# FACTORIAL EXPERIMENTAL DESIGNS AND GENERALIZED LINEAR MODEL

S. Dossou-Gbété and W. Tinsson

**Abstract.** This communication deals with the analysis of experimental designs using a generalized linear model. We prove that factorial designs provide a very simple estimation of polynomial model parameters for the Fisher scoring algorithm if we use a particular link function (called natural function). These results allow us to consider non-gaussian responses.

*Keywords:* Generalized linear model. Exponential family. Fisher-Scoring algorithm. Factorial designs. Regular fraction.

*AMS classification:* 62K15, 62J12

## §1. Introduction

Experimental designs are usually used in a linear context, *i.e.* assuming that the mean response can be correctly fitted by a linear model (polynomial of degree one or two in most cases). This assumption is often associated to the normality of the observed responses (in order to obtain an analysis of variance, a test of hypothesis, *etc*...). Then some classical configurations allow us to make the experiments at best (see the books of Box and Draper [1] or Khuri and Cornell [8]). However, it is clear that these linear assumptions are **inadequate** for some practical applications.

Then relaxing from linear model and gaussian model framework is needed. Many books and papers deal with this question (see, for example, the chapter 10 of the book of Khuri and Cornell [8] for a synthesis). But there are two main difficulties with this approach. First, the choice of a good nonlinear model is not always easy. In an other hand (assuming the nonlinear model is given) using a classical design (factorial, central composite, *etc*...) is not in general the best choice. This fact can be problematic when industrial results are first obtained with a classical design. If, at last, a linear model is inappropriate it is then impossible in general to make new experiments because they are too much expensive.

Our goal is to propose another class of solutions. These solutions have to be on one hand more general than the linear case and the gaussian framework and, on the other hand, easier to improve than nonlinear modeling. This intermediate solution consists of the choice of a **generalized linear model** (see, for example, the book of Grenn and Silverman [7]). In other words, we assume that the image of the mean response by a given "link function" can be

modelled *via* a linear relationship. Such an assumption allows us to consider any responses with a distribution in the exponential family (Bernoulli distribution, binomial distribution, Poisson distribution, Gamma distribution, *etc*...) and then we have not the restrictions of the classical linear case. These models have been studied in order to construct in general D-optimal designs (see the book of Pukelsheim [10] for the general problem of optimality). The main problem of this approach is due to the information matrix which depends on the unknown parameters of the model. Some authors have then develop Bayesian methods (see Chaloner and Larntz [4]) or constructions of robust designs (see Chipman and Welch [5] or Sebastiani and Settimi [11]) sometimes available only for a logistic regression. Our goal in this paper is to propose a general method of analysis with a simple information matrix, independent of the parameters of the model. We prove in the following that if we choose appropriate link functions then classical factorial designs can be advantageously used.

## §2. The generalized linear model

### 2.1. Definitions

We consider in the following a **generalized linear model** as it was introduced by Nelder and Wedderburn [9]. Suppose that we have $n$ observed responses $y_i$ ($i = 1...n$) associated to the independent random variables $Y_i$ with the same distribution, element of an exponential family. Denoting $m_i = E(Y_i)$, we have then a generalized linear model if and only if:

$$\forall\, i = 1...n\ ,\ g(m_i) = x_i^t \beta$$

where $x_i \in \mathbb{R}^p$ is the vector of predictors, $\beta \in \mathbb{R}^p$ is the vector of unknown parameters of the model and $g$ is the **link function** (assumed to be bijective and differentiable). Because $Y_i$ ($i = 1...n$) is an element of an exponential family we have then one of the two following likelihood:

$$\begin{aligned} & f(y_i, \theta_i, \phi) = h(y_i, \phi) \exp\left(\frac{y_i\theta_i - v(\theta_i)}{\phi}\right) \text{ with } \phi \text{ known or unknown,} \\ \text{or} \quad & f(y_i, \theta_i, \phi_i) = h(y_i, \phi_i) \exp\left(\frac{y_i\theta_i - v(\theta_i)}{\phi_i}\right) \text{ with all } \phi_i \text{ known.} \end{aligned} \tag{1}$$

Now we consider the first case with only one parameter $\phi$ but all the following results are still true with parameters $\phi_i$ instead of $\phi$. We say that $\theta_i$ is the canonical parameter of the distribution (associated to $Y_i$) and that $\phi$ is a dispersion parameter. It is usual to use the **canonical link function** which means that:

$$\forall\, i = 1...n\ ,\ g(m_i) = \theta_i.$$

Recall that for every element of an exponential family we have the following relations:

$$E(Y_i) = m_i = v'(\theta_i) \text{ and } \operatorname{Var} Y_i = \phi v''(\theta_i). \tag{2}$$

So we put $\operatorname{Var} Y_i = V(m_i)$ with $V(m_i) = \phi m_i'(\theta_i)$.

