,

* 본 연구는 1993년도 동아대학교 학술연구조성비에 의하여 이루어졌음.
** Department of Applied Statistics, Dong-A University Pusan, Korea 604-714

Schruben and Margolin's method (S-M method) may not be better than the independent streams for a certain case. The requisite assumptions as well as performance evaluations for their assignment rule have been documented by Schruben [11] and Tew and Wilson [13].

Also in simulation experiments, the method of control variates has been sucessfully applied to reduce the variability of the estimator of the mean response [1,4,7,10]. This method tries to take advantage of correlation between the response of interest and control variates (concomitant variables) obtained during the course of simulation run and to counteract the unknown variability of the estimator for the mean response of system. Thus, we consider the way of using an useful information of observed contol variates in keeping with the benifit of blocking of S-M method.

As Schruben and Margolin noted, the efficiency gain of their method highly depends on the induced correlation between two responses in the same block. Also the simulation efficiency of the control variates method is determined by the correlation between the response and a set of control variates. Typically the control variates are observed independently on each setting of factors in the factor space. Therefore, in effectively using the control variates under the S-M method, the key issuses are how to adjust the responses with the control variates and whether the adjusted responses show the similar correlation structure as the S-M method yields.

In this research, we propose a new method applying control variates to S-M rule in order to improve the simulation efficiency of the S-M method. We investigate conditions under which this method may yield better results than the S-M rule with respect to the unconditional variances of the estimator for the parameters of interest. We also explore the simulation efficiency of a proposed method through the simulation experiment on a selected model.

2. BACKGROUND AND NOTATION

In this section we provide the statistical framework necessary to formlly define a simulation experiment and its associated simulation model. Usually the goals of simulation study inculde either a determination of the settings for the input factors (experimental design points) that yields an optimal value of the system response or an understanding of the relationship between the response and the settings of input factors over the region of interest in the factor space. We consider simulation studies consisting of m design points and we run simulation n times at each design point. We let y_{ij} be the observed response from the jth replication at the ith design point. Also we let y_i be the mean response of the ith run. For the ith design point, the values of p interesting factors are specified by the design point $x_i = (x_{i1}, x_{i2},..., x_{ip})$ in the p-dimensional factor space.

In this paper, it is assumed that the relationship of the mean response to the factor settings is linear in unknown parameters. If we run a simulation m experimental points, the mean response at each desin point can be experess as follows:

$$y_i = \beta_0 + \beta_1 x_{i1}, + \beta_2 x_{i2} + ... + \beta_p x_{ip} + \varepsilon_i,\ i = 1, 2, ..., m, \tag{1}$$

where β_i is the model parameter, and ε_i is the error term. That is, the selection of m such experimental points constitutes the traditional experimental design. We now define the (mx (p+1)) design matrix X having a left column of 1's and element (i,k+1) equal to x_{ik}. Then the linear response model in (1) is given by

$$y = X\beta + \varepsilon , \tag{2}$$

where $y=(y_1,y_2,...,y_m)'$ is the vector of observed mean responses, $\beta = (\beta_0, \beta_1, ..., \beta_p)'$ is a vector of unknown parameters and $\varepsilon=(\varepsilon_1,\varepsilon_2,...,\varepsilon_m)'$ is a vector of random errors. It is assumed that ε_i's are IID $N(0,\sigma^2{}_y/n)$ for all experimental points. By a reparameterization of the factor level, the design matrix X may be chosen as an orthogonal matrix in the experimental design [12].

3. S-M METHOD

In computer simulation, an experimenter can control over the random number streams that drive a simulation model. The selection of streams of random numbers completely determines the simulation output. We let the random number streams r_{ij} denote the sequence of the random numbers having the uniform distribution U(0,1). Assume that the simulation model requires g such random number streams to drive all of its stochastic components for a single replication at each design point. Also let R_i be the set of g random number streams for the ith replication: $R_i=(r_{i1},r_{i2},...,r_{ig})$ (i=1,2,..., n).

When two replications are made with independent streams either at the design points or within each design point, two responses have zero correlation. If we use the common random numbers for the ith replication at two design points, a positive correlation of ρ_+ is induced between two observations of the responses [1,11,12]. On the other hand, using the antithetic set of streams for the ith replication at two design points induces a negative correlation of ρ_- between two observations of the responses [2,5,11,12].

For the first order linear model in (2), Schruben and Margolin exploited the random number assignment rule which uses a combination of common random numbers and antithetic streams in a simulation experiment designed to estimate the parameters β when the design matrix X admits orthogonal blocking into two blocks. For the ith replication, S-M rule uses (a) the same set of random R_i across all m_1 design points in the first block, and (b) the same set of antithetic random number streams $(1-R_i)$ across all m_2 design points in the second block $(m=m_1+m_2)$. Under the assumptions that (a) induced correlations ρ_+ and ρ_- are constant, and (b) $\rho_+ \geq -\rho_- > 0$, the covariance matrix of the mean responses $y=(y_1,y_2,...,y_m)'$ is given by

