BAYES ESTIMATE OF THE EQUILIBRIUM PARAMETER IN A MARKOV CHAIN: A MEDICAL APPLICATION

Carlos Alberto de Bragança PEREIRA*
André ROGATKO**

■ ABSTRACT: The results of a temporal study of 30 diabetic pregnant women are analyzed using a Markov Chain model and a Bayesian estimator for the parmeter of interest, the invariant probability measure, is obtained. It is shown that cartdiotocography results of a fetus are positively dependent on the mother's glycemic levels. Then, these levels can be used as a calibration tool for the expensive cardiotocography exam.

■ KEYWORDS: Trasition probability; invariant probability; high likelihood region.

1 Introduction

Fetal oxygenation is evaluated through cardiotocography – a very expensive method. In diabetic pregnant women the results of these tests, as well as glycemic levels, are very unstable and may suffer daily fluctuations. The aim of this study is to prove that in pregnant women with diabetes the glycemic level in the mother may affect the cardiotocography result. In more precise terms, what one wishes to prove is that the chance of a fetus having insufficient oxygenation increases when the mother has an abnormal glycemic level, and decreases when such level returns to normal.

The present study is based on data from 30 patients, diabetic pregnant women, submitted to weekly simultaneous tests of fetal oxygenation (cardiotocography) and glycemic level. The data are presented in Section 2 where some details are discussed.

According to the aims above, the paper developes a methodology to estimate, when selecting a diabetic pregnant women, the probability of finding a concordance between the results of the two tests: mother's glycemic level and fetal oxygenation. Concordance occurs when both levels.are normal or when both are abnormal. When contrary, the levels are said to be discordant.

* Instituto de Matemática e Estatística – USP – 05389-970 – São Paulo – SP – Brazil.
** Fox Chase Cancer Center – 7701 Burholme Av. – Philadelphia, PA 19111.

Section 2 describes the data, Section 3 describes the model, Section 4 presents the Bayesian numerical analysis, and final remarks are presented in Section 5. The Appendix presents some interesting analytical expressions.

2 Data description

Tests were performed in 30 diabetic pregnant patients from the Gynecology and Obstetrics Department, Faculty of Medicine, University of São Paulo, under the supervision of Doctor Marcelo Zugaib. The tests were done in the consecutive weeks in which the patients were under treatment. The patient was tested weekly till the end of or when she abandoned treatment.

Since the sample is not sufficiently large, the time of the beginning of tests was not controlled. Thus, the periods of pregnancy, as well as times of the year or patients' age, were not necessarily coincident. The only pre-fixed factor is a diabetic pregnancy.

Table 1 shows the data by patient, by week. This table includes also the likelihood factor of each patient. The "status" of the tests are represented by the letters N, A, S, and I, where N = normal glycemia, A = abnormal glycemia, S = sufficient fetal oxygenation, and I = Insufficient fetal oxygenation. Consequently, the pregnant woman after being submitted to the pair of tests is classified in one of the following ways: NS, NI, AS, and AI. The tests are considered concordant, C, if either NS or AI occurs, and discordant, if NI or AS occurs.

3 Model description

The two possible states of interest for a specific patient in a specific week are $C = \{\text{NS, AI}\}$, the concordant state, and $C' = \{\text{NI, AS}\}$, the discordant state.

Let $\pi = P(C)$ be the probability that the pair of tests concord when applied to a woman randomly chosen from a group of pregnant diabetic women. In such patient the weekly transition probabilities are represented by:

$p = P(C \to C)$ conditional probability of concordance in a week given that there was concordance in the previous week;

$1 - p = P(C \to C')$ conditional probability of discordance in a week given that there was concordance in the previous week;

$q = P(C' \to C')$ conditional probability of discordance in a week given that there was disconcordance in the previous week;

$1 - q = P(C' \to C)$ conditional probability of concordance in a week given that there was discordance in the previous week.

