Journal of Statistics: Advances in Theory and Applications
Volume 8, Number 2, 2012, Pages 105-130

ON ESTIMATED MEANS FOR 2^{k-p} EXPERIMENTS WITH BETA RESPONSE

LUIS FERNANDO GRAJALES[1], RAYDONAL OSPINA[2] and LUIS ALBERTO LÓPEZ[1]

[1]Department of Statistics
National University of Colombia
Bogotá
Colombia
e-mail: lfgrajalesh@unal.edu.co

[2]Department of Statistics
Federal University of Pernambuco
Recife/PE
Brazil

Abstract

Many factorial experiments include continuous responses restricted to the interval $(0, 1)$ such as percentages, rates, and indices mainly found in industry and medicine. In order to explain the response for some factors, usually the analysis of variance (ANOVA) is the most common technique because the response is considered normally distributed. However, normality and constant variance for errors is difficult to achieve for response in $(0, 1)$. This paper used beta regression models from frequentist and Bayesian perspectives to estimate mean response for 2^{k-p} factorial experiments with response in $(0, 1)$. The methodology is also compared with normal regression models. Numerical results exhibit that coverage of intervals with logit transformation was as good as the best function link for beta regression. Also, Bayesian and frequentist

2010 Mathematics Subject Classification: 62K15, 62J12, 62F15, 62F25, 93B17.

Keywords and phrases: Bayesian analysis, beta regression model, confidence intervals, experimental design, fractional factorial, rates, transformations.

Received December 20, 2012

lengths of intervals are similar for logit link. The Bayesian model with probit link presented the worst performance with respect to lengths of intervals.

1. Introduction

Some industrial and medical experiments have their response in $(0, 1)$, like percentages and rates. The normal model (or ANOVA) is the most widely used technique to predict the response by other variables through a linear regression structure; the main problem with this technique is not achieving normality and homogeneous variance of errors for responses in $(0, 1)$. A proposal to solve this problem has been doing an appropriate transformation of response, then applying ANOVA technique and finally doing the inverse transformation to analyze the results; although interpretation of results of inverse transformations are easier and good results has been reported (Patil and Kulkarni [20]). However, there are some potential problems related to transformations: (1) when inverse transformation is made to back the original units, it does not produce an unbiased estimate of the mean; (2) inverse transformations sometimes produce nonsensical results, for instance, fitted values for the response variable that exceed the lower or upper bound of variable (Myers and Montgomery [18]). For some factorial experiments with exponentially distributed response, but nonnormal, Lewis et al. [13, 14], and Patil and Kulkarni [20] presented their proposals to analyze those data by generalized linear models (GLM). They used Monte Carlo simulations and some examples to compare their results with normal and transformation approaches in some designed experiments. In all cases, GLM was better than the normal model. In most of their cases, GLM resulted better than transformations but, how can be in Patil and Kulkarni [20], testing significance of factorial effects in 2^k experiments, one of the transformations was as good as GLM approach. Because of the last results, we considered necessary to include two appropriate transformations for response in $(0, 1)$, logit and arcsin. However, as the current paper focuses in 2^{k-p} experiments whose response is continuously measured in $(0, 1)$, a logical approach is beta regression, where y in $(0, 1)$ is explained for a set of covariates, in our case expressed in a design matrix. Beta regression model (BRM) was explicitly developed by Paolino [19], Kieschnick and McCullough [11], Ferrari and Cribari-Neto [8], Vasconcellos and Cribari-Neto [24], Smithson and Verkuilen [23], Cribari-Neto and Zeileis [6], and Grün et al. [10], among others. Among the different specifications of the beta regression models, the proposed structure of Ferrari and Cribari-Neto [8] is the one presented here, as the parameterization, they use allows for direct modelling of the distribution mean using a linear predictor and a general link function, in a similar way to what is done in generalized linear models.

The paper is organized as follow: The data structure of the fractional factorial experiments and normal, transformed and beta regression model are considered in Section 2. Section 3 introduces expressions of confidence intervals on mean response for each model, the two empirical experiments analyzed from a frequentist viewpoint, and details and discussion of Monte Carlo simulation for a beta regression model. Section 4 contains details of Bayesian beta regression to estimate the mean response and its results compared with frequentist results. Finally, our conclusions are presented in Section 5.

2. Data Structure and Statistical Models

2.1 Structure and design matrix in 2^{k-p} experiments

In this paper, we discussed 2^{k-p} designed experiments with one replication, having n experimental units, observations, or runs and response in $(0, 1)$. Both experiments, complete and fractional factorial, have, as explanatory variables, k factors each at two levels, that can be dichotomous or originally continuous but coded as $+1$, usually for the highest level in continuous factors, and -1; this is a mandatory notation in these experiments, see, for instance, Box et al. [3]; the notation allows for building dimensionless covariates for the regression structure that gets a most effective and easier interpretation for fitted models. Usual linear regression, $Y = X\beta + \varepsilon$, is the proposed model to analyze 2^{k-p} experiments; those are supersaturated models, that consist in having more parameters than observations, which leads to non-identifiability problem; the common path to solve this problem is using the *aliased* or *confounded* effects, which means that some effects of lower order are equal than other higher order effects, and this reduces the number of parameters until it gets, at least, an $n-1$ parameters, that can be solved without problem.

The concept *aliased* and *confounded* depends on the *resolution* of experimental design, see Montgomery [16] and Box et al. [3] for review of those concepts. The 2^{k-p} experiments produce design matrices of order $n \times l$, here denoted by $X_{n \times l}$, which have the following properties: (i) its components $[x_{ij}] \in \{-1, 1\}$; (ii) its columns are orthogonal and $X^{\top}X = nI_l$, where I_l is the identity matrix of order l; (iii) the elements in any two different rows or in any two different columns of X coincides in exactly half the number of positions. A particular case happens in 2^k experiments when there are more of one replication and it is possible to achieve a Hadamard matrix X whit orthogonal rows. After will be discussed why is inappropriate using normal model in 2^{k-p} experiments with response in $(0, 1)$. The chief goal in this work is to consider frequentist and Bayesian models to estimate confidence (credibility) intervals on mean response for 2^{k-p} experiments and, then, to compare them by coverage or length of intervals. This strategy was used, for instance, in Lewis et al. [13, 14] although they only used frequentist analysis.