$$\mathrm{Cov}(y) = \sigma^2_y/n \begin{vmatrix} 1 & \rho_+ & : & \rho_+ & \rho_- & \rho_- & : & \rho_- \\ \rho_+ & 1 & : & \rho_+ & \rho_- & \rho_- & : & \rho_- \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho_+ & \rho_+ & : & 1 & \rho_- & \rho_- & : & \rho_- \\ \rho_- & \rho_- & : & \rho_- & 1 & \rho_+ & : & \rho_+ \\ \rho_- & \rho_- & : & \rho_- & \rho_+ & 1 & : & \rho_- \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho_- & \rho_- & : & \rho_- & \rho_+ & \rho_+ & : & 1 \end{vmatrix}$$

$$=\sigma^2_y/n\ [(\rho_+ + \rho_-)XGX'/2+(\rho_+ + \rho_-)zz'/2+(1-\rho_+)I_m], \qquad (3)$$

where G is a (p+1)x(p+1) matrix whose first row and first column entry is 1 with all other entries 0, and z is a (mx1) vector whose first m_1 elements are 1's and remaining m_2 elements are -1's (see equation (16) on p.245 in [12]). For the dispersion matrix given as above, it is known that the ordinary least squares (OLS) estimators of β in (2) and the weighted least squares (WLS) estimators are identical (see equation (63) in [9]). Schruben and Margolin [12] showed that the covariance matrix of estimator for β is given as follows

$$\mathrm{Cov}(\hat{\beta}) = \sigma^2_y/n\ [(\rho_+ + \rho_-)G/2 + (1-\rho_+)(X'X)^{-1}]. \qquad (4)$$

With respect to the design criterion of D-optimality, Schruben and Margolin's assignment rule (S-M rule) yields the OLS estimator of β with a smaller D-value than the assignment of (a) one common random numbers to all design points, or (b) a different streams to each design point for the ith replication under a certain condition, respectively. However, comparing the covariance matrix of the estimators indicates that S-M method may yield the greater variance of estimator for β_0 than independent streams for a certain case.

As we see the equation (4), the efficiency gain of S-M method highly depends on the induced correlation ρ_+ between two reponses in the same block. Thus, to be effective in applying the control variates method in conjuction with S-M method, we are concerned (a) whether the correlation between two adjusted responses in the same block is reduced or not, and (b) how effective the controls are in reducing the variance of the mean response of each design point.

4. APPLICATION OF CONTROL VARIATES TO S-M RULE

During the course of simulation run, often we can obtain the concomitant variables (control variates) which are strongly correlated with the response of interest at a little additional cost. Typically, controls are observed independently at each of the levels of factor variables during the experiment, and the mean of each control is known. We let c_{ij} be a vector of s controls corresponding to y_{ij} and c_i be the (sx1) mean vector of c_{ij}'s (j=1,2,...,n). When the length of the simulation run is sufficiently large at each design point, it seems reasonable to assume that the mean response y_i, (i=1, 2, ..., m) and s mean control vector c_i (i=1, 2, ..., m) have the (s+1)-variate normal distribution: $(y_i, c_i)' \sim N_{s+1}((\mu_{yi}, \mu_c)', \Sigma)$; where

$$\Sigma = \frac{1}{n}\begin{pmatrix} \sigma_y^2 & \sigma_{yc}' \\ \sigma_{yc} & \Sigma_c \end{pmatrix}; \tag{5}$$

σ_{yc} is a covariance between y_i and c_i; and Σ_c is a covariance of c_i [4,7,11].

Under this assumption, we try to counteract an unknown deviation of $(y_i - \mu_{yi})$ by substracting a known deviation $(c_i - \mu_c)$ from y_i. That is, the adjusted response of the ith design point is given by

$$y_i(\alpha_i) = y_i - \alpha_i'(c_i - \mu_c), \tag{6}$$

where α_i is a coefficient vector of controls. The value of α_i which minimizes $Var(y_i(\alpha_i))$ is known to be $\alpha_i = \Sigma_c^{-1}\sigma_{yc}$, and the resulting variance is $Var(y_i(\alpha_i)) = (1-R_{yc}^2)\sigma^2$, where $R_{yc}^2 = \sigma^{-2}\sigma'_{yc}\Sigma_c^{-1}\sigma_{yc}$[4].

Thus the adjusted responses $y_i(\alpha_i)$ (i=1,2,...,m) across m design points have the m-variates normal distribution. Also conditing on the controls, the adjusted mean response at each design point is an unbiased estimator of the mean response [3]. Therefore, similarly as we assume the linear response model of (2), the adjusted mean responses can be written as the linear model given by

$$y(A) = X\beta + \varepsilon^*, \tag{7}$$

where $y(A) = (y_1(\alpha_1), y_2(\alpha_2), ..., y_m(\alpha_m))'$ and ε^* is a vector of error terms.

