Thailand Statistician
January 2022; 20(1): 227-232
http://statassoc.or.th
Contributed paper

Bayesian Approach for the Estimation of Missing Responses in Randomized Block Design

Ajantha Rudhra*[a] and Bhatra Charyulu Nallan Chakravarthula [b]
[a] Geethanjali College of Engineering and Technology, Cheeryal, Hyderabad, India.
[b] Department of Statistics, University College of Science, Osmania University, Hyderabad, India.
*Corresponding author; e-mail: ajanta.rudra@gmail.com

Received: 18 January 2018
Revised: 18 March 2018
Accepted: 27 November 2019

Abstract

Many authors made their remarkable study on estimating missing observations. When an observation is missing in a randomized block design, resulting data is incomplete to carry out the analysis as per the original plan of the experiment. In this paper an attempt is made to estimate the missing observations using Bayesian approach to estimate the missing values in randomized block design and illustrated with a suitable example.

Keywords: Missing observations, RBD, Bayes theorem, Gibbs sampling.

1. Introduction

In well-planned experiments, in some situations, the responses may not be available due to natural or manmade causes like washed away by floods, destroyed by animals; birds etc., or failed to note the response or missed due to theft etc. If an observation is missing the resulting data is incomplete to carry out the analysis as per the original plan of the experiment, due to affecting the orthogonality in the data.

Randomized block design is one of the complete block design in which the experimental material is divided into ' b ' homogeneous groups called blocks, $B_1, B_2, \ldots, B_b$ so that each block contains ' v ' experimental units, $T_1, T_2, \ldots, T_v$ be ' v ' treatments applied randomly to the experimental units within the blocks i.e. the treatments are assigned to each block at random. It is flexible with respect to any number of treatments and blocks and provides more accurately than completely randomized design.

Let Y_{ij} is the observation corresponding to the i^{th} treatment in the j^{th} block, μ is the overall mean of a randomized block design, α_i is effect due to i^{th} treatment, β_j is effect due to j^{th} block. ε_{ij} is the random error corresponding to Y_{ij}. Let the number of observations be $N = vb$. The general linear model for a randomized block design is

$$\underline{Y} = X\underline{\beta} + \underline{\varepsilon},$$

where $\underline{Y} = \left[Y_{11} \quad Y_{12} \quad \ldots \quad Y_{1b} \middle| Y_{21} \quad Y_{22} \quad \ldots \quad Y_{2b} \middle| \ldots \quad \ldots \middle| Y_{v1} \quad Y_{v2} \quad \ldots \quad Y_{vb} \right]'$ vector of observations, X is the design matrix of order $(N \times 1 + v + b)$,

$$X = \begin{bmatrix} 1 & 1 & 0 & \ldots & 0 & 1 & 0 & \ldots & 0 \\ 1 & 1 & 0 & \ldots & 0 & 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & .. & \ldots & \ldots \\ 1 & 1 & 0 & \ldots & 0 & 0 & 0 & \ldots & 1 \\ 1 & 0 & 1 & \ldots & 0 & 1 & 0 & \ldots & 0 \\ 1 & 0 & 1 & \ldots & 0 & 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & 0 & 1 & \ldots & 0 & 0 & 0 & \ldots & 1 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & 0 & 0 & \ldots & 1 & 1 & 0 & \ldots & 0 \\ 1 & 0 & 0 & \ldots & 1 & 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & 0 & 0 & \ldots & 1 & 0 & 0 & \ldots & 1 \end{bmatrix}$$

$\underline{\beta} = [\mu \middle| \alpha_1 \quad \alpha_2 \ldots \quad \alpha_v \middle| \beta_1 \quad \beta_2 \ldots \quad \beta_k]'$ vector of parameters, $\underline{\varepsilon} = [\varepsilon_{11} \quad \varepsilon_{12} \ldots \varepsilon_{1b} \middle| \varepsilon_{21} \quad \varepsilon_{22} \ldots \varepsilon_{2b} \middle| \ldots \middle| \varepsilon_{v1} \quad \varepsilon_{v2} \quad \ldots \quad \varepsilon_{vb}]'$ vector of random errors. Assume $\underline{\varepsilon}$ follows $N(0, \sigma^2 I)$. The least square estimate of the vector of parameters is $\hat{\underline{\beta}} = (X'X)^{-1} X'Y$ and the variance of the estimated vector of parameters is $V(\hat{\underline{\beta}}) = (X'X)^{-1}\sigma^2$, where

$$X'X = \begin{bmatrix} N & vJ_{1\times v} & bJ_{1\times b} \\ vJ_{v\times 1} & vI_{v\times v} & J_{v\times b} \\ bJ_{b\times 1} & J_{b\times v} & bI_{b\times b} \end{bmatrix}.$$