Table 1 – Classification of tests per patient/week

Patient t	Weeks							Number of Weeks	Likelihood Factor
	1st	2nd	3rd	4th	5th	6th	7th		
01	NS	NS	AI					3	πp^2
02	NS	NI	NS	AI				4	$\pi(1-p)(1-q)p$
03	NS	NS	NS	AS	NS			5	$\pi p^2(1-p)(1-q)$
04	NS	NS	NS	NS	NS	NS		6	πp^5
05	NS	NS	NS	NS				4	πp^3
06	NS	NS						2	πp
07	NS	NI	NS	NS				4	$\pi(1-p)(1-q)p$
08	NS	NS	NS	NS	NS	NS		6	πp^5
09	NS	NS	NS					3	πp^2
10	NS	NS	NS	NS				4	πp^3
11	NI	NS	NS	NS	NS			5	$(1-\pi)(1-q)p^3$
12	AI	AI						2	πp
13	NS	NS	NS	NS				4	πp^3
14	NS	NS	NS	NS	NS	NS		6	πp^5
15	AI	AI	AI	AI				4	πp^3
16	AS	NS	NS	NS	NS			5	$(1-\pi)(1-q)p^3$
17	NS	NS	NS	AI	AI	AI		6	πp^5
18	AI	AI	AI	NS	NS			5	πp^4
19	AI	AI	NS	NS	NS	NS		6	πp^5
20	NS	NS	NS					3	πp^2
21	NI	NI	NI					3	$(1-\pi)q^2$
22	NI	NI	NI	NI	NI			5	$(1-\pi)q^4$
23	AS	AS	AS	AS				4	$(1-\pi)q^3$
24	AS	AS						2	$(1-\pi)q$
25	NS	NS	AS	AS				4	$\pi p(1-p)q$
26	NI	NI	NI					3	$(1-\pi)q^2$
27	NS	NI	NI	NI	AI	NS		6	$\pi(1-p)q^2(1-q)p$
28	AI	NI	NS	NS	AI	AI	NS	7	$\pi(1-p)(1-q)p^4$
29	NI	NI	NS					3	$(1-\pi)q(1-q)$
30	NI	NS	NS	NS	NS			5	$(1-\pi)(1-q)p^3$

The model to be considered from now on is a Markov Chain with two states (Feller, 1968). Each observation is a piece of the process defined by this chain. A defining condition of a Markov model is that the results of two weeks are conditionally independent given the result of an intermediate week. Another important restriction is to consider p and q as independent of the steps (time) of the process.

The probability π defined above is the well known invariant measure of the process.

Summing up and assuming p and q are neither 0 nor 1, the model adopted here is a Markov Chain with transition matrix

$$P = \begin{pmatrix} p & 1-p \\ 1-q & q \end{pmatrix} \tag{3.1}$$

with the invariant probability measure a solution of the following equation:

$$[\pi,.(1-\pi)]\, P = [\pi,\ (1-\pi)] \tag{3.2}$$

It is,

$$\pi = \frac{1-q}{(1-q)+(1-p)} \tag{3.3}$$

satisfying as well

$$1-q = (1-p)\frac{\pi}{1-\pi} \tag{3.4}$$

Each patient studied corresponds to an independent realization of a part of the chain. Since 30 distinct patients are considered, 30 independent observations of parts of the chain are obtained.

The likelihood corresponding to each individual observation is described in the last column of Table 1. Hence the observations being independent, the overall likelihood is given by the product of these functions, that is

$$L = \pi^{21}(1-\pi)^{9}p^{68}(1-p)^{6}(1-q)^{9}q^{16}, \tag{3.5}$$

where $0 < p < 1$, and $0 < q < 1$ and π satisfie equation (3.3).

