Matematika, 1999, Jilid 15, bil. 2, hlm. 111–117

Designing Optimal Set of Experiments by Using Bayesian Method in Methyl Iodide Adsorption

Zaitul Marlizawati Zainuddin
Jabatan Matematik, Fakulti Sains
Universiti Teknologi Malaysia
81310 UTM Skudai, Johor, Malaysia

Abstract This paper deals with the design and analysis of the best sets of experiments by using Bayesian Method in the presence of prior knowledge. However, the model presented is not completely general since it is restricted to solve this particular problem, that is to find the best conditions to adsorp Methyl Iodide released after a nuclear reaction.

Keywords Bayesian Method, Experimental Design, Bayesian Experimental Design

Abstrak *Kertas ini membincangkan penggunaan Kaedah Bayes di dalam merekabentuk dan menganalisis set ujikaji terbaik jika terdapat pengetahuan awalan. Walaubagaimanapun, model yang diberikan hanyalah untuk menyelesaikan masalah ini bukannya model secara umum, iaitu untuk mendapatkan syarat yang terbaik untuk penyerapan Methyl Iodide yang terbebas hasil daripada tindakbalas nuklear.*

Katakunci *Kaedah Bayes, Rekabentuk Ujikaji, Rekabentuk Ujikaji Bayes.*

1 Introduction

Stable and radioactive isotopes of iodine are formed as fission products in irradiated reactor fuel. Iodine can therefore be released to the environment as part of routine discharges, or under fault situations. As such, it creates a potential health risk. Although some eighteen isotopes of iodine are known to exist, the isotope of primary concern from operating power reactors is I_{131}. The physical form of the iodine may be as gas (e.g. Methyl Iodide), which is generally considered to be the most highly penetrating form. Therefore, potential release must be controlled by an iodine trap that is by using an activated impregnated carbon. Activated impregnated carbon can be manufactured either from coconut or coal based. Certain conditions such as face velocity of incident CO_2 , pressure, temperature, elution

time and iodine loading should be taken into consideration so to maximize the adsorption effectiveness. This paper deals with a method in finding the optimum conditions, that is Bayesian Experimental Design. The findings will help in minimizing the amount of test work required and also the time and money in testing all the charcoals needed for the adsorption of the iodine.

2 Design of The Mathematical Curve-Fitting Model

The model presented is not completely general since they are restricted to solve this particular problem. We need to maximize the 'K value' (a measure of adsorption effectiveness) of the charcoal. The K value is defined by

$$p = 10^{-Kt} = 10^{-Kz/v} \tag{1}$$

where p is the proportion of methyl iodide penetrating at a distance z down the column, v is the face velocity, t is the dwell time.
Hence,

$$\ln(p) = -ln(10)Kt = -2.303Kt,$$

and

$$\frac{dp}{dt} = -2.303Kp. \tag{2}$$

So, $2.303K$ is the 'hazard' of adsorption. Empirically, K is a function of z and changes slope at high z.
K is a function mainly of:

1. Face velocity of incident CO_2, v;
2. Pressure, P ;
3. Temperature, T ;
4. Elution time, μ ;
5. Iodine loading, τ.

The logarithm of K value for the existing charcoal was modelled as a sum of the five linear term x_i (i.e.P, T, v, μ and τ). In addition, the five quadratic terms ${x_i}^2$ and the ten interaction terms $x_i x_j$ were included i.e.

$$\log K = c_0 + \sum_{i=1}^{5} c_i x_i + \sum_{i=1}^{5}\sum_{j=1}^{5} c_{ij} x_i x_j. \tag{3}$$

The x variables are redefined as relative to a 'central' vector of values e.g. 40cm/sec, 20 bar, 150 degrees C, 0 minutes and 25 $\mu g/g$ loading. This is done to make the approximation to be as good as possible (more accurate), by letting it to be at the centre. Furthermore, by defining it this way, we make the contribution of the quadratic term smaller.

The equation (3),

$$\mu_0 = \log K = c_0 + \sum_{i=1}^{5} c_i x_i + \sum_{i=1}^{5}\sum_{j=1}^{5} c_{ij} x_i x_j,$$

derived for the existing charcoal still has the same form, but the numerical values of the coefficients will change. The new value of the pressure, temperature, velocity, elution time and iodine loading are noted by P_2, T_2, v_2, μ_2 and τ_2 respectively.

We now assume that the new charcoal has a prediction equation

$$\mu = \log K = c_0 + \theta_0 + \sum_{i=1}^{5} (c_i + \theta_i) x_i + \sum_{i=1}^{5}\sum_{j=1}^{5} c_{ij} x_i x_j. \tag{4}$$

So

$$\mu = \mu_0 + \theta_0 + \sum_{i=1}^{5} x_i \theta_i, \tag{5}$$

where μ_0 is the expected value of log K for the existing charcoal, to which a random error with (known) variance σ^2 must be added to obtain the true value of log K . The variable μ is the expected value of log K for a new charcoal, where the six θ values are offsets that have a prior distribution that is Normal, with mean ζ and variance S for i from 0 to 5. We assume that the regression equation fitted to the existing charcoal data is an exact representation of the reality. We have put probability distributions on the constant and linear terms, and assumed that the quadratic terms do not change between charcoals.