## 2.2. Estimation of the parameters

For a given generalized linear model, our problem is then to estimate the unknown parameters for the specifications of the mean. In a classic way, we can use the maximum likelihood method. Our goal is then to maximize the likelihood of the sample or (equivalently) its logarithm, that is:

$$L(y,\theta,\phi)=\frac{1}{\phi}\sum_{i=1}^{n}\left(y_i\theta_i-v\left(\theta_i\right)\right)+\sum_{i=1}^{n}\ln\left(h\left(y_i,\phi\right)\right). \tag{3}$$

The likelihood maximization involves nonlinear equation for which the solution is not in closed form. Nelder and Wedderburn [9] have proposed to use the **Fisher scoring** algorithm in order to find a numerical approximation of the maximum likelihood estimator $\hat{\beta}$. Fisher's method of scoring is the best known quasi-Newton method to solve the likelihood maximization problem (see Smyth [12]). It appear to be a special case of the Newton Raphson method when the link function is the canonical link. For the implementation of this algorithm we have to choose an initial value $\beta^{(0)}$ for the parameters of the model and then to apply iteratively the relation:

$$\forall\, k\in\mathbb{N}^*\,,\quad \beta^{(k+1)}=\beta^{(k)}+\left(X^tW^{(k)}X\right)^{-1}q^{(k)} \tag{4}$$

where $\beta^{(k)}\in\mathbb{R}^p$ is an approximation of the solution at the iteration $k$, $X$ is the model matrix (with $n$ rows and $p$ columns), $W^{(k)}=W\left(\beta^{(k)}\right)$ and $q^{(k)}=q\left(\beta^{(k)}\right)$ are such that $W(\beta)=diag(\omega_i\,,i=1...n)$ and $q(\beta)\in\mathbb{R}^p$ satisfy:

$$\omega_i=\frac{1}{\operatorname{Var}Y_i}\left(\frac{\partial m_i}{\partial\eta_i}\right)^2 \text{ with } \eta_i=g\left(m_i\right) \text{ and } q(\beta)=\left(\frac{\partial L(y,\theta,\phi)}{\partial\beta_j}\right)_{j=1...p}.$$

Note that the matrix $W$ has to be computed at every iteration (matrix $W^{(k)}$) because it depends on $m_i$ and $m_i=g^{-1}(x_i^t\beta)$ depends on the value of the approximation of the solution at the iteration $k$ (vector $\beta^{(k)}$).

*Remark.* It is also possible to find a vector $z^{(k)}$ such that relation $(4)$ becomes:

$$\beta^{(k+1)}=\left(X^tW^{(k)}X\right)^{-1}X^tW^{(k)}z^{(k)}.$$

In other words, the Fisher scoring algorithm is also an iteratively reweighted least squares (IRWLS) method.

## 2.3. Modified Fisher's method of scoring

Our goal is now to simplify the algorithm of Fisher scoring by dropping out he diagonal weighting matrix $W$. This can be done by a judicious choice of the link function. In fact our objective is :

$$W=I_d\Leftrightarrow\ \forall\, i=1...n\,,\ \frac{1}{\operatorname{Var}Y_i}\left(\frac{\partial m_i}{\partial\eta_i}\right)^2=1. \tag{5}$$

But we know, from relation $(2)$, that $\operatorname{Var}Y_i=V\left(m_i\right)$. Then $m_i=g^{-1}\left(\eta_i\right)$ implies that:

$$(5)\Leftrightarrow\frac{\partial m_i}{\partial\eta_i}=\sqrt{V\left(m_i\right)}\Longleftrightarrow\frac{1}{g'\left(m_i\right)}=\sqrt{V\left(m_i\right)}.$$

Our proposal relies on the following lemma:

