OPTIMAL EXPERIMENTAL DESIGN

The Construction of locally D-optimal designs by canonical forms to an extension for the logistic model.

Dr. Raúl Martín Martín
Irene García Camacha Gutiérrez

OUTLINE

- Introduction to OED
- Canonical Forms
- Real Case:
    1. Model.
    2. Design.
    3. Information matrix.
    4. Optimality Criterion.
    5. Canonical Forms.
- Results.
- References.

INTRODUCCTION TO OED

"Data analysis will be informative only if the data are themselves"
(Rodríguez Torreblanca C. et al. 1999)

- What does the OED consist of ?

To select **where** and **how many** observations we must collect in order to validate the study
(in sense of obtaining the results which we are pursuing)

- Model: $y(x) = \eta(x, \theta) + \varepsilon(x) \equiv \theta^t f(x) + \varepsilon(x)\,, \qquad x \in \chi$

$$E[\varepsilon(x)] = 0, \qquad Var[\varepsilon(x)] = \sigma^2(x) = \frac{\sigma^2}{\lambda(x)}$$

- Design space (closed and compact): χ.
- Parameters vector: $\theta \in \mathbb{R}^k$.
- What values of x ∈ χ should it be taken the observations y in? How many are necessaries?

- <u>Exact Designs:</u>

  $N \in \mathbb{N} =$ number of experiments (or observations)

$$\xi = \begin{Bmatrix} x_1 & x_2 & \ldots & x_m \\ n_1 & n_2 & \ldots & n_m \end{Bmatrix}, \quad n_1 + n_2 + \ldots + n_m = N$$

- <u>Approximate Designs:</u>

  Defining a probability measure on χ ,

$$\xi = \begin{Bmatrix} x_1 & x_2 & \ldots & x_n \\ p_1 & p_2 & \ldots & p_n \end{Bmatrix}, \quad \sum_{i=1}^{n} p_i = 1 \quad \left(defining\ p_i = \frac{n_i}{N} \right)$$

$$\Xi = \{design\ measures\}, \quad Support = S_\xi = \{x \epsilon \chi,\ \xi(x) > 0\}$$

- <u>Continuous Designs:</u>

  Considering whatever probability measure (Atkinson, 1992)

- Variances-Covariances Matrix: $cov(\hat{\theta}) = \sigma^2 N^{-1} M^{-1}(\xi)$
$M(\xi) = \sum_{x \in \chi} f(x) f^t(x) \xi(x)$ (Information Matrix).

- Assuming $\sigma^2 = 1$, the information matrix to a design:

$$M(\xi) = \sum_{i=1}^{n} p_i \frac{\partial \eta(x_i, \theta)}{\partial \theta} \frac{\partial \eta(x_i, \theta)}{\partial \theta^t} = X^t \Omega X$$

$$\Omega = diag(p_1, \dots, p_n)$$

- The information matrix set is compact and convex: $\mathcal{M} = \{M(\xi) : \xi \in \Xi\}$

- *Caratheodory's theorem*: Each element $\mathcal{M}$ can be expressed as a convex linear combination.

$$\sum_{i=1}^{m} \lambda_i f(x_i) f^t(x_i)$$

where $m \leq \frac{k(k+1)}{2} + 1$.

- Criterion function: $\Phi: \mathcal{M} \rightarrow \mathbb{R} \cup \{\infty\}$

-Convex: $\Phi[\gamma M_1 + (1-\gamma) M_2] \leq \gamma \Phi(M_1) + (1-\gamma) \Phi(M_2)$

-Decreasing: $M(\xi) \geq M(\eta) \Rightarrow \Phi[M(\xi)] \geq \Phi[M(\eta)]$

-Homogeneous: $\Phi[\delta M] = \frac{1}{\delta} \Phi[M], \qquad \delta > 0$

- The ϕ-optimal design ξ* is which minimizes ϕ.

- If ϕ is strictly convex (in non-singular matrixes), there is a only minimum M(ξ*).

CRITERIA:

- D-optimization:

$$\phi_D[M(\xi)] = \log[\det M(\xi)^{-1/k}]$$

(minimizes the volume of the confidence ellipsoid of the parameters)

- G–optimization:

$$\Phi_G[M(\xi)] = \max_x d(x,\xi) = \max_x f(x)M^{-1}(\xi)f^t(x)$$

(minimizes the maximum value of the variance)

- A-optimization:

$$\Phi_A[M(\xi)] = TrM^{-1}(\xi)$$

(minimizes the variance promethium of the parameter estimators)

- E-optimization:

$$\Phi_E[M(\xi)] = \frac{1}{\lambda_\xi} \; being \; \lambda_\xi \; the \; minimun \; eigenvalue \; of \; M(\xi)$$

(minimizes the axis of the confidence ellipsoid)

<u>Note:</u> There are other criteria in which the interest is to know some or a linear combinations of the parameters

How to know if a design is optimal ?