2. Bayes Method for Estimation of Missing

The classical definition of probability can be extended to continuous space based on the Bayes concept as the probability of any event is the ratio between the probability value at which parameter could impact on and the chance of the value would happen alone. It can express geometrically as the ratio of two areas as, measure of specified part of the region to measure of the whole region.

Let $\underline{y} = (y_1, y_2, ..., y_n)$ be the observed sample drawn from a population whose density is $P(\underline{y}|\theta)$ where the parameter θ follows a certain probability $P(\theta)$ then the probability of θ given $\underline{y}$ is

$$P(\theta \mid \underline{y}) = \frac{P(\theta)P(\underline{y} \mid \theta)}{P(\underline{y})} = \frac{P(\theta)P(\underline{y} \mid \theta)}{\int P(\theta)\, P(\underline{y} \mid \theta)\, d\theta}.$$

2.1. Posterior distribution of parameters in design model

Let $\underline{Y} = X\underline{\beta} + \underline{\varepsilon}$ be the general linear model for a complete block design where $\underline{Y}_{N\times 1}$ be the vector of observations corresponding to the design matrix X of size $(N \times p)$, Let $\underline{\beta}$ be the vector of

$(1+v+b)$ parameters, and $\underline{\varepsilon}$ be the vector of random error follows $N(0,\sigma^2 I)$. Assume the elements of X be 0 or 1 based on the absence or presence of particular effect in the observation.

In any complete block design if an observation is missing the resulting data is incomplete to carry out the analysis as per the original plan of the experiment, due to effect of orthogonality in the data. So, it is necessary to estimate the missing values to carry out the analysis by proper placing the observation.

Arrange the vector of observations $\underline{Y}$ as $[\underline{Y}_{N-m} \quad \underline{Y}_m]'$ where $\underline{Y}_{N-m}$ be the vector of $(N-m)$ known and $\underline{Y}_m$ is the vector of m missing values. Then the resulting partitioned general linear model be

$$\begin{bmatrix} \underline{Y}_{N-m} \\ \underline{Y}_m \end{bmatrix} = \begin{bmatrix} X_{N-m} \\ X_m \end{bmatrix} \underline{\beta} + \begin{bmatrix} \underline{\varepsilon}_{N-m} \\ \underline{\varepsilon}_m \end{bmatrix}, \tag{1}$$

where X_{N-m} is part of the design matrix corresponding to $\underline{Y}_{N-m}$ and X_m is the part of the design matrix corresponding to $\underline{Y}_m$. If $\underline{Y}$ follows $N(X\underline{\beta},\sigma^2)$, then the density of the sample observation Y_i is

$$f(Y_i,\beta,\sigma^2) = (2\pi\sigma^2)^{1/2} \exp\left\{-\frac{1}{2\sigma^2}\left[(Y_i - X\beta)'(Y_i - X\beta)\right]\right\}.$$

The likelihood function of known observed sample vector $\underline{Y}_{N-m}$ in terms of unknown vector of parameters is

$$L\left(\underline{Y}_{N-m} \mid \underline{\beta},\sigma^2\right) = (2\pi\sigma^2)^{-(N-m)/2} \exp\left\{-\frac{1}{2\sigma^2}\left[(\underline{Y}_{N-m} - X_{N-m}\underline{\beta})'(\underline{Y}_{N-m} - X_{N-m}\underline{\beta})\right]\right\} \tag{2}$$

$$= \frac{-(N-m)}{2}\log 2\pi\sigma^2 - \frac{1}{2\sigma^2}\left[\underline{Y}'_{N-m}\underline{Y}_{N-m} - \underline{Y}'_{N-m}X_{N-m}\underline{\beta} - \underline{\beta}'X'_{N-m}Y_{N-m} + \underline{\beta}'X'_{N-m}X_{N-m}\underline{\beta}\right].$$