2.2 Normal regression model

Denoting y_i as the response for the i experimental run (unit), the usual normal regression model is expressed as

$$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_l x_{il} + \varepsilon_i = X_i\beta + \varepsilon_i, \quad i = 1, \ldots, n, \tag{2.1}$$

where ε_i denotes the random errors, which are assumed independent and normally distributed with mean zero and constant variance σ^2. $\beta = (\beta_0, \beta_1, \ldots, \beta_l)^\top$, the vector of unknown regression parameters, under $X^\top X$ non-singular, is estimated by the ordinary least squares (OLS) method of this form

$$\hat{\beta} = (X^\top X)^{-1} X^\top y, \tag{2.2}$$

where $y = (y_0, y_1, \ldots, y_n)^\top$ contains the response values. The covariance matrix for $\hat{\beta}$ is given by

$$\text{Var} (\hat{\beta}) = \sigma^2 (X^\top X)^{-1} , \tag{2.3}$$

and σ^2 can be estimated by the mean squared error (MSE)

$$\widehat{\sigma^2} = \frac{1}{n-(l+1)} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 , \tag{2.4}$$

where $\hat{y}_i = X_i \hat{\beta}$, X_i the i-th row of X. The inference is done by ANOVA, that assumes normality and homogeneous variance for the errors, and consequently for y.

2.3 Transformations

When the response y is in $(0, 1)$, usually the variance is not constant, and is difficult to have normality. One alternative is transforming y to get variance-stabilizing and normality in order to do the common ANOVA with transformed y, and, finally, results are presented in the original scale, doing the appropriate inverse transformation. Although this method has some drawbacks in some situations, see, for example, Myers and Montgomery [18], also it is possible to find positive results as in Patil and Kulkarni [20]. In this paper, we transform the response variable as $\tilde{y} = \log(y / (1 - y))$ and $y^* = \arcsin(\sqrt{y})$, corresponding, respectively to the logit and arcsin transformations to compare with normal and beta regression models. For these models, we consider the same linear predictor, i.e., $\tilde{y}_i = X_i\beta + \varepsilon_i$ and $y_i^* = X_i\beta + \varepsilon_i$. The analysis are similar to the normal model, the inferential procedures are done by ANOVA taking as responses $\tilde{y}$ and y^*. At the end, when the ANOVA is completed, to back the original scale, is done the inverse transformation for each case.

2.4 Beta regression model

The beta regression model is a useful alternative for modelling when the response is in $(0, 1)$. This model is found in Ferrari and Cribari-Neto [8] and was implemented in R language as betareg package for many contributors and the latest versions are documented in Cribari-Neto and Zeileis [6] and Grün et al. [10]. Here are shown some details of the model. The expression of the beta density for the response variable is

$$f(y;\, \mu,\, \phi) = \frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1-\mu)\phi)}\, y^{\mu\phi - 1}(1-y)^{(1-\mu)\phi-1}, \quad y \in (0, 1), \tag{2.5}$$

with $0 < \mu < 1$, $\phi > 0$, and $\Gamma(\cdot)$ is the gamma function. If y follows the density (2.5), we denoted by $y \sim B(\mu, \phi)$, and this situation, the mean and variance of y are, respectively, $E(y) = \mu$ and $\text{Var}(y) = \frac{\mu(1-\mu)}{1+\phi}$. The parameter ϕ can be interpreted as the *precision parameter* since, for fixed μ, the larger ϕ the smaller the variance of y; ϕ^{-1} is a dispersion parameter. This last parameterization is more suitable for modelling purposes and it will be used in this paper. The beta regresion model is especified as follow: Let $y_1, \ldots, y_n$ be a random sample such that $y_i \sim B(\mu_i, \phi)$, $i = 1, \ldots, n$, where the systematic component is given by

$$g(\mu_i) = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_l x_{il}, \tag{2.6}$$

and ϕ is fixed for all observation (beta regression model with constant precision). Here, $g(\cdot) : (0, 1) \mapsto \mathbb{R}$ is a function strictly increasing and twice differentiable, named as the link function between mean response and the regressors. Note that $\mu_i = g^{-1}(X_i\beta)$ is a function of β for X_i known. Inference in the BRM is similar than GLM: parameter estimation is performed by maximum likelihood by using nonlinear optimization. The inference can be based on asymptotic results under regularity conditions. One difference with GLM is that in beta regression, the parameters ϕ and β are not orthogonal and they have to be estimated jointly. In this work, we consider the following four links functions: logit: $g(\mu_i) = \log(\mu_i / (1 - \mu_i))$; probit: $g(\mu_i) = \Phi^{-1}(\mu_i)$, where $\Phi(\cdot)$ is the cumulative distribution function of a standard normal random variable; cloglog: $g(\mu_i) = \log(-\log(1 - \mu_i))$, and cauchit: $g(\mu_i) = \tan(\pi[\mu_i - 0.5])$, where $\pi = 3.14159....$

3. Frequentist Confidence Intervals on Mean Response

Comparing analysis on mean response is a methodology commonly used in experimental data. For instance, Lewis et al. [13, 14], which analyzed experiments with other responses by using only the frequentist analysis. Here, we consider the next specifications.

Linear regression

In 2^k and 2^{k-p} experiments, we have $(X^\top X)^{-1} = \frac{1}{n} I_{l+1}$. Thus, the limits for the $(1 - \alpha / 2) \times 100\%$ confidence interval on mean response, $E(y|\mathbf{x} = x_i)$, using the normal model are given by

$$\hat{y}(x_i) \mp \hat{\sigma} t_{\alpha/2, n-(l+1)} \sqrt{\frac{1}{l+1}(x_i^\top x_i)}, \tag{3.1}$$

where $\hat{y}(x_i) = X_i \hat{\beta}$ is the i-th fitted value, $t_{\alpha/2, n-(l+1)}$ is the percentile of the t distribution corresponding to a cumulative probability of $(1 - \alpha / 2)$ and α is the significance level. In particular cases, this interval is simpler, when $x_i^\top x_i$ is constant, for all i, and consequently all confidence intervals have the same length. Further details of the normal model can be seen in Montgomery et al. [17].