- *General Equivalence theorem (GET):*

The following conditions are equivalents:

(1) $$\max_{\xi \in \Xi} |M(\xi)| = |M(\xi^*)|$$

(2) $$\min_{\xi \in \Xi} \sup_{x \epsilon \chi} d(x, \xi) = \sup_{x \epsilon \chi} d(x, \xi^*)$$

Moreover, $$\sup_{x \epsilon \chi} d(x, \xi^*) = \begin{Bmatrix} = k \; if \; x \in Supp(\xi^*) \\ < k \; if \; x \notin Supp(\xi^*) \end{Bmatrix}$$

$$\equiv \sup_{x \epsilon \chi} d(x, \xi^*) - k = \begin{Bmatrix} 0 \; if \; x \in Supp(\xi^*) \\ < 0 \; if \; x \notin Supp(\xi^*) \end{Bmatrix}$$

$\equiv$ directional derivative of ϕ_D

CANONICAL FORMS

- Optimal experimental design for non-linear problems depend on the values of the unknown parameters in the model.
- For various reasons, there is interest in providing explicit formulae for the optimal designs as a function of the unknown parameters.
- Certain class of generalized lineal models can be reduced to a canonical form to simplify the problem.
- The designs are constructed with a single variable using geometric and other arguments.

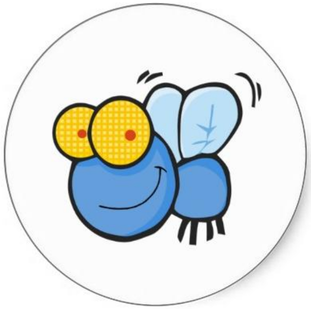
- It is necessary to have a design criterion invariant under transformation of the form: **x → z=Bx**, where B is a non-singular *kxk* matrix and x is mapped to z resulting in an introduced design space *Z*.

- The dependence of optimal design on the true value of Θ for given space χ is replaced by a design space which varies with Θ.

- In the literature it exists several studies with geometrical rules to construct the optimal design to the most important criteria and models.

Motivations

- There are many natural phenomena or external factors to which males and females respond differently.
- Atkinson et al. (1995) consider an experiment based on the dose-response to a fly insecticide males and females respond in a different way.



- The experiment consists of supplying a dose of insecticide on and analyzing its effectiveness.
- The characterization of this process is the impossibility to sex flies before and during the treatment application.
- In this work, it is proposed the use of canonical forms, Ford et al. (1992), in order to compute D-optimal designs.

1. MODEL

- Logistic Model for Binary Data:

$$y|x \sim Bi(1,\eta) \ where \ \eta(\theta,x) = \frac{e^{\alpha+\beta x+q\gamma}}{1+e^{\alpha+\beta x+q\gamma}} = F(\alpha+\beta x+\mathrm{q}\gamma) \ \ , a^* \leq x \leq b^*$$

being y: number of deaths

x: dose level

η: probability of death

q: factor (scores 0 males / 1 females)

$\vartheta=(\alpha,\beta,\gamma)$ parameter's vector

doing the logarithm of the probability ratio, it can be linearized :

$$log\left(\frac{\eta}{1-\eta}\right) = \alpha+\beta x+q\gamma = \mu$$

2. DESIGN

- Approximate design:

(Caratheodory's theorem sure us that the number of support points is finite and bounded)

$$\xi = \begin{Bmatrix} x_1 & x_2 & \dots & x_n \\ p_1 & p_2 & \dots & p_n \end{Bmatrix}$$

defined on the spatial region $\chi = [a^*, b^*]$ and

being the p_i the weight on the support points and verifying:

$$\sum_{i=1}^{n} p_i = 1$$

3. INFORMATION MATRIX

According to the previous design, the information matrix to a single observation x on an insect of known sex is :

$$M_{i,j}(x,\theta) = \omega(x,\theta)\frac{\partial\mu}{\partial\theta_i}\frac{\partial\mu}{\partial\theta_j} \quad \text{being} \quad \omega(x,\theta) = \lambda(x,\theta) = \eta(1-\eta)$$