Since the S-M rule uses the commom random numbers to drive the stochastic components of the simulation model for the factor setting in the same block, it allows the same controls for the design points in the same block. That is, the observation of the mean control variates across the m design points for the n replications is given by

$$C = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_{m_1} \\ c_{m_{1+1}} \\ \vdots \\ c_m \end{pmatrix} = \begin{pmatrix} c_1 \\ c_1 \\ \vdots \\ c_1 \\ c_{m_{1+1}} \\ \vdots \\ c_{m_{1+1}} \end{pmatrix},$$

where the first m_1 design points are in the first block, and the second $m_2(=m-m_1)$ design points are in the second block. The controls of two design points in the different blocks are negatively correlated. Thus a covariance structure of the adjusted mean responses is quite different from that of the mean resposnes obtained by S-M rule directly. To identify the covariance of $y(A)$, we first estabilish the relationships between the response and controls, and between controls across the design points and replicates as follows:

Property 1: Homogeneity of response across replicates,

$$Var(y_{ij}) = \sigma^2_y \text{ for } i=1, 2, ..., m \text{ and } j=1, 2, ..., n.$$

Property 2: Homogeneity of response covariances across replicates in either the same block or the different blocks,

$$Cov(y_{ij}, y_{kl}) = \begin{cases} \rho_+\sigma_y^2 & \text{for design points i and k in the same block, and } j=l. \\ \rho_-\sigma_y^2 & \text{for design points i and k in the different block, and } j=l. \\ 0, & \text{otherwise.} \end{cases}$$

Property 3: Homogeneity of response-control covariances across replicates in either the same block or the different blocks, and independence of the response and control variates observed on different replicates,

$$\mathrm{Cov}(y_{ij}, c_{ki}) = \begin{cases} \sigma_{yc}' & \text{for design points } i \text{ and } k \text{ in the same block, and } j=l. \\ \sigma_{yc}^{*\prime} & \text{for design points } i \text{ and } k \text{ in the different block, and } j=l. \\ 0', & \text{otherwise.} \end{cases}$$

Property 4: Homogeneity of control variates covariances across design points and replicates,

$\mathrm{Cov}(c_{ij}) = \Sigma_c$ for $i = 1, 2, ..., m$, and $k = 1, 2, ..., n$.

Property 5: Homogeneity of control-variates covariances across replicates in either the same block or the different blocks, and independence of the control variates observed on different replicates,

$$\mathrm{Cov}(c_{ij}, c_{kl}) = \begin{cases} \Sigma_c & \text{for design points } i \text{ and } k \text{ in the same block, and } j=l. \\ \Sigma_c^{*} & \text{for design points } i \text{ and } k \text{ in the different block, and } j=l. \\ O_{sxs,} & \text{otherwise.} \end{cases}$$

Properties 1 and 2 are adopted from the asssumptions estabilished by Schruben and Margolin [12]. Similarly to these relationships, we estabilsh Properties 3, 4 and 5. Under the Properties 1-5, we consider the covariance structure of the adjusted responses for the cases that the optimal coefficient vector of control variates is known and unknown.

When the optimal value of $\alpha = \Sigma_c^{-1}\sigma_{YC}$ is known, the variance of the mean controlled response at the ith design point given by

$$\mathrm{Var}(y_i - c_i'\alpha) = \mathrm{Var}(\sum_{j=1}^{n} y_{ij}/n - \sum_{j=1}^{n} c_{ij}'\alpha/n) = \sum_{j=1}^{n}\mathrm{Var}(y_{ij} - c_{ij}'\alpha)/n^2 \quad (8)$$

by Property 3. If we develop this equation, and replace α with $\Sigma_c^{-1}\sigma_{YC}$, then we find, by Property 3,

$$\mathrm{Var}(y_i - c_i'\alpha) = \sum_{j=1}^{n}\{\mathrm{Var}(y_{ij}) - 2\mathrm{Cov}(y_{ij}, c_{ij}'\alpha) + \alpha'\mathrm{Cov}(c_{ij})\alpha\}/n^2$$
$$= \sum_{j=1}^{n}(\sigma_Y^2 - 2\sigma_{YC}'\alpha + \alpha'\Sigma_C\alpha)/n^2 = \sigma_Y^2(1-R_{YC}^2)/n, \quad (9)$$

where $R_{YC}^2 = \sigma^{-2}{}_Y\sigma_{YC}'\Sigma_C^{-1}\sigma_{YC}$ is the multiple correlation coefficient between y_{ij} and c_{ij}. This equation shows that the response adjusted by the control variates with known coefficient α has a variance reduced by $(1-R_{YC}^2)$ over that given by in (3). By the Corollary 5.2.1 in Mood, Graybill and Boes [6], the covariance between two controlled responses of the design points i and k in the same block is also given by

$$\mathrm{Cov}(y_i - c_i'\alpha, y_k - c_k'\alpha) = \mathrm{Cov}(y_i, y_k) - \mathrm{Cov}(y_i, c_k'\alpha) - \mathrm{Cov}(y_k, c_i'\alpha) + \alpha'\mathrm{Cov}(c_i, c_k)\alpha, \quad (10)$$

where each term is develpoed as follows:

$$\mathrm{Cov}(y_i, y_k) = \mathrm{Cov}(\sum_{j=1}^{n} y_{ij}/n, \sum_{l=1}^{n} y_{kl}/n) = \sum_{j=1}^{n}\sum_{l=1}^{n}\mathrm{Cov}(y_{ij}, y_{kl})/n^2$$
$$= \sum_{j=1}^{n}\mathrm{Cov}(y_{ij}, y_{kj})/n^2 = \rho_+\sigma_Y^2/n \text{ (by properties 1 and 2);} \quad (11)$$

$$\mathrm{Cov}(y_i, c_k'\alpha) = \mathrm{Cov}(\sum_{j=1}^{n} y_{ij}/n, \sum_{l=1}^{n} c_{kl}'\alpha/n) = \sum_{j=1}^{n}\sum_{l=1}^{n}\mathrm{Cov}(y_{ij}, c_{kl}'\alpha)/n^2$$

which reduces to, by Property 3 and replacement of α into this equation,

$$\mathrm{Cov}(y_i, c_k'\alpha) = \sum_{j=1}^{n}\mathrm{Cov}(y_{ij}, c_{kj}'\alpha)/n^2 = \sigma'_{YC}\,\Sigma_C^{-1}\sigma_{YC}/n; \quad (12)$$