Logit transformation

We can use a logit transform for the response variable, $y^* = \log(y/(1-y))$ with inverse transformation given by $y = 1/(1+e^{-y^*})$. In this case, the $(1-\alpha/2)\times 100\%$ confidence interval is given by

$$[1/(1+e^{A}),\ 1/(1+e^{B})],$$

where $A = -\widehat{y}^*(x_i) + \widehat{\sigma}^* t_{\alpha/2, n-l-1}\sqrt{\frac{1}{l+1}(x_i^\top x_i)}$ and $B = \widehat{y}^*(x_i) - \widehat{\sigma}^* t_{\alpha/2, n-l-1} \times\sqrt{\frac{1}{l+1}(x_i^\top x_i)}$.

Arcsin transformation

Having $\tilde{y} = \arcsin(\sqrt{y})$ for the response variable and its inverse transformation given by $y = \sin^2(\tilde{y})$, we obtain the $(1-\alpha/2)\times 100\%$ confidence interval

$$\left[\sin^2\left\{(\hat{\tilde{y}}(x_i) - \hat{\tilde{\sigma}} t_{\alpha/2, n-l-1}\sqrt{\frac{1}{l+1}(x_i^\top x_i)})\right\},\right.$$
$$\left.\sin^2\left\{(\hat{\tilde{y}}(x_i) + \hat{\tilde{\sigma}} t_{\alpha/2, n-l-1}\sqrt{\frac{1}{l+1}(x_i^\top x_i)})\right\}\right].$$

Note that in these transformations, $\widehat{y}^*$ and $\widehat{\sigma}^*$, $\hat{\tilde{y}}$, $\hat{\tilde{\sigma}}$ have to be computed from the regression model applied on the respective transformed responses. More details of the transformations approaches can be seen in Montgomery et al. [17].

Asymptotic confidence intervals based on beta regression

The four frequentist $(1-\alpha/2)\times 100\%$ confidence intervals on mean response, $E(y|\mathbf{x} = x_i)$, using beta regression models are given by

(1) Link logit:

$$\left(\frac{1}{1+\exp\{-\hat{\eta} + Z_{\alpha/2}\text{s.e.}(\hat{\eta})\}}, \frac{1}{1+\exp\{-\hat{\eta} - Z_{\alpha/2}\text{s.e.}(\hat{\eta})\}}\right);$$

(2) Link probit:

$$\left[\Phi(\hat{\eta} - Z_{\alpha/2}\text{s.e.}(\hat{\eta})),\ \Phi(\hat{\eta} + Z_{\alpha/2}\text{s.e.}(\hat{\eta}))\right];$$

(3) Link cloglog:

$$\left[1 - \exp\{-\exp(\hat{\eta} - Z_{\alpha/2}\text{s.e.}(\hat{\eta}))\},\ 1 - \exp\{-\exp(\hat{\eta} + Z_{\alpha/2}\text{s.e.}(\hat{\eta}))\}\right];$$

(4) Link cauchit:

$$\left[\frac{1}{2} + \frac{\tan^{-1}(\hat{\eta} - Z_{\alpha/2}\text{s.e.}(\hat{\eta}))}{\pi},\ \frac{1}{2} + \frac{\tan^{-1}(\hat{\eta} + Z_{\alpha/2}\text{s.e.}(\hat{\eta}))}{\pi}\right];$$

where $Z_{\alpha/2} = \Phi^{-1}(1 - \alpha/2)$ indicates the upper $\alpha/2$ percentage point of the standard normal distribution; $\hat{\eta} = x^T\hat{\beta}$ is the fitted linear predictor; and $\text{s.e.}(\hat{\eta}) = \sqrt{x^T\widehat{\text{Cov}}(\hat{\beta})x}$ is the standard error of the linear predictor in that $\text{Cov}(\hat{\beta})$ is obtained from the inverse of Fisher's information matrix evaluated at the maximum likelihood estimates by excluding the row and column of this matrix corresponding to the precision parameter.

3.1. Empirical data

The trial paint data

The paint trial experiment is found in Box et al. [3]. In order to develop a paint for certain vehicles, a customer requires that the paint has high glossiness and acceptable abrasion resistance. Glossiness was measured on a scale of 1 to 100 and abrasion resistance on a scale of 1 to 10. In this work is only analyzed the glossiness. The experimenters were helped by paint technologist to decide that was necessary to include eight factors each with two levels. To carry out the complete factorial experiment 2^8, they required 256 runs or experimental units; however, they only had sixteen experimental units. The solution was to do a fractional factorial 2^{8-4} experiment, which has several characteristics: (i) it require exactly sixteen experimental units, (ii) all 8 factors can be used, and (iii) it is necessary to put some restrictions for the statistical analysis because the linear regression model will contain more parameters (over 36) than observations (16). For instance, the model includes 8 parameters for 8 main effects $(x_1, \ldots, x_8)$, 28 for two-factors interactions $(x_1x_2, \ldots, x_7x_8)$, etc. Restrictions for the analysis are based on optimum experimental designs. In our case, this 2^{8-4} experiment has *resolution* IV, denoted by 2^{8-4}_{IV}, which means that during the statistical analysis, the parameters for estimating main effects are equal than the parameters for estimating three-factor interactions effects and some parameters for two-factor interactions are equal between them, which reduces the number of parameters to estimate. When two effects are estimated by the same parameter, they are named *aliased* effects, Box et al. [3]. The design matrix and glossiness for the paint trial experiment are shown in Table 1.

Table 1. Design matrix and response for paint trial data

Run	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	$(y-1)/99$
1	–	–	–	–	–	–	–	–	0.525
2	+	–	–	–	+	+	+	–	0.596
3	–	+	–	–	+	+	–	+	0.677
4	+	+	–	–	–	–	+	+	0.778
5	–	–	+	–	+	–	+	+	0.475
6	+	–	+	–	–	+	–	+	0.667
7	–	+	+	–	–	+	+	–	0.545
8	+	+	+	–	+	–	–	–	0.778
9	–	–	–	+	–	+	+	+	0.485
10	+	–	–	+	+	–	–	+	0.677
11	–	+	–	+	+	–	+	–	0.606
12	+	+	–	+	–	+	–	–	0.808
13	–	–	+	+	+	+	–	–	0.515
14	+	–	+	+	–	–	+	–	0.697
15	–	+	+	+	–	–	–	+	0.646
16	+	+	+	+	+	+	+	+	0.818

The semiconductors data

The second example is an eight-run 2^{4-1} fractional factorial assessing the performance's percentage of a process. The design matrix and response are presented in Table 2, which are presented in Melo et al. [15].