From (3.4), we finally have

$$L(\pi, p) = \pi^{30} p^{68} (1-p)^{15} \left\{ 1 - \frac{\pi(1-p)}{1-\pi} \right\} , \qquad (3.6)$$

where $0 < p < 1$ and $0 < (1-p)\frac{\pi}{(1-\pi)} < 1$ since $0 < (1-q) < 1$. Hence, the parametric space for (π, p) is defined by

$$\Omega = \{(\pi, p);\ 0 < \pi < 0.5 \text{ and } 0 < p < 1\} \cup \left\{(\pi, p);\ 0.5 < \pi < 1 \text{ and } \left(2 - \frac{1}{\pi}\right) < p < 1\right\},$$

that is represented by the left hand side in Figure 1. Figure 2 presents the smallest region of the likelihood function that contains approximately an area, under the likelihood function, of 95% of the total area. Taking the partial derivatives of the logarithm of the likelihood function (3.6) and using the Broyden System Solver we obtain the maximum likelihood estimator of the vector (π, p) as

$$(\hat{\pi}, \hat{p}) = (0.755857,\ 0.898571).$$

The computations of this section, including the set of high likelihood presented in Figure 2, were obtained by a discretization of function (3.6) where we considered the variation from 0 to 1 with steps of 0.01 for both parameters π and p.

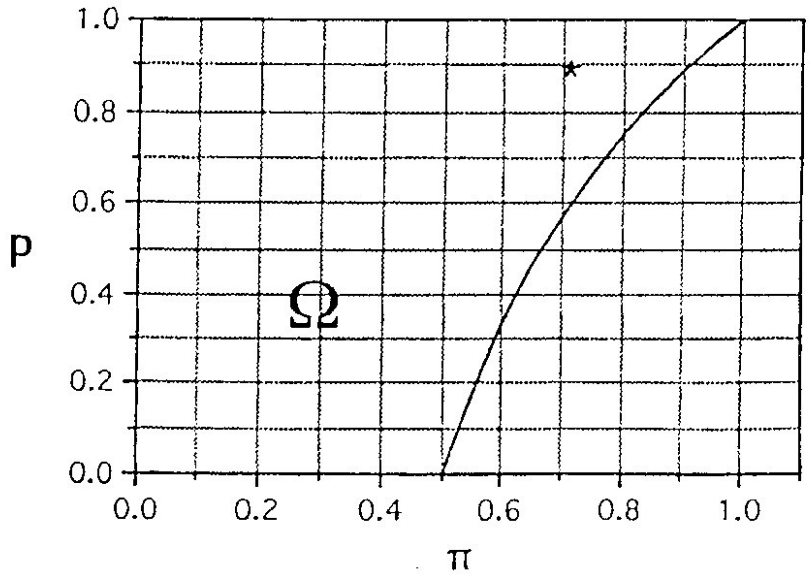

FIGURE 1 – Parametric space of (π,p) and the maximum likelihood estimate.

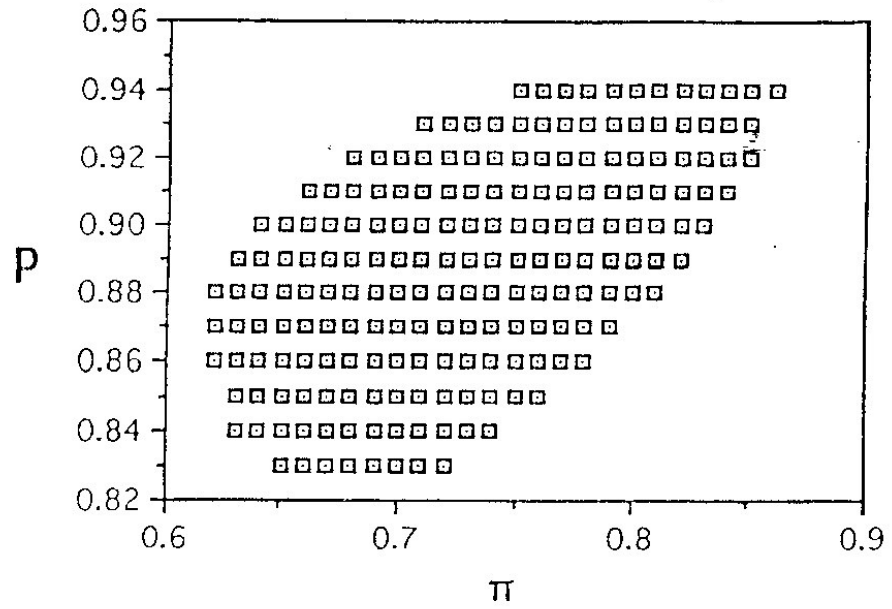

FIGURE 2 - 95% higher likelihood region.