Similar to this equation, we have

$$\mathrm{Cov}(y_k, c_i'\alpha) = \sigma'_{YC}\,\Sigma_c^{-1}\sigma_{YC}/n; \quad (13)$$

$$\alpha'\mathrm{Cov}(c_i, c_k)\alpha = \alpha'\mathrm{Cov}(\sum_{j=1}^{n} c_{ij}/n, \sum_{l=1}^{n} c_{kl}/n) = \alpha'\sum_{j=1}^{n}\mathrm{Cov}(c_{ij}, c_{kj})\alpha = \sigma_{YC}'\Sigma_c^{-1}\sigma_{YC}/n, \quad (14)$$

by Properties 4 and 5, and substitution of α with $\Sigma_c^{-1}\sigma_{yc}$. Plugging (11)-(14) into (10) gives that

$$Cov(y_i-c_i'\alpha,\ y_k-c_k'\alpha)\ =(\rho_+\sigma_y^2-\sigma_{yc}'\Sigma_c^{-1}\sigma_{yc})/n=(\rho_+-R^2_{yc})\sigma^2_y/n. \quad (15)$$

As we see in this equation, the covariance between two controlled responses in the same block also decreases by the same amount as in the variance reduction of the controlled response in (3). If two controlled responses at design points i and k are not in the same block, in the similar procedures given above, we find each term in (10) as follows:

$$Cov(y_i,y_k)\ =\ \rho_-\sigma_y^2/n \text{ (by Property 2)}; \quad (16)$$

$$Cov(y_i,c_k'\alpha)=\sum_{j=1}^{n}Cov(y_{ij},\ c_{kj}'\alpha)/n^2\ =\ \sigma_{yc}^{*\prime}\Sigma_c^{-1}\sigma_{yc}\ /n \text{ (by Property 3)}; \quad (17)$$

$$Cov(y_i,c_k'\alpha)=\sum_{j=1}^{n}Cov(y_{ij},\ c_{kj}'\alpha)/n^2\ =\ \sigma_{yc}^{*\prime}\Sigma_c^{-1}\sigma_{yc}\ /n \text{ (by Property 3)}; \quad (18)$$

$$\alpha'Cov(c_i,\ c_k)\alpha\ =\ \sigma_{yc}'\ \Sigma_c^{-1}\Sigma_c^{*}\Sigma_c^{-1}\ \sigma_{yc}\ /n \text{ (by Properties 3, 4 and 5)}. \quad (19)$$

Substituion of each term in (10) with equations (16)-(19) yields

$$Cov(y_i-c_i'\alpha,y_k-c_k'\alpha)=(\rho_-\sigma^2_y\ -\ 2\sigma_{yc}'\Sigma_c^{-1}\sigma^*_{yc}\ +\ \sigma_{yc}'\Sigma_c^{-1}\Sigma_c^{*}\Sigma_c^{-1}\sigma_{yc})/n=\sigma^2_y(\rho_-\ -\ R^*_{yc})/n, \quad (20)$$

where $R^*_{yc}=\sigma^{-2}_y(2\sigma_{yc}'\Sigma_c^{-1}\sigma^*_{yc}\ -\ \sigma_{yc}'\Sigma_c^{-1}\Sigma_c^{*}\Sigma_c^{-1}\sigma_{yc})$. The term $\sigma^2_yR^*_{yc}$ can be interpreted as a difference between the covariance of the two responses and that of the two responses adjusted by the control variates when there exist the correlations among y_i, c_i, y_k, and c_k. For the case of a single control variate (s=1), the term R^*_{yc} can be written as

$$R^*_{yc}\ =\ \sigma_y^{-2}\sigma_c^{-2}\{2\sigma_{yc}\sigma^*_{yc}\ -\ \sigma^2_{yc}(-\rho_c\sigma_c^2)\sigma_c^{-2}\ \}\ =\ 2\rho_{yc}\rho^*_{yc}+\rho_c\rho_{yc}^2; \quad (21)$$

where ρ_{yc} is the correlation coefficient between y_{ij}, and c_{ij}; ρ^*_{yc} is the correltion coefficient between y_{ij}, and c_{kj} which are in two different blocks; $-\rho_c$ ($\rho_c>0$) is the correlation coefficient between c_{ij} and c_{kl} in two different blocks. Instead of identifying the relationship among ρ_{yc}, ρ^*_{yc} and ρ_c analytically, we compute this relationship based on the data set obtained from the simulation run given the example. The computational results show that $\rho^*_{yc}\fallingdotseq\rho_c\rho_{yc}$. When this relationship holds, we get

$$R^*_{yc}\ \fallingdotseq\ -2\rho_c\rho^2_{yc}\ +\ \rho_c\rho_{yc}^2\ =\ -\rho_c\rho_{yc}^2\ <\ 0, \quad (22)$$

which implies the negative correlation between the two controlled responses in the two different blocks is reduced by approximately $\rho_c\rho_{yc}^2$ for the case of a single control variate.