Table 2. Design matrix and response for semiconductors data

Run	x_1	x_2	x_3	x_4	y
1	–	–	–	–	0.07
2	+	–	–	+	0.10
3	–	+	–	+	0.32
4	–	+	–	+	0.55
5	–	–	+	+	0.18
6	+	–	+	–	0.20
7	–	+	+	–	0.40
8	+	+	+	+	0.61

3.2. Comparing frequentist confidence intervals on means

Models and results for paint trial data

In order to compare the current results with the normal model found in Box et al. [3] for the paint trial data, the seven models considered in the present work have the same systematic part $\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2}$, assuming independent and normal errors for the three normal models (normal, logit, and arcsin transformations models). This assumption is not necessary for the four beta models because in those cases is modelled a function of the mean of y. Using the notation $\eta_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2}$ for the linear predictor, and $E(y_i) = \mu_i$, explicitly, we fitted the following seven models for the paint trial data:

(1) $y_i = \eta_i + \varepsilon_i$ (normal).

(2) $\text{logit}(y_i) = \eta_i + \varepsilon_i$ (logit transformation).

(3) $\sin^{-1}(\sqrt{y_i}) = \eta_i + \varepsilon_i$ (arcsin transformation).

(4) $\log(\frac{\mu_i}{1 - \mu_i}) = \eta_i$ (beta logit).

(5) $\Phi^{-1}(\mu_i) = \eta_i$ (beta probit).

(6) $\log(-\log(1 - \mu_i)) = \eta_i$ (beta cloglog).

(7) $\tan(\pi[\mu_i - \frac{1}{2}]) = \eta_i$ (beta cauchit).

For the paint trial experiment, the lengths of the confidence intervals on the mean response for each of the sixteen experimental units using the seven models are shown in Table 3. Inspections of the Table 3 indicate that the lengths of the confidence intervals show a substantial improvement in the precision of estimation using the beta approach with its different link functions in comparison with the normal and transformations models.

Table 3. Lengths of the 95% confidence intervals on the mean response in paint trial data. Normal, transformed and beta models

Run	Normal	Transformed logit	Transformed arcsin	Beta link = logit	Beta probit	Beta cloglog	Beta cauchit
1	17.73	20.27	18.81	6.43	6.29	5.82	7.49
2	17.73	17.98	17.72	5.94	5.86	5.73	6.67
3	17.73	18.92	18.19	6.15	6.04	5.84	7.04
4	17.73	13.77	15.35	4.95	4.96	4.99	4.94
5	17.73	20.27	18.81	6.43	6.29	5.82	7.49
6	17.73	17.98	17.72	5.94	5.86	5.73	6.67
7	17.73	18.92	18.19	6.15	6.04	5.84	7.04
8	17.73	13.77	15.35	4.95	4.96	4.99	4.94
9	17.73	20.27	18.81	6.43	6.29	5.82	7.49
10	17.73	17.98	17.72	5.94	5.86	5.73	6.67
11	17.73	18.92	18.19	6.15	6.04	5.84	7.04
12	17.73	13.77	15.35	4.95	4.96	4.99	4.94
13	17.73	20.27	18.81	6.43	6.29	5.82	7.49
14	17.73	17.98	17.72	5.94	5.86	5.73	6.67
15	17.73	18.92	18.19	6.15	6.04	5.84	7.04
16	17.73	13.77	15.35	4.95	4.96	4.99	4.94

Models and results semiconductors data

For semiconductors data, Melo et al. [15] applied a common linear regression model using the response on 0 a 100 scale. Their model had the covariates $x_1, x_2, x_3,$ and the interaction X_1X_2. In the current paper, using the same covariates, we modelled the response as $((y-1)/99),$ and using the same seven models studied in last section, we calculated the eight 95% confidence intervals on mean response. Then, using the notation $\zeta_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_{12} x_{i1} x_{i2}$ for the linear predictor, and $E(y_i) = \mu_i,$ explicitly, we fitted the following seven models for the semiconductors data:

(1) $y_i = \zeta_i + \varepsilon_i$ (normal).

(2) $\text{logit}(y_i) = \zeta_i + \varepsilon_i$ (logit transformation).

(3) $\sin^{-1}(\sqrt{y_i}) = \zeta_i + \varepsilon_i$ (arcsin transformation).

(4) $\log(\frac{\mu_i}{1-\mu_i}) = \zeta_i$ (beta logit).

(5) $\Phi^{-1}(\mu_i) = \zeta_i$ (beta probit).

(6) $\log(-\log(1-\mu_i)) = \zeta_i$ (beta cloglog).

(7) $\tan(\pi[\mu_i - \frac{1}{2}]) = \zeta_i$ (beta cauchit).

The confidence intervals on mean response for semiconductors data are shown in the Table 4. At analyzing the Table 4, the best results were encountered in the beta models. The logit transformation was the worse, producing greater lengths than normal model; however, it is not surprisingly because the parameter for interaction, β_{12}, was not significant ($p-$value = 0.20). This characteristic was also encountered in the beta model with cauchit link, ($p-$value = 0.70) for β_{12}; this fact can explain some lengths greater than the normal model. A little consideration should have the beta model with cloglog link: two experimental units had lengths greater than normal model, ($p-$value = 0.057) for β_{12}.

Table 4. Lengths of the 95% confidence intervals on mean response in semiconductors data. Systematic part of the models $\beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2$

		Transformed	Transformed	Beta	Beta	Beta	Beta
Run	Normal	logit	arcsin	logit	probit	cloglog	cauchit
1	0.13	0.19	0.16	0.03	0.03	0.04	0.05
2	0.13	0.23	0.18	0.07	0.06	0.07	0.08
3	0.13	0.44	0.26	0.12	0.10	0.15	0.30
4	0.13	0.51	0.29	0.12	0.10	0.16	0.36
5	0.13	0.29	0.21	0.07	0.06	0.08	0.11
6	0.13	0.34	0.22	0.08	0.07	0.09	0.12
7	0.13	0.50	0.28	0.10	0.09	0.12	0.16
8	0.13	0.47	0.28	0.10	0.08	0.12	0.15

After, we conduced a simulation study in order to compare the previous results for the paint trial data.