**Lemma 1.** The matrix $W$ is the identity matrix if and only if the link function $g$ satisfies:

$$\forall\, i = 1...n\,,\, g'\left(m_i\right) = V^{-1/2}\left(m_i\right).$$

Such a function is then called **natural**

| *Distribution of* $Y_i$ | *Function* $V$ | ***Natural link fct.*** | *Canonical link fct.* |
|---|---|---|---|
| Bernoulli $(p)$ | $t\left(1-t\right)$ | $\sqrt{n}\arcsin\left(2t-1\right)$ | $\ln\left(\dfrac{t}{1-t}\right)$ |
| Binomial $\mathcal{B}\left(n,p\right)$ | $t\left(1-\dfrac{t}{n}\right)$ | $\sqrt{n}\arcsin\left(\dfrac{2t}{n}-1\right)$ | $\ln\left(\dfrac{t}{n-t}\right)$ |
| Neg. Bin. $(n,p)$ | $t\left(\dfrac{t}{n}+1\right)$ | $\sqrt{n}\arg\cosh\left(\dfrac{2t}{n}+1\right)$ | $\ln\left(\dfrac{t}{n+t}\right)$ |
| Poisson $\mathcal{P}\left(\lambda\right)$ | $t$ | $2\sqrt{t}$ | $\ln t$ |
| Gamma $\mathcal{G}\left(a,p\right)$ | $\dfrac{t^2}{p}$ | $\sqrt{p}\ln t$ | $\dfrac{p}{t}$ |

**Table 1:** Natural link function for different distributions.

Table 1 gives, for some exponential families of distributions, the natural link functions (depending on $t$) verifying the differential equations of lemma 1 (with the additive constant chosen to be zero). We also recall in this table the classical canonical link function.

*Remark.* We have seen in subsection 2.2 that the Fisher scoring algorithm is in fact an iteratively reweighted least squares method. Then, the use of the natural link function allows us to have simply an iteratively ordinary least squares method.

The algorithm of Fisher scoring needs also the use of a vector $q$ such that $q\left(\beta\right) = \left(\partial L\left(y,\theta,\phi\right)/\partial\beta_j\right)$ for $j = 1...p$ (see subsection 2.2). In a classical way, we have the following relation (from the chain rule):

$$\frac{\partial L}{\partial\beta_j} = \frac{\partial L}{\partial\theta_i}\frac{\partial\theta_i}{\partial m_i}\frac{\partial m_i}{\partial\eta_i}\frac{\partial\eta_i}{\partial\beta_j}.$$

Then we obtain immediately for the likelihood of every sample of the exponential family (see formula (3)):

$$\forall\, j = 1...p\,,\, \frac{\partial L}{\partial\beta_j} = \sum_{i=1}^{n}\frac{\left(y_i - m_i\right)}{\operatorname{Var} Y_i}\frac{\partial m_i}{\partial\eta_i}\left[X\right]_{ij}$$

where $\left[X\right]_{ij}$ is the element of row $i$ and column $j$ of the matrix model $X$. This general relation can be simplified in our case because we have:

$$\eta_i = g\left(m_i\right) \text{ with } g'\left(m_i\right) = V^{-1/2}\left(m_i\right) \text{ so } \frac{\partial m_i}{\partial\eta_i} = \sqrt{V\left(m_i\right)}.$$

Thus, we can state the following lemma:

**Lemma 2.** *If the link function is the* **natural** *link function the vector* $q$ *is then defined by:*

$$\forall\, j = 1...p\,,\, \frac{\partial L}{\partial \beta_j} = \sum_{i=1}^{n} [X]_{ij}\, y_i^* \text{ with } y_i^* = \frac{y_i - m_i}{\sqrt{V(m_i)}}.$$

We see from lemma 2 that the vector $q$ has then a very simple expression when the natural link function is used. It needs only the observations $y_i^*$ in their standardized and centered form.