We now consider the covariance matrix of the controlled responses across the m design points. From equations (9), (15) and (20), if we divide the covariance matrix by the variance of the controlled response in (9), then we obtain the covariance matrix of the m controlled responses as follows:

$$Cov(y(A))=(1-R_{yc}^2)\sigma^2_y/n \begin{vmatrix} 1 & r & \vdots & r & q & q & \vdots & q \\ r & 1 & \vdots & r & q & q & \vdots & q \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ r & r & \vdots & 1 & q & q & \vdots & q \\ q & q & \vdots & q & 1 & r & \vdots & r \\ q & q & \vdots & q & r & 1 & \vdots & r \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ q & q & \vdots & q & r & r & \vdots & 1 \end{vmatrix} \quad (23)$$

where $y(A)=y-C\alpha$, $r=(\rho_+-R^2_{yc})/(1-R^2_{yc})$ and $q=(\rho_--R^*_{yc})/(1-R^2_{yc})$. This covariance matrix can be written in another form given by

$$Cov(y(A))=(1-R_{yc}^2)\sigma_y^2/n\ \ [(r+q)XGX'/2\ +\ (r+q)zz'/2\ +\ (1-r)I_m], \quad (24)$$

where G and z are defined in (3). For a dispersion matrix having the above pattern, the weighted least squares estimator for β is equal to the OLS estimator: $\beta=(X'X)^{-1}X'y(A)$. Taking the variance operation on the estimator of β and substituing $Cov(y(A))$ yields

$$\begin{aligned}Cov(\hat{\beta}) &= (X'X)^{-1}X'Cov(y(A))X(X'X)^{-1} \\ &= \sigma^2_y(1-R_{yc}^2)(X'X)^{-1}[(r+q)XGX'/2 + (r+q) zz'/2 + (1-r)I_m]X(X'X)^{-1} \\ &= \sigma^2_y(1-R_{yc}^2)/n\ [(r+q)G/2 + (1-r)(X'X)^{-1}]\end{aligned}$$

since $X'z=z'X=0$. Substituiton for r and q in (23) into the above equation finally gives

$$\begin{aligned}Cov(\hat{\beta}) &= \sigma^2_y(1-R_{yc}^2)/n\ [\{(\rho_+ - R_{yc}^2)/(1-R_{yc}^2) + (\rho_- - R_{yc}^*)/(1-R_{yc}^2)\}G/2 \\ &\quad + \{1-(\rho_+ - R_{yc}^2)/(1-R_{yc}^2)\}(X'X)^{-1}] \\ &= \sigma^2_y[\{(\rho_+ + \rho_-) - (R_{yc}^2+R_{yc}^*)\}G/2 + (1-\rho_+)(X'X)^{-1}],\end{aligned}$$

where $R^*_{yc}=\sigma_y^{-2}(2\sigma_{yc}'\Sigma_c^{-1}\sigma^*_{yc}-\sigma_{yc}'\Sigma_c^{-1}\Sigma_c^*\Sigma_c^{-1}\sigma_{yc})$. Comparing this equation with equation (3), we note that the variances of $(\beta_1,\beta_2,...,\beta_p)$ are same as those obtained by S-M rule, but the variance of β_0 is less than that in (4) if $R_{yc}^2+R^*_{yc} > 0$. Under the relationship in (22), this condition holds for the case that $s=1$ since

$$R_{yc}^2 + R^*_{yc} = \rho^2_{yc} - \rho_c\rho_{yc}^2 = (1-\rho_c)\rho_{yc}^2 > 0. \quad (26)$$

We next consider a more practical situation that the optimal value of α is unknown. In this case, we have to estimate it by its sample analogue. That is, for the ith design point, we estimate it by

$$\hat{\alpha}_i=\hat{\Sigma}_{c(i)}\quad\hat{\sigma}_{yc(i)} \quad (27)$$

where $\hat{\Sigma}_{c(i)}=\sum_{j=1}^{n}(c_{ij}-c_i)'(c_{ij}-c_i)/(n-1)$ and $\hat{\sigma}_{yc(i)}=\sum_{j=1}^{n}(y_{ij}-y_i)(c_{ij}-c_i)/(n-1)$.

Then the adjusted mean response at the ith design point is given by

$$y_i(\hat{\alpha}_i)=(y_i-\hat{\alpha}_i'c_i). \quad (28)$$

It is known that the $y_i(\hat{\alpha}_i)$ is the unbiased estimator of mean response for the ith design point and has the normal distribution [4]. Given the control variates c_i, we can obtain the sample covariance matrix of the adjusted response as follows:

$$\hat{Cov}((y(A))=\sum_{j=1}^{n}\{y(j) - y(\hat{\alpha})\}\{y(j) - y(\hat{\alpha})\}'$$

where $y(j)=[y_{1j}(\hat{\alpha}_1), y_{2j}(\hat{\alpha}_2), \cdot, y_{mj}(\hat{\alpha}_m)]'$ and $y(\hat{\alpha})=\sum_{i=1}^{m} y_i(\hat{\alpha})/n$ is the mean vector of the adjusted responses across the m design points.