3.3. Monte Carlo simulation: details and discussions

In this section, it was chosen and generated a 16×1 response vector distributed beta, based on an target beta regression model with covariate matrix $X_{pt} = (\mathbf{1}, x_1, x_2)$, being x_1 and x_2 the first and second columns of the design matrix in paint trial data, i.e., the systematic part considered was $\eta_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2}$, and for this case was used the logit link function, in that $\mu_i = 1/(1 + \exp(-\eta_i))$. Then, the model contains a known mean μ_i for each experimental run (sixteen runs in a replication of the experiment). By Monte Carlo simulation, the actual simulated response y_i is obtained by adding an error drawn at random from a specified distribution to the linear predictor, namely, $y_i = 1/(1 + \exp(-\eta_i)) + \xi_i$ is generated with some initial values for $(\beta_0, \beta_1, \beta_2)$. When the beta regression model is applied throughout the sixteen rows of X_{pt}, the 16×1 response vector, y is finally generated. The precision parameter used was $\phi = 50$ for all experimental scenarios. This 16×1 simulated vector is called a *replication* of the experiment.

Each of the replications of the experiment is simulated 5,000 times and after the seven models used for the paint trial data were fitted. Summarizing, the *ideal* beta regression model has logit link and the *inappropriate* use of other models is measured by coverage of confidence intervals on the parameters. For each of the seven models, and for each of the three β parameters, 95% confidence intervals are examined. Therefore, for one replication of the experiment (16 runs), a total of 80,000 points observational are used in the coverage calculations for each parameter, and for each model. Moreover, in order to study how much are affected the results for the sample size also we considered simulations for 2, 3, and 4 replications of the experiment, concatenating the design matrix one, two, and three times horizontally leading to have 32, 48, and 64 observations, respectively. Also, each experimental situation was simulated 5,000 times and the final result is, again, coverage of confidence intervals. For instance, in the experimental design situation for 64 observations, are fitted $5{,}000 \times 7$ models. The coverage of confidence intervals on β parameters are shown in the Table 5 (16, 32, 48, and 64 runs). Table 5 indicates that in spite of logit transformation model had a bad performance with respect to lengths of confidence intervals for mean response (Table 5), this model was the best at considering the coverage of confidence intervals, because it was near to the target percentage (95%) in the three parameters for each sample size. The beta model with logit link, although was the target model, had only a satisfactory performance because it needed more sample sizes to achieve better coverage.

At this point, speaking theoretically, asymptotical results are assumed for distribution of $\hat{\beta}$ in the beta regression model. This could be justify the latter result. On another hand, the logit transformation could have achieved the normality and homogeneous variance for response, which would lead to $\hat{\beta}$ normally distributed with variance constant and, in that case, a good coverage should be achieved with a little sample size too. The beta model with cauchit had an acceptable coverage for the β_1 and β_2 parameters for all sample sizes; however, for the intercept parameter, his coverage was illogically decreasing as sample size was increasing. Three models had bad results: normal, beta with probit, and cloglog links. They never achieved coverage greater than 37% and some of them had coverage equal to zero%. A special analysis deserves the behaviour of the arcsin transformation model because in all simulations for all sample sizes, and for the three parameters, it had a coverage of zero percent indicating that all confidence intervals never included the true parameter. This result is surprisingly, taking into account that the logit transformation model presented the best performance and both transformations have the same theoretical objectives, getting normality and homogeneous variance. On another hand, this result does not surprise when Table 2 is analyzed, because the results for length were bad for the arcsin transformation model.

Table 5. Simulated results for intervals coverage (%) in normal, transformed and beta models for 2^{8-4}_{IV} factorial fractional experiment. 1, 2, 3, and 4 replications of the experiment

			Transformed	Transformed	Beta	Beta	Beta	Beta
Runs	Parameters	Normal	logit	arcsin	logit	probit	cloglog	cauchit
16	β_0	36.12	92.22	0	89.62	0.18	0	78.60
16	β_1	0	92.98	0	90.20	5.52	7.40	87.02
16	β_2	0	92.84	0	90.28	21.10	23.52	88.10
32	β_0	10.2	93.54	0	92.78	0	0	70.82
32	β_1	0	93.36	0	92.90	0.24	0.34	86.24
32	β_2	0	93.50	0	92.74	4.64	5.86	87.62
48	β_0	2.46	92.42	0	93.02	0	0	60.78
48	β_1	0	93.12	0	93.04	0	0	84.26
48	β_2	0	93.92	0	93.38	0.72	1.10	86.56
64	β_0	0	92.52	0	93.70	0	0	52.44
64	β_1	0	93.18	0	93.56	0	0	81.44
64	β_2	0	94.20	0	93.98	0	0.12	84.92

4. Bayesian Beta Regression and Credibility Intervals

One of the drawbacks of frequentist analysis of data from experimental designs is that they are based on large-samples inference, which is difficult to get in realistic sample-sizes experiments for different reasons, e.g., low budget to do the runs. The main idea of this paper is not to generate discussion about paradigms, just to have two different perspectives to analyze the same data. Readers interested on philosophical and methodological discussion between frequentist and Bayesian paradigms can encounter some papers, e. g., Bayarri and Berger [1]. However, the concepts of Bayesian statistics can be encountered in many books, e.g., Carlin and Louis [5] and Kruschke [12]. From a Bayesian approach, parameters are modelled as random variables. A summary of concepts and notation is θ: Parameter vector; $L(\mathbf{y}|\theta)$: Likelihood function (data's information); $p(\theta)$: Prior distribution about θ (before doing the experiment); $p(\theta|\mathbf{y})$: Posterior distribution; $p(\theta|\mathbf{y}) \propto L(\mathbf{y}|\theta)p(\theta)$. In this part is considered the Bayesian beta regression model (BBRM) for the paint trial's data. Although we used the same systematic part $\beta_0 + \beta_1 x_1 + \beta_2 x_2$ of frequentist model, here, the parameters are taken as random variables. The aim is to compute credibility intervals (Bayesian confidence intervals) on the mean response in order to *compare* them with frequentist confidence intervals by their lengths.