## §3. Application to factorial designs

### 3.1. Definitions

Consider a random phenomenon of $m$ factors that may be checked by the experimenter. We have seen that the choice of the natural link function allows us to put the matrix ${}^tXX$ in place of the initial matrix $X^tWX$ in the algorithm of Fisher scoring. So, our goal is now the find an experimental configuration such that $X^tX$ has a simple form. The optimal situation is reached when $X^tX$ is a diagonal matrix (*i.e.* when the design is orthogonal). We consider in the following the two models given below:

$$(L)\ ,\ \forall\, i = 1...n\ ,\ g(m_i) = \beta_0 + \sum_{j=1}^{m} \beta_j x_{ij}\ ,$$

$$(I)\ ,\ \forall\, i = 1...n\ ,\ g(m_i) = \beta_0 + \sum_{j=1}^{m} \beta_j x_{ij} + \sum\sum_{j<l} \beta_{jk} x_{ij} x_{il}.$$

$(L)$ and $(I)$ are then two generalized linear models with a polynomial linear part, of degree one for $(L)$ and of degree two with interactions for $(I)$. Classicaly, we assume that the variables $x_{ij}$ are coded in a such ways that their values always belong to the interval $[-1, 1]$. In other words, if a factor is associated to a variable $X$ with extreme levels $X_{\min}$ and $X_{\max}$, the associated coded variable $x$ is obtained by the following transformation (see the chapter 2 of Khuri and Cornell [8]):

$$x = \frac{2X - (X_{\max} + X_{\min})}{X_{\max} - X_{\min}}.$$

Then it is well known that every complete **factorial design** (*i.e.* constituted by all the vertices of the cube $[-1, 1]^m$) gives us an orthogonal configuration for the two models $(L)$ and $(I)$. Nevertheless, using such designs is not possible when the number of factors $m$ becomes high (because of the $2^m$ experimental units). So we also consider in the following some **regular fractions** of these factorial designs (see Box and Hunter [2] and [3]). In other words, we are now working with configurations given by:

1) $2^{m-k}$ vertices of the cube $[-1, 1]^m$,
2) $n_0$ central replications of the experimental domain.

We know that the matrix model $X$ is of full rank (*i.e.* $X^tX$ is regular) if and only if we used a regular fraction of resolution at least $III$ for the model $(L)$ and at least $V$ for the model $(I)$.

## 3.2. Iterative treatment

In this part we are going to find the simplified form of the Fisher scoring algorithm used with the natural link function and a complete factorial design or a regular fraction of it. We denote by $n$ the number of experiences (*i.e.* $n = 2^{m-k} + n_0$) and by $D$ the design matrix (*i.e.* the $n \times m$ matrix with the row $i$ made up from the $m$ coordinates of the $i$-th design point, so the matrix of the $x_{ij}$). Let $D_j$ ($1 \leq j \leq m$) be the $j$-th column of this matrix and $\odot$ the Hadamard product operator (also called term to term product because if $u, v \in I\!R^n$ then $u \odot v \in I\!R^n$ verifies $\forall$ $i = 1 \dots n$ , $(u \odot v)_i = u_i v_i$). Denoting $Q_{jl} = D_j \odot D_l$ ($1 \leq j < l \leq m$), the model matrix is then:

$$X = [1_n \mid D_1 \dots D_m] \text{ for the model } (L)\,,$$
$$X = \left[1_n \mid D_1 \dots D_m \mid Q_{12} \dots Q_{(m-1)m}\right] \text{ for the model } (I)\,.$$

Then it is well known that, if we use appropriate regular fraction (such that $X$ is a full rank matrix), these factorial designs are orthogonal for the model $(L)$ or $(I)$ and we have:

$$X^t X = \text{diag}\left(2^{m-k} + n_0, 2^{m-k}, \dots, 2^{m-k}\right).$$

This result implies, combined with relation $(4)$ and lemma $1$ and $2$, the following form for the components of the unknown vector $\beta$ in the Fisher scoring algorithm:

**Proposition 3.** Consider the model $(L)$ or $(I)$ used with the natural link function. For a complete factorial design or a regular fraction of resolution at least $III$, the Fisher scoring algorithm is given for the model $(L)$ by (with $k \in \mathbb{N}^*$)*:*

$$1) \quad \beta_0^{(k+1)} = \beta_0^{(k)} + \frac{1}{2^{m-k} + n_0} \sum_{i=1}^{n} y_i^*,$$
$$2) \quad \forall\, j = 1 \dots m\,,\ \beta_j^{(k+1)} = \beta_j^{(k)} + \frac{1}{2^{m-k}} \sum_{i=1}^{n} x_{ij} y_i^*.$$

For a complete factorial design or a regular fraction of resolution at least $V$, the algorithm of Fisher scoring for the model $(I)$ verifies, in addition of the two previous relations:

$$3) \quad \forall\, j, l = 1 \dots m \text{ with } j < l\,,\ \beta_{jl}^{(k+1)} = \beta_{jl}^{(k)} + \frac{1}{2^{m-k}} \sum_{i=1}^{n} x_{ij} x_{il} y_i^*.$$

*Remark.* These three results have a very simple expression. We can also note that factorial designs have only two levels, so the values for the coded variables $x_{ij}$ are only $-1$, $0$ (for possible central experiments) or $1$.