4 Bayesian estimation

Up to this point, a classical procedure of analysis was used. That is, a model is considered and based on it, the likelihood function is obtained and the maximum likelihood estimate is computed. The problem arises when one needs to describe the precision of the estimate. In a problem like the one discused here, it is not easy to describe the sample space since the stopping rule cannot be simply defined. Consequently, the sample distribution and the moments of the estimators cannot be simply defined. Hence, standard frequentist analyses do not apply. Cases like the one described are non-problems for a Bayesian statistician however. Bayesian analysis nedds only the likelihood function to calibrate the prior probability function that represents the uncertainty of the scientist about the parameter involved [(p, q) for instance].

It should be noted that the "credible" set presented in Figure 2 can be interpreted as Bayesian if we consider that a priori (π, p) has a uniform distribution over the unit square. Also the maximum likelihood estimate would be a Bayesian point estimate since it is the mode of the posterior obtained from this uniform prior calibrated by the likelihood. This prior however is not realistic since the support of a distribution for (π, p) must be contained in the space presented by Figure 1, that is different from a unit rectangle. Consequently, π and p cannot be independent. Now we describe a simple Bayesian coherent estimate of π, the parameter of interest.

Let the uncertainty about the transition probabilities, p and q, be represented by two independent Beta distributions with parameters $(a + 1, b + 1)$ and $(c + 1, d + 1)$, respectively. This independence restriction is adequate as the parameters p and q are

related to two different subpopulations. It is not difficult to see that the posterior density for (p, q) is proportional to

$$f(p, q) = \frac{p^{a+68}(1-p)^{b+15}q^{c+16}(1-q)^{d+30}}{[(1-p)+(1-q)]^{30}}, \quad 0 < p, q < 1. \tag{4.1}$$

Looking at the parameter of interest, π, we consider the following parametrization:

$$\pi = \frac{1-q}{(1-q)+(1-p)} \quad \text{and} \quad \theta = (1-q)+(1-p) \tag{4.2}$$

whose Jacobian is equal to θ. We obtain the posterior density of (π, θ) by a transformation of (4.1). The result is a joint density that is proportional to the following function:

$$g(\pi, \theta) = \pi^{d+30}(1-\pi)^{b+15}\theta^{b+d+16}(1-\theta\pi)^{c+16}[1-\theta(1-\pi)]^{a+68}, \tag{4.3}$$

whose support is $\Omega^* = \{(\pi, \theta);\ \theta < (1-\pi)^{-1}$ if $0 < \pi < 0.5$ and $\theta < \pi^{-1}$ if $0.5 < \pi < 1\}$. This set is displayed in Figure 3.

To illustrate the procedure that will follow after obtaining the joint density (4.3), we consider the uniform (over the unit square) prior distribution for (p, q), that is, $a = b = c = d = 0$.