With independent data, the likelihood function for the beta regression model given by (2.5) and (2.6) is

$$L(\beta, \phi) = \prod_{i=1}^{n} \frac{\Gamma(\phi)}{\Gamma(F(x_i^{\top}\beta)\phi)\Gamma(\phi(1 - F(x_i^{\top}\beta))} y_i^{(\phi F(x_i^{\top}\beta))^{-1}} (1 - y_i)^{\phi(1-F(x_i^{\top}\beta))^{-1}},$$

where $F(\cdot)$ represent the inverse link function $g^{-1}(\cdot)$ and $n = 16$ is the sample size. The joint posterior distribution is given by

$$p(\beta, \phi|\mathbf{y}) = \frac{L(\beta, \phi)p(\beta, \phi)}{\int L(\beta, \phi)p(\beta, \phi)d\beta d\phi},$$

and eliminating terms not depending on $\boldsymbol{y}$:

$$p(\beta, \phi|\mathbf{y}) \propto p(\beta, \phi)[\Gamma(\phi)]^n \prod_{i=1}^{n} [\Gamma(F(x_i^\top \beta)\phi)\Gamma(\phi(1 - F(x_i^\top \beta)))]^{-1} y_i^{\phi F(x_i^\top \beta)}$$

$$(1 - y_i)^{\phi(1-F(x_i^\top \beta))}, \quad (4.1)$$

which is analytically untractable, i.e., does not present closed form, thus must be approximated numerically with methods based on Markov chains Monte Carlo (MCMC), such as Metropolis-Hasting and the Gibbs sampler, see Robert and Casella [21]. According to Branscum et al. [4], and Figueroa-Zúniga et al. [9], Gibbs sampling can be used to generate a Monte Carlo sample from $p(\beta, \phi|\mathbf{y})$; in this context, the Gibbs sampler involves iteratively sampling the *full conditional distributions*:

$$p(\beta|\phi, \mathbf{y}) \propto L(\beta, \phi)p(\beta), \quad (4.2)$$

$$p(\phi|\beta, \mathbf{y}) \propto L(\beta, \phi)p(\phi). \quad (4.3)$$

In this paper is used the Walk Random for Metropolis Hasting (WRMH)'s method (Robert and Casella [21]) and was implemented in software R. The WRMH's method appears detailed in the Appendix. For the analysis, we considered 50,000 Monte Carlo iterations and the results were got accounting in the last 45,000 iterations. Additionally, the necessary diagnostic tests were performed (such as convergence, autocorrelation, history), from which desirable behaviours were observed in the chains. The process to compute the length of credibility intervals on mean response for the four considered beta models can summarized thus: (i) modelling βs as normal and ϕ as gamma mentioned earlier; (ii) applying the Bayes' rule can achieve the posterior distribution and, then to estimate the Bayesian βs; (iii) using the Bayesian estimated βs and the respective design matrix were estimated the (Bayesian) mean responses for the 4 different links in beta regression; (iv) at obtaining 45,000 estimates of Bayesian mean responses, they are ordered and then are computed the length of 95% credibility intervals (with 2.5 and 97.5

percentiles); (v) the necessary diagnostics tests were performed (convergence, autocorrelation, history): desirable behaviours were observed; Figure 1 shows the diagnostics for β_1 parameter with cauchit link: upper graph presents the convergence of chain; on another hand, because we observed high autocorrelation (20th order) in chains, we used each 20 observations and got low autocorrelation, see Figure 1 (middle and lower graphs). A similar behaviour was found in the other parameters.

The lengths of credibility intervals achieved for the paint trial data are shown in Table 6 together the frequentist results. Inspection of Table 6, we can observe that similar values between Bayesian and frequentist lengths for the logit link were obtained. The Bayesian beta regression model with probit link had the worst lengths, because they are greater than the normal length: (i) this bad result is surprisingly because the Akaike's Information Criterium (AIC) for the frequentist fit for beta model presented similar values for both functions: logit link (AIC = – 54.14) and probit link (AIC = – 54.55); (ii) on another hand, this bad result is not surprisingly at comparing with results of Monte Carlo's method (Table 5) because the probit link had coverage lower than 37%, and some of them presented coverage equal to zero for some parameters and sample sizes. These contradictory results leads to suggest more studies using the probit link for beta regression models.

Bayesian models with cloglog and cauchit links had intermediate results: wider lengths and also narrower intervals than frequentist's lengths for some experimental units; it is interesting and contradictory because the cloglog link achieved the best (AIC = – 55.53) and the cauchit link got the worst (AIC = – 50.80), which could lead to judge these link functions as unstable for Bayesian analysis on mean response.