## 3.3. Dispersion of the estimations

We know (see Green and Silverman [6]) that asymptotically the maximum likelihood estimator of $\beta$ has a Gaussian distribution and a dispersion given by:

$$\text{Var}\,\hat{\beta} = \phi \left(X^t W X\right)^{-1}.$$

If $\phi$ is unknown then it can be estimated by means of pearson statistics. This result is very interesting in our case because we know that $X^t W X$ is a diagonal matrix and the diagonal elements are given in the last subsection. So we have the following proposition:

**Proposition 4.** *C*onsider the model $(L)$ or $(I)$ used with the natural link function and a complete factorial design or an appropriate regular fraction (of resolution at least $III$ for $(L)$ and at least $V$ for $(I)$). The maximum likelihood estimator $\hat{\beta}$ satisfies asymptotically the following properties: its components are **non-correlated**, and its dispersion is given by:

$$\text{Var}\,\hat{\beta}_0 = \frac{\phi}{2^{m-k}+n_0} \text{ and } \forall\, j,l = 1...m\ j < l\ , \text{Var}\,\hat{\beta}_j = \text{Var}\,\hat{\beta}_{jl} = \frac{\phi}{2^{m-k}}.$$

*Remark.* The dispersion parameter $\phi$ is needed in order to obtain these different dispersions. This is not a serious problem in practice because the observations are often associated to usual dispersions and then $\phi$ has a simple form (for example, $\phi = 1$ for a binomial distribution, a Poisson distribution, a negative-binomial distribution...).

## 3.4. Implementation of the algorithm

The implementation of the Fisher scoring algorithm is then very simple in our case, we only have to apply iteratively results from proposition 3. No use of matricial calculus is needed, and in particular we do not have to invert any matrix. The development of the algorithm is given by figure 1. Note that the algorithm has to be initialized by judicious values for $\beta^{(0)}$. This can be done, for example, by a classic linear regression on the transformed response (*i.e.* on the $g(y_i)$ with $g$ natural link function in place of the $y_i$).

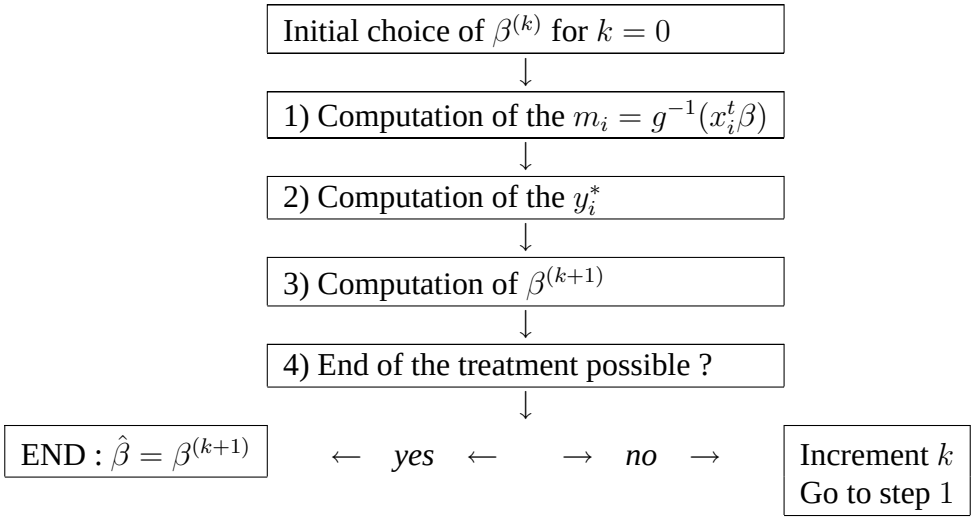


**Figure 1** : Fisher scoring algorithm.

This iterative algorithm must also be stopped (step 4) and we can use many different conditions to do that. It can be stopped when the likelihood seems to be non-increasing (*i.e.* when $L_{\max}^{(k+1)} < L_{\max}^{(k)}$), when the likelihood seems to be constant (*i.e.* when $|L_{\max}^{(k+1)} < L_{\max}^{(k)}| < \varepsilon$ with $\varepsilon$ small positive real), when the estimated parameters seem to be constant (*i.e.* when $||\beta^{(k)} - \beta^{(k+1)}|| < \varepsilon$ where $\|.\|$ is a chosen norm), *etc*...