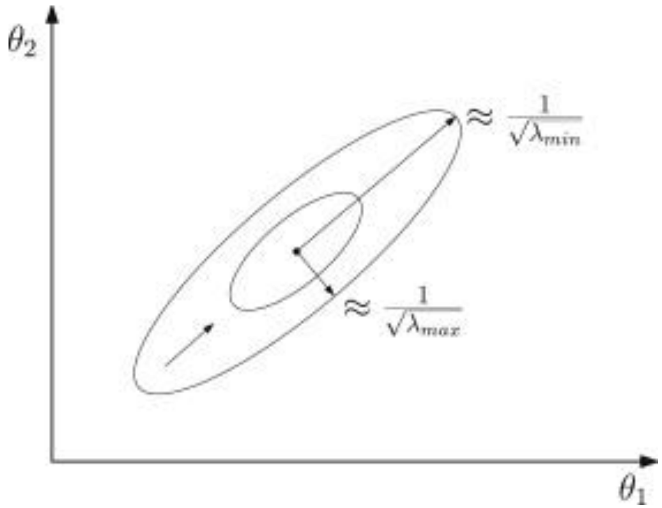
In spite of the experimental limitations about the lack of sex knowledge, it necessary to modify the above matrix to take into account this uncertainly. The information matrix for a fly whose sex is unknown is:

$$M(x,\theta) = 0.5\ \omega_M \begin{pmatrix} 1 & x & 1 \\ x & x^2 & x \\ 1 & x & 1 \end{pmatrix} + 0.5\ \omega_F \begin{pmatrix} 1 & x & 0 \\ x & x^2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

4. OPTIMALITY CRITERION

- D - optimality:

  D-optimal criterion minimizes the volume of the confidence ellipsoid of the parameters.



$$\phi_D[M(\xi,\theta)] = \log[\det M(\xi,\theta)^{-1/k}]$$

where *k* is the number of parameters in the model.

5. CANONICAL FORMS

- Reformulating the problem :

$$\eta(\theta, x) = \frac{e^{\alpha+\beta x+\mathrm{q}\delta}}{1+e^{\alpha+\beta x+\mathrm{q}\delta}} = F(\alpha+\beta x+\mathrm{q}\delta) \text{ con } a^* \leq x \leq b^*$$

<u>Note</u>: Knowing the insect' sex, next step is valid for both males and females.

The change of variable $z=\alpha+\beta x$ reduces the model:

$$\eta = F(z+\mathrm{q}\delta) = \frac{e^{z+\mathrm{q}\delta}}{1+e^{z+\mathrm{q}\delta}},$$

$$\alpha+\beta a^* \leq z \leq \alpha+\beta b^* \text{ or } \alpha+\beta a^* \geq z \geq \alpha+\beta b^* \; if \; \beta < 0$$

- It is necessary to apply this type of changes to write the information matrix: $M_x(\xi) = \sum_{i=1}^{n} p_i \, v(x_i) v(x_i)^t$

Thus,

$$v(x) = \frac{1}{\sqrt{\lambda(x,\theta)}} \underbrace{\left\{ \frac{\partial F(z+\mathrm{q}\delta)}{\partial(z+\mathrm{q}\delta)} \frac{\partial(z+\mathrm{q}\delta)}{\partial\alpha}, \frac{\partial F(z+\mathrm{q}\delta)}{\partial(z+\mathrm{q}\delta)} \frac{\partial(z+\mathrm{q}\delta)}{\partial\beta}, \frac{\partial F(z+\mathrm{q}\delta)}{\partial(z+\mathrm{q}\delta)} \frac{\partial(z+\mathrm{q}\delta)}{\partial\delta} \right\}^t}_{H(x_i)} = \frac{f(z+\mathrm{q}\delta)}{\sqrt{F(z+\mathrm{q}\delta)(1-F(z+\mathrm{q}\delta))}} \begin{pmatrix} 1 \\ x \\ \mathrm{q} \end{pmatrix}$$

the information matrix for n dose levels is:

$$M_x(\xi) = \sum_{i=1}^{n} \frac{p_i}{F(z_i+\mathrm{q}\delta)\big(1-F(z_i+\mathrm{q}\delta)\big)} H(x_i)H(x_i)^t = \sum_{i=1}^{n} p_i \frac{f(z_i+\mathrm{q}\delta)^2}{F(z_i+\mathrm{q}\delta)\big(1-F(z_i+\mathrm{q}\delta)\big)} \begin{pmatrix} 1 \\ x_i \\ \mathrm{q} \end{pmatrix} (1 \quad x_i \quad \mathrm{q}) =$$

$$= \sum_{i=1}^{n} p_i \, \omega(z_i+\mathrm{q}\delta) \begin{pmatrix} 1 \\ x_i \\ \mathrm{q} \end{pmatrix} (1 \quad x_i \quad \mathrm{q}) = \sum_{i=1}^{n} p_i \, v(x_i) v(x_i)^t$$