The maximum likelihood estimates of parameters $\underline{\beta}$ and σ^2 from observed sample is

$$\frac{\partial \log L}{\partial \underline{\beta}} = 0 \Rightarrow \hat{\underline{\beta}} = \left(X'_{N-m}X_{N-m}\right)^{-1} X'_{N-m}\underline{Y}_{N-m}, \tag{3}$$

$$\frac{\partial \log L}{\partial \sigma^2} = 0 \Rightarrow \frac{-N-m}{2\sigma^2} + \frac{1}{2\sigma^4}\left[\left(\underline{Y}_{N-m} - X_{N-m}\underline{\beta}\right)'\left(\underline{Y}_{N-m} - X_{N-m}\underline{\beta}\right)\right] = 0,$$

$$\Rightarrow \quad \hat{\sigma}^2 = \frac{1}{N-m}(\underline{Y}_{N-m} - X_{N-m}\underline{\beta})'(\underline{Y}_{N-m} - X_{N-m}\underline{\beta}) \cdot \tag{4}$$

The likelihood function of observed sample $\underline{Y}_{N-m}$ with the estimated parameters $\hat{\beta},\hat{\sigma}^2$ is

$$L(\underline{Y}_{N-m}|\hat{\underline{\beta}},\hat{\sigma}^2) = (2\pi\hat{\sigma}^2)^{-(N-m)/2} \exp\left\{-\frac{1}{2\hat{\sigma}^2}\left[(\underline{Y}_{N-m} - X_{N-m}\hat{\underline{\beta}})'(Y_{N-m} - X_{N-m}\hat{\underline{\beta}})\right]\right\}$$

$$= (2\pi\hat{\sigma}^2)^{-(N-m)/2} \exp\left\{-\frac{1}{2\hat{\sigma}^2}\left[(N-m-k)\sigma^2 + (\hat{\underline{\beta}} - \underline{\beta})' X'_{N-m}X_{N-m}(\hat{\underline{\beta}} - \underline{\beta})\right]\right\}$$

$$\left(\because (N-m-k)\sigma^2 = (\underline{Y}_{N-m} - X_{N-m}\beta)'(\underline{Y}_{N-m} - X_{N-m}\beta)\right)$$

$$= (2\pi)^{-(N-m)/2}(\hat{\sigma}^2)^{-((N-m)+1)/2} \exp\left\{-\frac{1}{2\hat{\sigma}^2}\left[(N-m-k)\hat{\sigma}^2 + (\hat{\underline{\beta}} - \underline{\beta})' X'_{N-m}X_{N-m}(\hat{\underline{\beta}} - \underline{\beta})\right]\right\}. \tag{5}$$

We have $P[\underline{Y},\underline{\beta},\sigma^2] = P[\underline{Y} \mid \underline{\beta},\sigma^2].\, P[\underline{\beta} \mid \underline{Y}].\, P[\sigma^2 \mid \underline{Y}]$, where $P(\underline{\beta}) \propto 1$ and $P(\sigma^2) = 1/\sigma$. (Refer Bayesian parametric inference by Bansal (2007)).

Then the posterior distribution of $\hat{\underline{\beta}}$ and $\hat{\sigma}^2$ can be obtained as

$$f(\hat{\underline{\beta}}) = (2\pi)^{-(N-m)/2} (\hat{\sigma}^2)^{-((N-m)+1)/2} \exp\left\{ -\frac{1}{2(X'_{N-m}X_{N-m})^{-1}\hat{\sigma}^2} (\hat{\underline{\beta}} - \underline{\beta})' (\hat{\underline{\beta}} - \underline{\beta}) \right\} \quad (6)$$

$$f(\hat{\sigma}^2) = (\hat{\sigma}^2)^{-((N-m)-k+1)/2} \exp\left\{ -\frac{1}{2\hat{\sigma}^2} (N-m-k)\sigma^2 \right\}. \quad (7)$$

i.e., $\hat{\underline{\beta}} \sim MVN(\underline{\beta}, (X'_{N-m}X_{N-m})^{-1}\hat{\sigma}^2)$ and $\hat{\sigma}^2 \sim IG\left(N-m-k, \frac{1}{2}(N-m-k)\sigma^2 \right)$.