## §4. Example: application to the Bernoulli distribution

Consider a first application with the case, very usual in practice, of binary responses. It corresponds, for example, to a medical experience where our goal is to predict the appearance or not of a distension for the blood vessels in accordance with the injection of three different constituents (called factors one, two and three in the following). We assume that every observed response $y_i$ is a realization of a Bernoulli distribution of parameter $p_i$ (unknown) and we denote $0$ when there is no distension, $1$ in the other case. Recall that such a distribution belongs to the exponential family because its density satisfies relation $(1)$ with:

$$\theta_i = \ln \frac{p_i}{1-p_i}\ ,\ v(\theta_i) = -\ln\left(1+e^{\theta_i}\right)\ ,\ \phi = 1 \text{ and } h(y_i,\phi) = \mathbb{I}_{\{0,1\}}(y_i,\phi)\,.$$

From the results of subsection $2.3$ the natural link function associated to this distribution is $g$ with $g(t) = \arcsin(2t-1)$ and then we can consider the following model (with $m_i = \mathrm{E}\,Y_i = p_i$):

$$\forall\, i = 1...n\ ,\ \arcsin(2m_i - 1) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3}.$$

We assume that the experimenter has made the experiments corresponding to a complete factorial design with two center points (the low number of factors allow us to consider the complete design in our case). We have then a total of $10$ experiences given in table $2$. Recall that the factors are in the coded form, so $-1$ is associated to the lowest quantity injected and $+1$ is associated to the biggest ($0$ is then associated to the mean of these two quantities). Responses given in table $2$ are obtained by the following formula for the $p_i$ parameters:

$$\forall\, i = 1...n\ ,\ p_i = 0.2x_{i1} - 0.1x_{i2} - 0.1x_{i3} + 0.6.$$

In other words, we assume that the probabilities associated to each Bernoulli distribution can be correctly fitted by a Taylor series of order one in the experimental domain. We also assume that the effects of factors $2$ and $3$ are opposite (and lower) to the effect of the factor $1$ on the response. Table $2$ gives the probabilities $p_i$ associated to each experimental unit (column $p_i$) and simulated results for the different responses (column $y_i$).

| Exp | Fac. 1 | Fac.2 | Fac. 3 | $p_i$ | $\mathbf{y}_i$ | $\hat{p}_i$ | $\hat{\mathbf{y}}_i$ |
|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 0.60 | **1** | 0.54 *(0.75)* | **1** *(1)* |
| 2 | −1 | 1 | 1 | 0.20 | **0** | 0.27 *(0.00)* | **0** *(0)* |
| 3 | 1 | −1 | 1 | 0.80 | **1** | 1.00 *(1.00)* | **1** *(1)* |
| 4 | 1 | 1 | −1 | 0.80 | **1** | 1.00 *(1.00)* | **1** *(1)* |
| 5 | −1 | −1 | 1 | 0.30 | **0** | 0.03 *(0.00)* | **0** *(0)* |
| 6 | −1 | 1 | −1 | 0.30 | **0** | 0.03 *(0.00)* | **0** *(0)* |
| 7 | 1 | −1 | −1 | 1.00 | **1** | 0.59 *(1.00)* | **1** *(1)* |
| 8 | −1 | −1 | −1 | 0.60 | **1** | 0.60 *(0.75)* | **1** *(1)* |
| 9 | 0 | 0 | 0 | 0.60 | **1** | 0.57 *(0.75)* | **1** *(1)* |
| 10 | 0 | 0 | 0 | 0.60 | **0** | 0.57 *(0.75)* | **1** *(1)* |

**Table 2** : Results for the complete factorial design.

Now, we are going to fit a generalized linear model. If we have no information concerning the choice for the initial values of the algorithm, we can take (for example):

$$\beta_0^{(0)} = 1 \text{ , } \beta_1^{(0)} = \beta_2^{(0)} = \beta_3^{(0)} = 0.$$