3 Estimation of Good Prior Probability Distributions for θ

The θ_i's are offsets or just the unknown regression coefficients. θ is considered to be a random variable, that have a prior distribution that is Normal, with mean ζ and variance S for i from 0 to 5. If we have the data for two types of charcoal, to find the value for ζ and S , we combine the data of both existing charcoals that we have in one model and have a new variable h = 0 for Charcoal 1 and h=1 for Charcoal 2. The variable h is only added to the linear terms because it is assumed that the quadratic and interaction terms are fixed. The only terms that vary between charcoals are the linear terms.
The equation is now

$$\log K = y = c_0 + hc_0{}' + \sum_{i=1}^{5} (c_i + hc_i{}') x_i + \sum_{i=1}^{5}\sum_{j=1}^{5} c_{ij} x_i x_j. \tag{6}$$

The last part of the equation (6) $\sum_{i=1}^{5}\sum_{j=1}^{5} c_{ij} x_i x_j$ is assumed known. For experimental design, they are ignored.

Minitab regression is run using this equation, and the mean and variance of θ can be estimated from the regression coefficient, where

$$S = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2}{2 - 1} = \frac{1}{2}(x_1 - x_2)^2, \tag{7}$$

and

$$\bar{x} = \zeta = \frac{x_1 + x_2}{2}, \tag{8}$$

where x_1 is the coefficient of the variable when h = 0 and x_2 is when h = 1. This method is called Empirical Bayes Method because we are using the data that we already have from previous experiments in estimating the variances for the prior distribution. If we are using Bayesian approach fully, hierarchical prior is used, instead of just estimating the values for the mean and the variance of the prior distribution (Chaloner and Verdinelli [2]).

4 Expected Utility Function (Design Criteria)

We have to choose experiments that will minimise the expected value of the loss function over the posterior distribution of θ . The expected loss function is defined as:

$$R(X) = \int f(y,\theta) l(\theta, x, y) dy d\theta, \tag{9}$$

where $y = \log K$, and

$$l(\theta, x, y) = \int_{\Gamma} (\mu - E_\theta(\mu))^2 dx. \tag{10}$$

Here

$$E_\theta(\mu) = \int \mu(\theta) f(\theta|y) d\theta, \tag{11}$$

i.e. it is the mean over the posterior distribution of θ .

The integral denotes a 5-dimensional integral over a range of values for each x_i. This loss function is thus an average over the envelope of x values of the squared error in the predicted value of log K (the value before measurement error is added) due to uncertainty over the value of the θ_i after doing the experiments. As new information is obtained, it is combined with any previous information to form the basis for statistical procedures.
We have

$$f(y,\theta) = f(y|\theta) f(\theta), \tag{12}$$

where the likelihood function

$$f(y|\theta) = \prod_{j=1}^{8} (1/\sqrt{2\pi\sigma^2}) \exp\left\{ -(y_j - \mu_0 - \theta_0 - \sum_{i=1}^{5} x_{ij}\theta_i)^2 / 2\sigma^2 \right\}, \tag{13}$$

and the prior distribution for θ is expressed in terms of the normal distribution:

$$f(\theta) = \prod_{i=0}^{5} 1/(2\pi S)^{1/2} \exp\left\{ -\sum_{i=0}^{5} (\theta_i - \zeta)^2 / 2S \right\}. \tag{14}$$

We find $E_\theta(\mu) = \int \mu(\theta) f(\theta|y) d\theta$ using Bayes' theorem, where

$$f(\theta|y) = f(y|\theta) f(\theta) \Big/ \int f(y|\theta) f(\theta) d\theta. \tag{15}$$

Hence

$$R(x) = \int_\Gamma dx \left\{ \int \mu(\theta)^2 f(\theta) d\theta - \int \{E_\theta(\mu)\}^2 f(y) dy \right\}, \tag{16}$$

where the first term is a constant over all sets of experiments.

The expected loss function has now been obtained as an integral over the y_j's , the six θ_i's and the five ranges of the x_i .
Let prior distribution be

$$\theta : N[\underline{\zeta}, \underline{S}], \tag{17}$$

distribution of

$$y|\theta : N\left[X\theta, \sigma^2 I\right], \tag{18}$$

where $y_i = \sum_j X_{ij}\theta_j + \xi_i$,
and distribution of

$$\theta|y : N[\underline{\alpha}, \underline{M}^{-1}], \tag{19}$$

where

$$\underline{M} = S^{-1} + \frac{X^T X}{\sigma^2}$$

and

$$\underline{\alpha} = M^{-1}\left(S^{-1}\zeta + \frac{X^T y}{\sigma^2}\right).$$