Different from the known case of α, it is not easy to get the covariance matrix of the estimator β unconditionally on the observed control variates by using similar procedures for obtaining the equation (25). Thus, instead of deriving the unconditional covariance of β, given c_i $(i=1, 2, ..., n)$, we compute the sample covariance of the estimator,

$$Cov(\hat{\beta})=(X'X)^{-1}X'\hat{Cov}(y(A))X(X'X)^{-1}$$

and compare this result with that obtained by S-M method without applying control variates. We note that for a single design point, if we estimate the unknown α_i, then the variance of its mean response is increased by the amount of $(n-2)/(n-s-2)$, where s is the number of control variates (see the discussions of loss factor in [4]). Therefore, if this factor is not negligible, the variance of the mean response at each design point achieved by the control variates should compensate the variance increasement of the estimator due to the estimation of α_i for the preference of this method to S-M method.

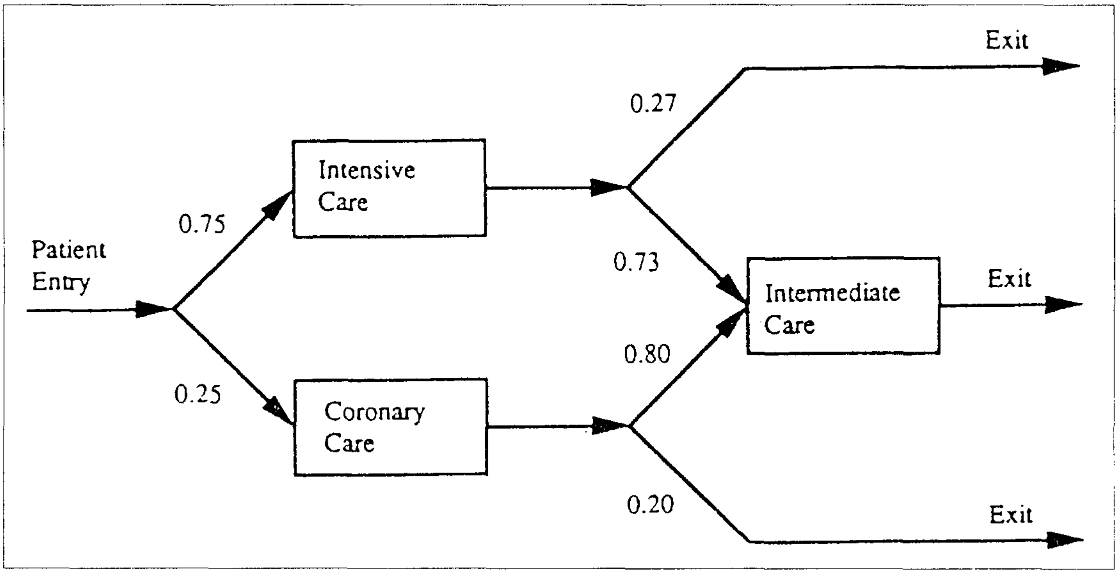
5. EXAMPLE

5.1 Description of System and Model

To compare the performances of the proposed method and the S-M method, we conduct the simulation experiment on the hospital resource allocation model originally given by Schruben and Masrgolin [12]. ⟨Figure 1⟩ shows the operation of the hospital unit in terms of patient path types of resource (see Figure A in [12]). The hospital unit consists of three types of resources that are devoted to specialized care: intensive care, coronary care and intermediate care. Patients arrive at the hospital unit according to a poisson process with an arrival rate of 3.3 per day. Upon entering the hospital, 75% of the patients need intensive care, and 25% need coronary care. The service time distribution at intensive care is lognormal with mean 3.4 days and standard deviation 1.6 days. After intensive care, 27% of the patients leave the hospital and 73% go to the intermediate care unit. Intermediate care stay for intensive care is distributed lognormally with mean 15.0 days and standard deviation 7.0 days. Finally, the length of intermediate care for coronary patients is distributed lognormally with mean 17.0 days and standard deviation 3.0 days. When the patients admission to special care units which are unavailable, they can not be accommodated and balk from the system.

The hospital administration now considers constuction of a new facility to provide better service to the patients. The administration's decision is complicated by conflicting interest of several groups since no one knows how the numbers of each type of bed will affect the frequency with which the patients can not be accomodated. To help resolve this conflict, a statistically designed simulation experiment is conducted.

To estimate the effect of the number of beds of three types to failure rate of the patients, Schruben and Margolin considered a linear model including a overall mean, all main effects and pairwise interations, and implemented 2^3 factorial design: three factors (three types of beds) having two levels for each factor. The experimental conditions for the eight design points are given in ⟨Table 1⟩. Their simulation results show that two factor interaction effects are negligible.