Then the iterative treatment of proposition $3$ leads us very quickly (in two iterations) to a maximal likelihood obtained for:

$$\hat{\beta}_0 = 0.143 \text{ , } \hat{\beta}_1 = 1.376 \text{ , } \hat{\beta}_2 = -0.719 \text{ and } \hat{\beta}_3 = -0.719.$$

In other words, the best model satisfies ($\forall\, x_1, x_2 \in [-1, 1]$):

$$\hat{p}\,(x_1, x_2) = \frac{\sin\left(0.143 + 1.376x_1 - 0.719x_2 - 0.719x_3\right) + 1}{2}.$$

Predicted values of the probabilities $p_i$ are given in table $2$ (column $\hat{p}_i$) with the predicted responses (column $\hat{y}_i$), that is the values of $\hat{p}_i$ rounded to the nearest integer. We also present, in brackets, results obtained by the classical analysis with the canonical link function (these results come from the SAS software). We observe the good global quality of the results since observed responses $y_i$ and predicted responses $\hat{y}_i$ are always the same (except, of course, for the two last experiences where it is impossible to predict at once $0$ and $1$). If we consider probabilities $p_i$ associated to Bernoulli distributions we note, on one hand, that predictions are very good for half of the experiments (*i.e.* experiences $1, 2, 8, 9$ and $10$). On the other hand, these results are not so good for experiences $3$ and $4$ and they are bad for experiences $5, 6$ and $7$. These problems of prediction are principally due to the small number of experiences, and also to the nature of the responses which give poor informations because they have only two levels. We can finally note that the adjusted model allows us to find again the correct effect of each factor (*i.e.* the factor $1$ has a preponderant effect on the response and factors $2$ and $3$ have equal effects, opposite to factor $1$).

Note also that the choice of this natural link function gives us a robust estimation procedure. Indeed we have verified with this example and also with other distributions (see Dossou-Gbete and Tinsson $[6]$ for more details) that estimations using the natural link function are **very closed** to estimations using the classical canonical link function and moreover the natural link function allows us, in general, to obtain faster convergence rate.

## References

[1] BOX G. AND DRAPER N. *Empirical Model-Building and Responses Surfaces*. John Wiley, New-York, 1987.

[2] BOX G. E. P. AND HUNTER J. S. The $2^{k-p}$ fractionnal factorial designs, Part $I$. Technometrics, 3 (1961), 311-351.

[3] BOX G. E. P. AND HUNTER J. S. The $2^{k-p}$ fractionnal factorial designs, Part $II$. Technometrics, 3 (1961), 449-458.

[4] CHALONER K. AND LARNTZ K. Optimal bayesian design applied to logistic regression experiments. *Journal of Statistical Planning and Inference*, 21 (1989), 191-208.

[5] CHIPMAN H.A. AND WELCH W.J.. D-Optimal Design for Generalized Linear Models. *Unpublished.* http://www.stats.uwaterloo.ca/~hachipma/publications.html.

[6] DOSSOU-GBÉTE S. AND TINSSON W. Factorial experimental designs ans generalized linear model. *Submitted* (2004).

[7] GREEN P.J. AND SILVERMAN B.W. *Nonparametric Regression and Generalized Linear Models*. Chapman & Hall, Monographs on Statistics and Applied Probability, 58, 1994.

[8] KHURI A. AND CORNELL J. *Responses Sufaces : Designs and Analyses*. Dekker, Statistics: textbooks and monographs, 152, New-York 1996.

[9] NELDER J.A. AND WEDDERBURN R.W.M. Generalized linear models. *Journal of Royal Statistical Society,* A, 135 (1972), 370-384.

[10] PUKELSHEIM F. *Optimal Designs of Experiments*. John Wiley, New-York, 1993.

[11] SEBASTIANI P. AND SETTIMI R. A note on D-optimal designs for a logistic regression model. *Journal of Statistical Planning and Inference*, 21 (1997), 359-368.

[12] SMYTH G.K. Optimization. *Encyclopedia of Environmetrics*, Vol. 3 (2002), 1481-1487.

Simplice Dossou-Gbété and Walter Tinsson
Laboratoire de Mathématiques Appliquées, FRE 2570
Université de Pau et des Pays de l'Adour
IPRA, BP1155, 64013 Pau Cedex, France
simplice.dossou-gbete@univ-pau.fr and walter.tinsson@univ-pau.fr