Estimators are exhibited for the parameters in the multiple linear regression model, which, besides just using data from the present set of observations, the data obtained in previous experiments was also incorporated. These estimators have lower Bayes risk than the usual least squares estimators (Wind [6]).
The expected value of the loss function is

$$R(X) = Var\ y = \underline{x}^T M^{-1} \underline{x} = \sum (M^{-1})_{ij} x_i x_j = Tr(M^{-1}A)$$

where $A_{ij} = x_i x_j$.
Therefore,

$$\begin{aligned} R(X) &= Tr\left[M^{-1} \int_v x_i x_j dx_1 \cdots dx_5\right] \\ &= Tr[M^{-1}A]. \end{aligned} \tag{20}$$

This is an A-optimal design. Matrix A is then estimated to be

$$A_{ij} = \frac{\int x_i dx_i}{x_{ih} - x_{il}} \frac{\int x_j dx_j}{x_{jh} - x_{jl}} = \frac{1}{4}(x_{ih} + x_{il})(x_{jh} + x_{jl}), \tag{21}$$

and

$$A_{ii} = \frac{\int_{x_l}^{x_h} {x_i}^2 dx_i}{x_{ih} - x_{il}} = \frac{1}{3}\frac{({x_{ih}}^3 - {x_{il}}^3)}{x_{ih} - x_{il}} = \frac{1}{3}({x_{ih}}^2 + x_{il}x_{ih} + {x_{il}}^2). \tag{22}$$

Therefore,

$$\Delta A_{ii} = \frac{1}{12}{x_{ih}}^2 - \frac{1}{6}x_{ih}x_{il} + \frac{1}{12}{x_{il}}^2 = \frac{(x_{ih} - x_{il})^2}{12} \tag{23}$$

where, x_h is the highest value for a variable and x_l is the lowest value for a variable. However, the value for $A_{11} = 1$.

5 Numerical Results

Experiments are done under certain envelope conditions for pressure, temperature, velocity, elution time and iodine loading, that are 0 bar to 40 bar, 0 degree C to 233 degree C, 0 to 70m/s, 0 to 4 hours and 0 to 50 $\mu g/g$ respectively.

By using the multiple regression model fitted in (6), a Minitab regression is run for the data obtained from the experiments on the adsorption of radioactive iodide in charcoal. It is computed that the experiment done on the new charcoal under the same pressure, temperature, velocity, elution time and iodine loading will give the higher value for log K. The R^2 value of 96.3% indicates this model fits the data very well. The standard deviation of the random error is stated to be 0.1474. The covariance matrix S is calculated to be:

$$S = \begin{pmatrix} 5.02 \times 10^{-2} & 0 & 0 & 0 & 0 & 0 \\ 0 & 2.21 \times 10^{-6} & 0 & 0 & 0 & 0 \\ 0 & 0 & 6.27 \times 10^{-7} & 0 & 0 & 0 \\ 0 & 0 & 0 & 7.56 \times 10^{-5} & 0 & 0 \\ 0 & 0 & 0 & 0 & 6.88 \times 10^{-4} & 0 \\ 0 & 0 & 0 & 0 & 0 & 8.65 \times 10^{-5} \end{pmatrix}$$

After the estimations have been made, we have to compute the value of the expected value of the loss function given any conditions of the experiment. Then, assuming that we have four values for every variable, we want to find the combination of the conditions that will give us the minimum expected value of the loss function. A programming language called Fortran95 is used.

After running this program, the lowest expected value of the loss function is obtained to be 0.1104 at pressure = 20 bar, temperature = 150 degree C, velocity = $70ms^{-1}$, elution time = 2.5 hour and iodine loading = 30μg/g . Therefore, these are the conditions under which the experiments on different charcoals should be carried out.

6 Conclusion and Recommendation

Even though it is difficult to use Bayesian method in practical implementation, it is better to use a method that is fundamentally sound, rather than to work with one having fundamental flaws (Chaloner and Verdinelli [2]). According to Pilz [4], Bayesian approach which allows for an efficient and exhaustive use of prior knowledge and its combination with actual observations, will lead to an improved estimation and a possibly considerable reduction of experimental efforts.

This research could be done by using the hierarchical prior instead of just estimating the mean and variance of the prior distribution.

References

[**1**] K. Chaloner, *Optimal Bayesian Experimental Design for Linear Models*, Ann. Statist, **Vol.12, No.1** (1984), 283–300.

[**2**] K. Chaloner & I. Verdinelli, *Bayesian Experimental Design : A Review*, Statistical Science, **Vol.10, No.3** (1995), 273–304.

[**3**] F. Corradi, *Bayesian Forecasting*, Masters Dissertation, University of Manchester (1987).

[**4**] J. Pilz, *Bayesian Estimation and Experimental Design in Linear Regression Models*, John Wiley and Sons Limited, Chichester (1991).

[**5**] P. Whittle, *Some General Points in The Theory of Optimal Experimental Design*, J. Roy. Statist. Soc, **Ser B35** (1973), 123–130.

[**6**] S. L. Wind, *An Empirical Bayes Approach to Multiple Linear Regression*, Ann. Statist. **1** (1973), 93–103.

[**7**] R. L. Winkler, *An Introduction to Bayesian Inference and Decision*, Holt, Rinehart and Winston Inc., Indiana (1972).