The mean of the known observed sample $\underline{Y}$ follows $N(\overline{Y}, \sigma^2/N-m)$. The precession of randomized block design follows Gamma(a,b), where $\mathbf{a} = (vb-m-v-b-1)/2$, $\mathbf{b} = 2/(vb-m-v-b-1)\hat{\sigma}^2$. Posterior distribution can be evaluated using $P(\underline{\beta}|\underline{Y}_{N-m}) = \frac{P(\underline{Y}_{N-m}|\underline{\beta})\,P(\underline{\beta})}{\int P(\underline{Y}_{N-m}|\underline{\beta})\,P(\underline{\beta})\,d\beta}$, by generating a sequence of samples, as more and more sample values as possible with initial values for the parameters $\hat{\underline{\beta}} = (X'_{N-m}X_{N-m})^{-1}X'_{N-m}Y_{N-m}$ and $(N-m)\hat{\sigma}^2 = (\underline{Y}_{N-m} - X_{N-m}\underline{\beta})'(\underline{Y}_{N-m} - X_{N-m}\underline{\beta})$, such that the distribution of sample values more closely approximates the desired distribution and is used to evaluate the normalized constant $P(\underline{y}) = \int P(\underline{\beta}).P(\underline{y}\,\underline{\beta})\,d\beta.$

3. Bayesian Estimation of Parameters and Estimation of Missing Response in RBD Model

The posterior estimate of parameters can be obtained using Win-BUG software by writing the program for the implementation of following procedure.

Let $\underline{Y} = (Y_1, Y_2, \ldots, Y_{N-m})$ be the vector of known observations. Evaluate the Mean and Precision of the known observations. The sample mean $(\underline{Y})$ follows $N(\underline{Y}, \sigma^2/N-m)$. The precession of randomized block design follows Gamma$(\mathbf{a},\mathbf{b})$, where $\mathbf{a} = (vb-m-k)/2$, $\mathbf{b} = 2/(vb-m-k)\hat{\sigma}^2$ and k is number of parameters. i.e., $k = (v+b+1)$. Set the values for the parameters for gamma distribution (a,b). Then evaluate the initial estimate for the vector of parameters from known partitioned using $\hat{\underline{\beta}} = (X'_{N-m}X_{N-m})^{-1}X'_{N-m}Y_{N-m}$. Set the values for the parameters for normal distribution based on the design. Generate a large sample (repeatedly) from the distribution using Win-BUG program Hastings (1970). Estimate the vector of parameter each time and compute the average of parameter $\hat{\underline{\beta}}$. Estimate the 'm' missing observations using the normal equation

$$\hat{\underline{Y}}_m = X_m\hat{\underline{\beta}}.$$

The method of implementation is illustrated with suitable examples for RBD in the following example.

Example. Graham (2004) studied the problem of cost of legitimate music representation prices on five albums/artists (D12 world, Damita-J0, 30 number1Hits, Feels like home, Up-Shania Twan) with

five digital music services (I-tunes, Watmart, Musicnow, Music match, Napster) were examined are presented in Table 1 with two missing values y_1 and y_2.

Table 1 Prices on five albums/artists with five digital music services

Album	Music services				
	I-tunes	Watmart	Musicnow	Musicmatch	Napster
D_{12} World	11.99	9.44	13.99	20.79	9.95
Damita J0	13.99	y_1	13.99	12.49	13.95
30#1 Hits	9.99	27.28	y_2	9.99	9.95
Feels like home	12.87	9.44	9.99	11.99	13.95
Up-Shania Twan	9.99	17.44	13.99	9.99	18.81

The partitioned design matrix corresponding missing observations is

$$X_m = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}.$$

The parameters of normal population are 13.31522, and 18.77646. The precision follows Gamma distribution with parameters 11 and 0.004841. The estimated pairs of mean and variances for vector of parameter are [(9.62974, 0.81636), (1.6763, 3.75529), (2.7024, 3.75529), (2.6704, 3.75529), (0.0923, 3.75529), (2.4883, 3.75529), (0.2103, 3.75529), (4.5384, 3.75529), (1.6204, 3.75529), (1.4943, 3.75529), (1.766, 3.75529)]. A sample is simulated using Win-BUG program.

The estimated $\underline{\beta}$ and missing responses as [8.769, 1.489, 2.589, 2.598, 0.2481, 2.209, 0.379, 4.178, 1.79, 1.44, 1.711, 13.33, 0.08022]′ and [15.527, 13.157]′.