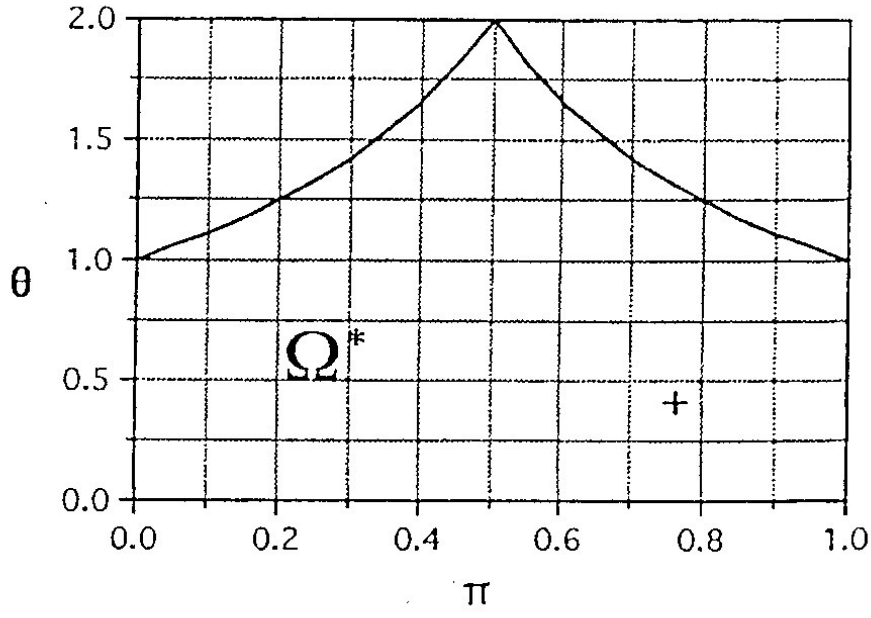

FIGURE 3 - Mode and support of the posterior density with the mode of this distribution

To estimate the parameter of interest we may consider the mode of the density (4.3) using the same method we have used to obtain the maximum likelihood estimates given in Section 3. The posterior mode for the uniform prior is

$$(\pi', \theta') = (0.7593, 0.4383). \tag{4.4}$$

By a simple numerical (Romberg's) method, we obtain the posterior marginal density of π, the parameter of principal interest. This density is presented in Figure 4. The mode of this density is also 0.76, that agrees with the joint mode (4.4). An exact method to obtain this marginal is described in Appendix.

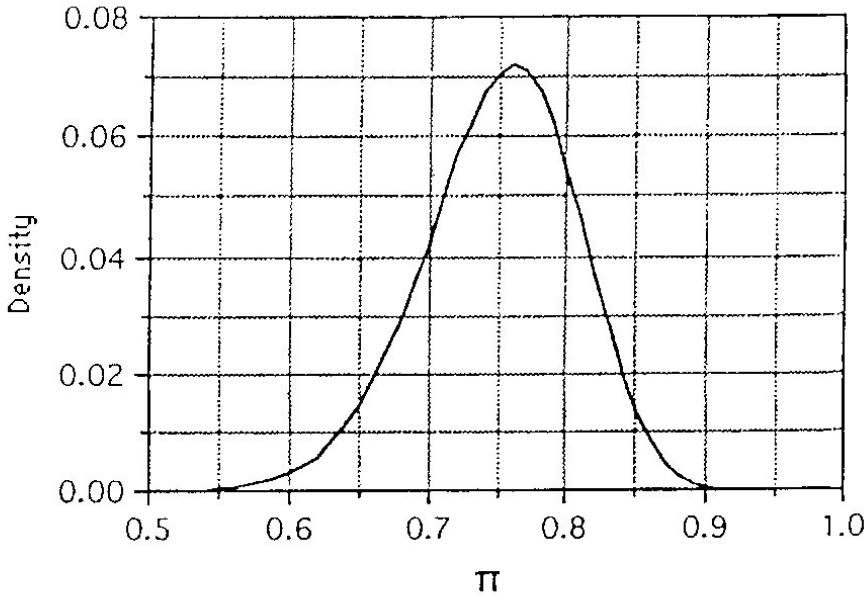

FIGURE 4 - Posterior marginal distribution of π.

Another common estimate is the posterior mean that, in our case is $E(\pi \mid data)$ = 0.75. The posterior standard deviation is 0.05916. The following credible intervals are also of same interest:

95.4% credible interval	(0.64; 0.85)
99% credible interval	(0.60; 0.87)

Another important fact that we can obtain from the posterior density is that the support of this density is (almost surely) in the upper half of the unit interval. That is, the two kinds of medical evaluation, glycemia and cardiotocography, agree with probability greater than 0.60.