⟨Figure 1⟩ Hospital Resource Allocation Model

Based on these results, we also consider the linear model consisting only of the overall mean (β_0) and all main main effects ($\beta_1, \beta_2, \beta_3$). We used the same simulation output obtained by S-M method and addtionally collected a single standardized control of the interarrival time of the patients to the system (see the definition of the standardized control variates in [14]). This control variate is independent of three

⟨Table 1⟩ Experimental Design Points in 2^3 Factorial Design

Experimental DesignPoint		Number of Beds (Intensive)	Number of beds (Coronary)	Number of Beds (Intermediate)
Block 1	1	13(-1)	4(-1)	15(-1)
	2	13(-1)	6(+1)	17(+1)
	3	15(+1)	4(-1)	17(+1)
	4	15(+1)	6(+1)	15(-1)
Block 2	5	13(-1)	4(-1)	17(+1)
	6	13(-1)	6(+1)	15(-1)
	7	15(+1)	4(-1)	15(-1)
	8	15(+1)	6(+1)	17(-1)

factor variables since we use different number streams for driving the arrival process of the patient to the system. The adjusted response of interst (failure rate of patient) at the ith design point is given by

$$y_i(\hat{\alpha}_i) = y_i - \hat{\alpha}_i c_i,$$

where y_i is the mean response observed by the S-M method, $\hat{\alpha}_i$ is the estimator of the coefficient of the control (see equation (27)), and c_i is the mean control variate. We assume that the adjusted response, $y_i(\hat{\alpha}_i)$, can be described by the linear model as follows:

$$y_i(\hat{\alpha}_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \varepsilon_i,\ i = 1, 2, \ldots, 8, \quad (30)$$

where β_0 is the overall mean; β_j is the main effect of the jth factor variable (number of the the specialized care beds); x_{ij} is 1 (-1) if the jth factor is at the high (low) level for a design point i (by a reparamerization of the factor variables); and the ε_i is the error term. Clearly the (8x4) design matrix $X=(x_{ij})$ in ⟨Table 1⟩ admits the orthogonal blocking into two blocks. We partitioned the X into two blocks: the first block includes the design points 1-4, and the second block includes the design points 5-8.

We coded this model in SLAM II and conducted the 200 replications independently at a given design point according to the random number assignment rule of S-M method. We simulate this system for 1500 days. To reduce the initial bias, we collected the necessary statistics after a warm period of length 300 time units.

5.2 Experimental Results and Inferences

The variances of the responses at the eight points are in the range from 1.88 to 2.19 for the S-M method, and those of the controlled responses for the proposed method are in the range from 0.43 to 0.67. As we expected, the latter method substantially reduces the variances in estimating the mean responses across all design points. Tables 2 and 3 show the sample correlation matrices obtained by two methods. The S-M method yields the correlation coefficients in the range from 0.98 to 0.99 between responses in the same block, and from -0.54 to -0.51 between responses in different blocks. From ⟨Table 3⟩, we note that the correlations between two controlled responses in the same block are in the range from 0.92 to 0.98, and those from two different blocks are in the range from -0.24 to -0.20 for the proposed method. To explore the notion that the induced correlations are consistent with that developed in equation (23), we estimated the appropriate statistics: $\hat{\rho}_+ = 0.985$ and $\hat{R}_{yc} = -0.867$. Then, from equation (23), the correlation coefficient between the two controlled responses in the same block is esmated by

$$(\hat{\rho}_+ - \hat{R}^2_{yc})/(1-\hat{R}^2_{yc}) = (0.985 - 0.767)/(1 - 0.767) = 0.94.$$

〈Table 2〉 Correlation Matrix of Responses: S-M method

	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8
y_1	1.000	0.986	0.989	0.981	-0.540	-0.528	-0.539	-0.527
y_2	0.986	1.000	0.980	0.989	-0.540	-0.526	-0.540	-0.526
y_3	0.989	0.980	1.000	0.986	-0.525	-0.514	-0.526	-0.513
y_4	0.981	0.989	0.986	1.000	-0.530	-0.517	-0.531	-0.518
y_5	-0.540	-0.540	-0.525	-0.530	1.000	0.990	0.992	0.986
y_6	-0.528	-0.526	-0.514	-0.517	0.990	1.000	0.988	0.992
y_7	-0.539	-0.540	-0.526	-0.531	0.992	0.988	1.000	0.990
y_8	-0.527	-0.526	-0.513	-0.518	0.986	0.992	0.990	1.000

〈Table 3〉 Correlation Matrix of Responses: Proposed Method

	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8
y_1	1.000	0.941	0.954	0.920	-0.219	-0.236	-0.229	-0.234
y_2	0.941	1.000	0.916	0.954	-0.194	-0.201	-0.210	-0.209
y_3	0.954	0.916	1.000	0.943	-0.208	-0.226	-0.223	-0.226
y_4	0.920	0.954	0.943	1.000	-0.214	-0.226	-0.233	-0.231
y_5	-0.219	-0.194	-0.208	-0.214	1.000	0.965	0.971	0.955
y_6	-0.236	-0.201	-0.226	-0.226	0.965	1.000	0.957	0.976
y_7	-0.229	-0.210	-0.223	-0.233	0.971	0.957	1.000	0.969
y_8	-0.234	-0.209	-0.226	-0.231	0.955	0.976	0.969	1.000