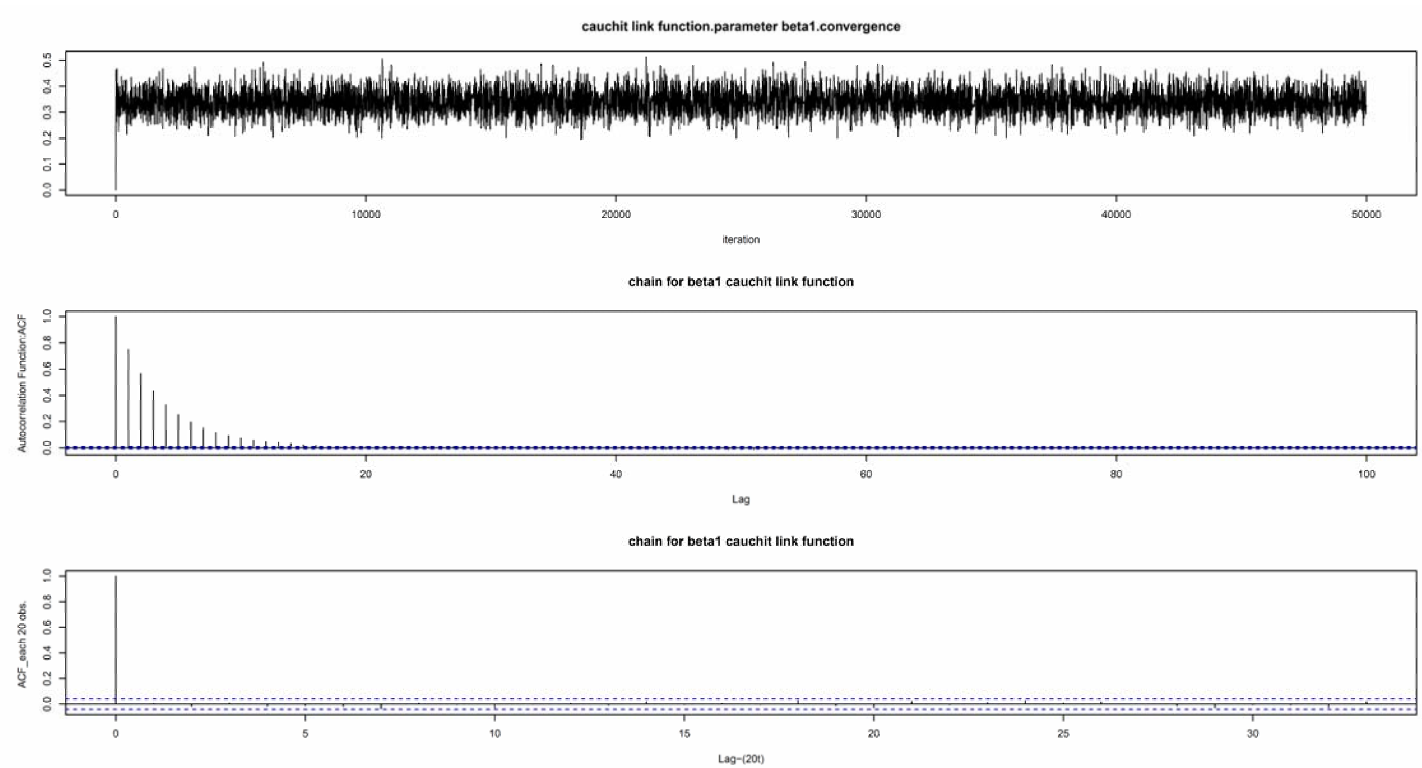

Figure 1. Some diagnostic tests for Bayesian analysis.

Table 6. Lengths of 95% frequentist and Bayesian confidence intervals on mean. Paint Trial Data

		logit	logit	probit	probit	cloglog	cloglog	cauchit	cauchit
RUN	Normal	freq	Bayes	freq	Bayes	freq	Bayes	freq	Bayes
1	17.73	**6.43**	**6.83**	6.29	22.47	5.82	10.02	**7.49**	**7.54**
2	17.73	5.94	6.42	5.86	20.14	5.73	7.67	6.67	6.76
3	17.73	6.15	6.69	6.04	21.35	5.84	8.87	7.04	6.87
4	17.73	4.95	5.33	4.96	16.92	**4.99**	**3.02**	**4.94**	**4.47**
5	17.73	6.43	6.83	6.29	22.47	5.82	10.02	7.49	7.54
6	17.73	5.94	6.42	5.86	20.14	5.73	7.67	6.67	6.76
7	17.73	6.15	6.69	6.04	21.35	5.84	8.87	7.04	6.87
8	17.73	4.95	5.33	4.96	16.92	4.99	3.02	4.94	4.47
9	17.73	6.43	6.83	6.29	22.47	5.82	10.02	7.49	7.54
10	17.73	5.94	6.42	5.86	20.14	5.73	7.67	6.67	6.76
11	17.73	6.15	6.69	6.04	21.35	5.84	8.87	7.04	6.87
12	17.73	4.95	5.33	4.96	16.92	4.99	3.02	4.94	4.47
13	17.73	6.43	6.83	6.29	22.47	5.82	10.02	7.49	7.54
14	17.73	5.94	6.42	5.86	20.14	5.73	7.67	6.67	6.76
15	17.73	6.15	6.69	6.04	21.35	5.84	8.87	7.04	6.87
16	17.73	4.95	5.33	4.96	16.92	4.99	3.02	4.94	4.47

5. Concluding Remarks

In this paper, 2^{k-p} experiments with response in $(0, 1)$ were considered; although the frequentist beta regression was enough to estimate the mean response, we also did the estimations using Bayesian beta regression in order to compare the classical results because in these experiments is common to have small samples. We compare frequentist and Bayesian confidence intervals extracted from seven models: normal, two transformations logit and arcsin, and the beta regression models with four link functions logit, probit, cloglog, and cauchit. Lengths were computed and analyzed for simulated and real data. For the paint trial data set, which is a 2_{IV}^{8-4} experiment, we did three analysis: frequentist and Bayesian beta regression for original data and frequentist analysis for simulated data in scenarios based on the original data. For the semiconductors data, a 2^{4-1} experiment, we only did the classical analysis. A different model was fitted for each experiment in order to compare with normal models found in the literature. At analyzing the results, we encountered several interesting conclusions. First, comparing the frequentist results, the lengths of confidence intervals from beta models were better than normal and transformed models for empirical paint trial data in all observations; for the semiconductors data, the cloglog and cauchit links for beta models had some greater lengths than the normal models for some observations, which is considered an unexpected and a bad result for us. After simulating with 1, 2, 3, and 4 replications of the beta logit model, using the systematic part of the paint trial data, we computed the intervals coverage for βs parameters applying the same seven models already mentioned; results for the simulated data can be summarized as follows: (i) the logit transformation model presented the best result because coverage was greater than 90% for the three parameters of the fitted model in all sample sizes (replications) of the experiment; (ii) the beta cauchit model was barely acceptable; (iii) the other four models: normal, arcsin, beta probit, and beta cloglog presented very bad results. Finally, considering the reported drawback for frequentist beta regression models with respect to the small sample sizes, we also fitted a Bayesian beta regression model for the paint trial data and computed length of credibility intervals on mean responses and they were compared with the lengths of frequentist confidence intervals for the same sixteen observations produced for the four beta models already considered. The logit link presented similar results for fits Bayesian and frequentist; the cloglog and cauchit links had contradictory results and, unfortunately, the probit link had credibility intervals greater than normal confidence intervals for almost observations; at joining these results with the AIC for frequentist models, we suggested further studies in order to elucidate some contradictory findings in our simulated and real data.

Acknowledgement

LFGH and LALP belong to the Experimental Statistics Group supported by COLCIENCIAS (Colombia), and RO gratefully acknowledges the grant from CNPq (Brazil).

References

[1] M. J. Bayarri and J. O. Berger, The interplay of Bayesian and frequentist analysis, Statistical Sciences 19 (2004), 58-80.

[2] J. M. Bernardo and A. F. Smith, Bayesian Theory, John Wiley and Sons Inc., Wiltshire, Great Britain, 2000.

[3] G. E. P. Box, W. G. Hunter and J. S. Hunter, Statistics for Experimenters, John Wiley and Sons Inc., New Jersey, USA, 2005.

[4] A. J. Branscum, W. O. Johnson and M. C. Thurmond, Bayesian beta regression: Applications to household expenditure data and genetic distance between foot-and-mouth disease viruses, Australian and New Zealand Journal of Statistics 3 (2007), 287-301.