Remarks:

1. Win-BUG Code:

```
model
{
for(i in 1:23)
{
mu.y[i]<-
beta0+beta1*x[1,i]+beta2*x[2,i]+beta3*x[3,i]+beta4*x[4,i]+beta5*x[5,i]+beta6*x[6,i]+beta7*x
[7,i]+beta8*x[8,i]+beta9*x[9,i]+beta10*x[10,i]
y[i]~dnorm(mu.y[i], prec)
}
beta0~dnorm(9.62974,0.81636)
beta1~dnorm(1.6763,3.75529)
beta2~dnorm(2.7024,4.69411)
beta3~dnorm(2.6704,4.69411)
beta4~dnorm(0.0923,3.75529)
beta5~dnorm(2.4883,3.75529)
beta6~dnorm(0.2103,3.75529)
beta7~dnorm(4.5384,4.69411)
beta8~dnorm(1.6204,4.69411)
beta9~dnorm(1.4943,3.75529)
```

```
beta10~dnorm(1.766,3.75529)
prec~dgamma(11,0.004841)
mu~dnorm(13.31522,0.81636)
s2<-1/prec
}
list(y=c(11.99,9.44,13.99,20.79,9.95,13.99,13.99,12.49,13.95,9.99,27.28,9.99,9.95,12.87,9.
44,9.99,11.99,13.95,9.99,17.44,13.99,9.99,18.81), x=structure(.Data= c(1,1,1,1,1, 1,1,1,1, 1,1,1,1,
1,1,1,1,1, 1,1,1,1,1, 1,1,1,1,1, 0,0,0,0, 0,0,0,0, 0,0,0,0,0, 0,0,0,0,0, 0,0,0,0,0, 1,1,1,1, 0,0,0,0, 0,0,0,0,0,
0,0,0,0,0, 0,0,0,0,0, 0,0,0,0, 1,1,1,1, 0,0,0,0,0, 0,0,0,0,0, 0,0,0,0,0, 0,0,0,0, 0,0,0,0, 1,1,1,1,1, 0,0,0,0,0,
0,0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0,0, 1,1,1,1,1, 1,0,0,0,0, 1,0,0,0, 1,0,0,0, 1,0,0,0,0, 1,0,0,0,0, 0,1,0,0,0,
0,0,0,0, 0,1,0,0, 0,1,0,0,0, 0,1,0,0,0, 0,0,1,0,0, 0,1,0,0, 0,0,0,0, 0,0,1,0,0, 0,0,1,0,0, 0,0,0,1,0, 0,0,1,0,
0,0,1,0, 0,0,0,1,0, 0,0,0,1,0, 0,0,0,0,1, 0,0,0,1, 0,0,0,1, 0,0,0,0,1, 0,0,0,0,1), .Dim= c(11,23)))
list(mu=13.31522,prec=0.81636,beta0=9.62974,beta1=1.6763,beta2=2.7024,beta3=2.6704,
beta4=0.0923,beta5=2.4883,beta6=0.2103,beta7=4.5384,beta8=1.6204,beta9=1.4943,beta10=1.766)
```

2. Bayesian is a simulated posterior estimate uses large sample generated from the distribution whereas least square only depends on the small sample used.

3. Bayesian approach depends on the prior distribution of the parameter and likelihood of observed sample, whereas least estimate is depends on normality and does not plays its distribution function in estimation of missing observations.

4. Comparison of least square and Bayesian methods is presented below.

Table 2 Comparison of least square and Bayesian methods

Approach	Before		Estimated	After		
	Mean ($\bar{y}$)	Variance	$\hat{y}$	Mean ($\bar{y}$)	Variance	MSE
Least Square	13.31522	17.83938	16.87059 & 13.92059	13.48165	17.72487	25.03850
Bayesian	13.31522	17.83938	15.52700 & 13.15400	13.39736	17.40960	25.15370

Acknowledgements

The authors are thankful to the referees for improving the final version of the manuscript and also UGC for providing financial assistance to carry out this work under BSR-RFMS.

References

Bansal AK. Bayesian parametric inference. Delhi: Narosa Publishing House; 2007.

Graham E. Prices going up, but it's not gas; it's online music. USA Today. 2004; 18.

Hastings WK. Monte Carlo sampling methods using Markov chains their applications. Biometrica. 1970; 57(1): 97-109.

Park J. Estimation of missing values in linear models. Ann Stat. 1998; 3(2): 155-164.

Subramani J, Ponnuswamy NK. A non iterative least squares estimation of missing values in experimental designs. J Appl Stat. 1989; 16(1): 77-86.

Yates F. The analysis of replicated experiments when the field results are incomplete. Empire J Exp Agr. 1933; 1: 129-142.