5 Final remarks

The Bayes estimate of π is 0.75 and the 95.4% credible interval is (0.64; 0.85). This is a strong evidence that the value of p is greater than 0.5. Translating into the language of the medical problem, it means that the probability of concordance in the pair of tests is high. Thus, if a patient shows abnormal glycemia, at a given time, the fetal oxygenation at that moment has a high probability of being insufficient. On the other hand, a normal glycemia indicates, with high probability, normal results in cardiotocography, which reflects the fetal oxygenation.

A better schedule for collecting the samples, a larger number, and a stricter control of patients would contribute to a better evaluation of the probabilities. This paper, however, emphasizes the fact that the clinician should be concerned with the well-being of the fetus whenever the mother shows an abnormal glycemia.

Acknowledgement

To Dr. M. Zugaib from the University of São Paulo, Medical School, who kindly provided the data analyzed.

PEREIRA, C. A. de B., ROGATKO, A. Estimativa Bayesiana do parâmetro de equilíbrio em uma Cadeia de Markov: uma aplicação médica. *Rev. Mat. Estat. (São Paulo)*, v.14, p.103-112, 1996.

▪ *RESUMO: Os resultados de um estudo temporário com trinta mulheres grávidas e diabéticas são analisados usando um modelo de Cadeia de Markov e uma estimativa bayesiana para os parâmetros de interesse, a medida invariante de probabilidade é obtida. Mostramos que os resultados da cardiotocografia do feto é positivamente dependente do nível de glicemia da mãe. Assim, esses níveis podem ser usados como um instrumento de calibração para a cardiotocografia, que é um exame caro.*

▪ *PALAVRAS-CHAVE: Probabilidade de transição; medida invariante de probabilidade; região de máxima verossimilhança.*

References

1 FELLER, W. *A introduction to probability theory and its applications*. New York: v.1. John Wiley, 1968. v.1.

2 ZACKS, S. *Parametric statistical inference:* Basic theory and modern approaches. Oxford: Pergamon, 1981.

Recebido em 28.8.1995.

Appendix
Analytical expression for the posterior marginal of π

To obtain the marginal distribution of π we first obtain the joint posterior distribution of (π, p). We note first that the Jacobian for the transformation of (p, q) into (π, p) is

$$\frac{1-p}{(1-\pi)^2}.$$

Using now function (4.1) we obtain the posterior density of (π, p) which is proportional to the following function whose domain is presented in Figure 1:

$$g(\pi, p) = (1-\pi)^{-18}\pi^{30}p^{68}(1-p)^{16}[(1-2\pi) + p\pi]^{16}. \qquad (A.1)$$

Using properties of the Beta function and expression (2.3.3) of Zacks (1981), we obtain the desired function, $f(\pi \mid data)$. This posterior marginal density of π is proportional to the following function:

$$f(\pi) = \begin{cases} f(\pi) = \dfrac{\pi^{30}(1-2\pi)^{16}}{(1-\pi)^{18}} \displaystyle\sum_{k=0}^{16} \frac{B(69+k;\,1)}{k\,B(k;\,17)} (\pi^{-1}-1)^{-k}, \quad 0 < \pi < 0.5, \\ f(\pi) = \dfrac{\pi^{30}(1-2\pi)^{16}}{1-\pi)^{18}} \displaystyle\sum_{k=0}^{16} \frac{B(69+k;\,1)}{k\,B(k;\,17)} \sum_{i=17}^{85-k} \frac{(\pi^{-1}-1)\,(2-\pi^{-1})^{85+k-i}}{i\,B(86+k-i;\,i)}, \quad 0.5 < \pi < 1. \end{cases}$$

where $B(a; b)$ is the Beta function evaluated at point (a, b).