〈Table 4〉 Estimators for Model Parameters

Parameter	S-M rule	Proposed Method
β_0	45.722	45.670
β_1	-0.291	-0.290
β_2	-0.378	-0.378
β_3	-1.805	-1.805

〈Table 5〉 Variances of Estimators and D-Value of Cov(β)

Parameter	S-M rule	Proposed Method
β_0	0.472	0.200
β_1	0.003	0.003
β_2	0.005	0.005
β_3	0.001	0.001
D-Value	9.4×10^{-9}	3.9×10^{-9}

This result indicates that the simulation result in 〈Table 3〉 is consistent with that given equation (23) for two responses in the same block. However, the same sort of conjecture is difficult to make for correlations between two controlled responses in the two different blocks since R_{YC}^{**} in (23) is a complex function of the control variates and responses.

We now compare the performances of the two methods with repect to the sample variances of the estimators, and the D-values of the sample covariance matices. 〈Table 4〉 provides the estimators for the model parameters obtained by two methods, and 〈Table 5〉 presents their variances and the D-values of the estimators' covariance matrix. From 〈Table 4〉, we note (a) in estimating the overall mean response, the proposed method is superior to S-M method, and (b) in estimating the main factor effects, the performances of both methods are similar.

6. CONCLUSIONS

In applying the S-M method, the magnitude of the correlation between two responses in the same block is critical to the efficiency of this method in reducing the variances of the estimators for the main effects and interaction effects of the factor variables. The S-M method inflates the variance of estimator of the overall mean response when the differnce between ρ_+ and ρ_- is not small. The proposed method focuses on reducing the variance of the estimator for the overall mean response by the use of control variates in the application of the S-M rule. When the relationships between the response and controls are known, the combined method always yields better results than S-M method. Simulation results on the selected model indicate a promising evidence that a combined method may yield better results than S-M method even though we have to estimate the unknown coefficients of control variates. For the case that an effective set of controls can be identified and synchronization of the random number streams is difficult to achieve in the model, we consider that the combined method utilizing the control variates with S-M method yields better results than S-M method.

REFERENCES

[1] Bratley, P., B. L. Fox and L. E. Schrage, A Guide to Simulation (2nd Ed.), Spring-Verlag, New York, 1987.

[2] Cheng, R. C. H., "The Use of Antithetic Variates in Computer Simulations," *Journal of Operational Research 31*, (1982), pp. 229-237.

[3] Kwon, C. and J. D. Tew, "Strategies for Combining Antithetic Variates and Control Variates in Designed Simulation Experiments," *Management Science 40*, (1994), pp. 1021-1034.

[4] Lavenverg, S. S., T. L. Moeller and P. D. Welch, "Statistical Results on Control Variates with Application to Queuing Simulation," *Operations Research 27*, (1982), pp. 182-202.

[5] Law, A. M. and W. D. Kelton, Simulation Modeling and Analysis (2nd Ed.), McGraw-Hill, New York, 1991.

[6] Mood, A. M., F. A. Graybill and D. C. Boes, Introduction to the Theory of Statistics, McGraw-Hill, New York, 1974.

[7] Nozari, A., S. F. Arnold and C. D. Pegden, "Control Variates for Multipopulaion Simulation Experiments," *IIE Transactions 16*, (1984), pp. 159-169.

[8] Pritsker, A. A. B., Introduction to Simulation and SLAM II, John Wiley & Sons, New York, 1986.

[9] Rao, C. R.,"Least Squares Theory Using an Estimated Dispersion Matrix and its Application to Measurement of Signals," *Proceeding of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, (1967), pp. 355-372.

[10] Rubinstein, R. Y. and R. Marcus, "Efficiency of the Multivariates Control Variates in Monte Carlo Simulation," *Operations Research 33*, (1985), pp. 661-677.

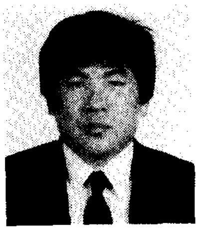
[11] Schruben, L. W., "Designing Correlation Induction Strategies for Simulation Experiments," Current Issuses in Computer Simulation, Academic Press, New York, 1979.

[12] Schruben, L. W. and Margolin, B. H., "Pseudorandom Number Assignment in Statistically Designed Simulation and Distribution Sampling Experiments," *Journal of the American Statistical Association 73*, (1978), pp. 504-525.

[13] Tew, J. D. and J. R. Wilson, "Validation of Simulation Analysis Methods for the Schruben-Margolin Correlation Induction Strategy," *Operations Research 40*, (1992), pp. 87-103.

[14] Wilson, J. R., "Variance Reduction Techniques for the Simulation for Digital Simulation," *American Journal of Mathematical and Management Sciences 4*, (1984), pp. 227-312.

● 저자소개 ●



권치명
1978 서울대학교 산업공학과 졸업
1983 서울대학교 대학원 산업공학과 석사
1990 VPI & SU 산업공학과 박사
현재 동아대학교 응용통계학과 부교수
관심분야 : Simulation Output Analysis, Simulation Optimization, FMS, Quality Management