[5] B. P. Carlin and T. A. Louis, Bayesian Methods for Data Analysis, CRC Press, Boca Raton, Florida, USA, 2009.

[6] F. Cribari-Neto and A. Zeileis, Beta regression in R. Journal of Statistical Software 2 (2010), 1-24.

[7] H. Eves, Elementary Matrix Theory, Dover Publications Inc., New York, USA, 1980.

[8] S. L. P. Ferrari and F. Cribari-Neto, Beta regression for modelling rates and proportions, Journal of Applied Statistics 31 (2004), 799-815.

[9] J. I. Figueroa-Zúniga, R. B. Arellano-Valle and S. L. P. Ferrari, Mixed beta regression: A Bayesian perspective, Preprint, 2, arXiv:1201.2375v2 (2012).

[10] B. Grün, I. Kosmidis and A. Zeileis, Extended beta regression in R: Shaken, stirred, mixed, and partitioned, Journal of Statistical Software 2 (2011), 1-24.

[11] R. Kieschnick and B. D. McCullough, Regression analysis of variates observed on (0, 1): Percentages, proportions and fractions, Statistical Modelling 3 (2003), 193-213.

[12] J. K. Kruschke, Doing Bayesian Data Analysis: A Tutorial with R and BUGS, Academic Press, Burlington, USA, 2011.

[13] S. L. Lewis, D. C. Montgomery and R. H. Myers, Examples of designed experiments with nonnormal responses, Journal of Quality Technology 3 (2001a), 265-278.

[14] S. L. Lewis, D. C. Montgomery and R. H. Myers, Confidence interval coverage for designed experiments analyzed with GLM, Journal of Quality Technology 3 (2001b), 279-292.

[15] O. O. Melo, L. A. López and S. E. Melo, Diseno de Experimentos (métodos y aplicaciones), Universidad Nacional de Colombia, Bogotá, Colombia, 2007.

[16] D. C. Montgomery, Design and Analysis of Experiments, John Wiley and Sons Inc., New Jersey, USA, 2005.

[17] D. C. Montgomery, E. A. Peck and G. G. Vining, Introduction to Linear Regression Analysis, John Wiley and Sons Inc., New Jersey, USA, 2006.

[18] R. H. Myers and D. C. Montgomery, A tutorial on generalized linear models, Journal of Quality Technology 29 (1997), 274-291.

[19] P. Paolino, Maximum likelihood estimation of models with beta-distributed dependent variables, Political Analysis 9 (2001), 325-346.

[20] S. C. Patil and H. V. Kulkarni, Analysis of 2nd factorial experiments with exponentially distributed response, Applied Mathematical Sciences 10 (2011), 459-476.

[21] C. P. Robert and G. Casella, Monte Carlo Statistical Methods, Springer, New York, USA, 2004.

[22] G. E. Shilov, Elementary Real and Complex Analysis, Dover Publications Inc., New York, USA, 1973.

[23] M. Smithson and J. Verkuilen, A better lemon-squeezer? Maximum likelihood regression with beta-distributed dependent variables, Psychological Methods 11 (2006), 54-71.

[24] K. L. P. Vasconcellos and F. Cribari-Neto, Improved maximum likelihood estimation in a new class of beta regression models, Brazilian Journal of Probability and Statistics 19 (2005), 13-31.

Appendix

Description of the method Random Walk Metropolis-Hasting

Although is difficult to order the concepts, here they are introduced according to Robert and Casella [21]. A Markov chain Monte Carlo method (MCMC), for the simulation of a distribution f is any method producing an ergodic Markov chain $(W^{(t)})$, whose stationary distribution is f. (Current use: It is possible to obtain a sample $W_1, \ldots, W_n$ approximately distributed from f without directly simulating from f). An ergodic Markov chain is a particular stochastic process. The Metropolis-Hasting (MH) is an algorithm to sample MCMCs, and is preferred rather Gibbs sampling when the prior $p(\theta)$ and the likelihood $L(\mathbf{y}|\theta)$ do not belong to the same distributional family, i.e., they are not a *conjugate pair*. The Metropolis-Hasting algorithm (MH) starts with the (target) objective density f. A conditional density $q(y|w)$, defined with respect to the dominating measure for the model, is then chosen. The MH algorithm can be implemented when $q(\cdot|w)$ is easy to simulate from and is either explicitly available or symmetric $(q(y|w) = q(w|y))$. The target density f must be available to some extent; a general requirement is that the ratio $\frac{f(y)}{q(y|w)}$ is known up to a constant independent of w. The MH algorithm associated with the target density f and the conditional density q produces a Markov chain $(W^{(t)})$ through the following transition, named Algorithm 24 in Robert and Casella [21]. Given $w^{(t)}$,

(1) Generate $Y_t \sim q(y|w^{(t)})$.

(2) Take $W^{(t+1)} = \begin{cases} Y_t & \text{with probability } \rho(w^{(t)}, Y_t), \\ w^{(t)} & \text{with probability } 1 - \rho(w^{(t)}, Y_t), \end{cases}$

where $\rho(w, y) = \min\{\frac{f(y)}{f(w)} \frac{q(w|y)}{q(y|w)}, 1\}$. The q distribution is called the *proposal* distribution and the probability $\rho(w, y)$ the Metropolis-Hasting *acceptance probability*. The random walk Metropolis-Hasting (RWMH) is a modification of the (MH) algorithm taking into account the value previously simulated to generate the following value; that is, to consider a *local* exploration of the neighbourhood of the current value of the Markov chain. The change consists in the first choice to simulate Y_t, according to

$$Y_t = X^{(t)} + \epsilon_t, \tag{A.1}$$

where ϵ_t is a random perturbation, independent of $X^{(t)}$. Then, the Markov chain in the MH algorithm associated with q is a *random walk* on the support of f density; in the RWMH method, is chosen a symmetric function g (that is, such that $g(t) = g(-t)$), which leads to the following simpler algorithm:

Given $w^{(t)}$,

(1) Generate $Y_t \sim g(|y - w^{(t)}|)$.

(2) Take $W^{(t+1)} = \begin{cases} Y_t & \text{with probability } \min\{1, \frac{f(Y_t)}{f(w^{(t)})}\}, \\ w^{(t)} & \text{otherwise.} \end{cases}$

■