Note: $\omega(z+\mathrm{q}\delta) = \dfrac{f(z+\mathrm{q}\delta)^2}{F(z+\mathrm{q}\delta)\big(1-F(z+\mathrm{q}\delta)\big)}$

Considering now the linear transformation in the components:

$$\begin{pmatrix}1\\ z\\ \mathrm{q}\end{pmatrix} = \begin{pmatrix}1 & 0 & 0\\ \alpha & \beta & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}1\\ x\\ \mathrm{q}\end{pmatrix} = B\begin{pmatrix}1\\ x\\ \mathrm{q}\end{pmatrix}$$ and denoting

$$g(z) = Bv(x) = \frac{f(z+\mathrm{q}\delta)}{\sqrt{F(z+\mathrm{q}\delta)\big(1-F(z+\mathrm{q}\delta)\big)}}\begin{pmatrix}1 & 0 & 0\\ \alpha & \beta & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}1\\ x\\ \mathrm{q}\end{pmatrix} = \frac{f(z+\mathrm{q}\delta)}{\sqrt{F(z+\mathrm{q}\delta)\big(1-F(z+\mathrm{q}\delta)\big)}}\begin{pmatrix}1\\ z\\ \mathrm{q}\end{pmatrix}$$

So that we can write: $v(x) = B^{-1}g(z)$ (B is not singular; β≠0)

Then, the information matrix can be written equivalently as

$$M_x(\xi)=B^{-1} M_z(\xi)\, B.$$

Due to the D-optimal criterion does not vary by non-singular linear transformations of the design space, the maximization problem of $M_x(\xi)$ determinant reduces to maximize $M_z(\xi)$.

- The Information matrix for *n* dose levels and unknown sex will be:

$$M_z(\xi) = \sum_{i=1}^{n} p_i\, \omega(z_i + q\gamma) \begin{pmatrix} 1 \\ z_i \\ \mathrm{q} \end{pmatrix} (1 \quad z_i \quad \mathrm{q}) = \sum_{i=1}^{n} p_i\, g(z_i) g(z_i)^t, \quad \omega(z_i + \mathrm{q}\gamma) = \frac{f(z_i + \mathrm{q}\gamma)^2}{F(z_i + \mathrm{q}\gamma)\big(1 - F(z_i + \mathrm{q}\gamma)\big)}$$

- Adding it the uncertainty about sex:

$$M_z(\xi) = 0.5 \sum_{i=1}^{n} p_i [\, g(\, z_i) g(\, z_i)^t + \, h(\, z_i) h(\, z_i)^t]\,, \; g(z) = \begin{pmatrix} \sqrt{\omega(z+\delta)} \\ z\sqrt{\omega(z+\delta)} \\ \sqrt{\omega(z+\delta)} \end{pmatrix} \text{and } h(z) = \begin{pmatrix} \sqrt{\omega(z)} \\ z\sqrt{\omega(z)} \\ 0 \end{pmatrix}$$

- Expressed in a simplified way:

$$M_z(\xi) = 0.5 \sum_{k=1}^{2n} \xi(z_k)\, f_i(z_k) f_j(z_k)$$

6. RESULTS

- Formula for computing the determinant:

Considering $\chi = \{x_1, \ldots, x_n\}$ with $n \geq k$:

$$\det M(\xi) = \sum_{k_1 < \cdots < k_k} \xi(x_{k_1}) \ldots \xi(x_{k_k}) \det\left[\left\{f_i(x_{k_j})\right\}\right]^2$$

being k_i the S_n*'s* elements, which is the symmetric group of the four-order permutations.

6.1 TWO POINTS DESIGN

Basing on that, it results:

$$\left|M_{z_1,\,z_2}(\xi)\right| = 0.5[\,p^2(1-p)\,E + p(1-p)^2\,F]$$

where E and F are the squares of the determinants which result of combining the column matrixes $g(z_i)$, $h(z_i)$ with $i=1,2$ and operating them conveniently.

- Analytical Expression to the Optimal Weighs:

$$\frac{\partial\left|M_{z_1,\,z_2}(\xi)\right|}{\partial p} = 0.5\,[\,3(F-E)p^2 + 2(E-2F)p + F\,] = 0\,,$$

$$p^* = \frac{2F - E \pm \sqrt{F^2 - EF + E^2}}{3(F-e)}$$

Maximizing the determinant expression, it results:

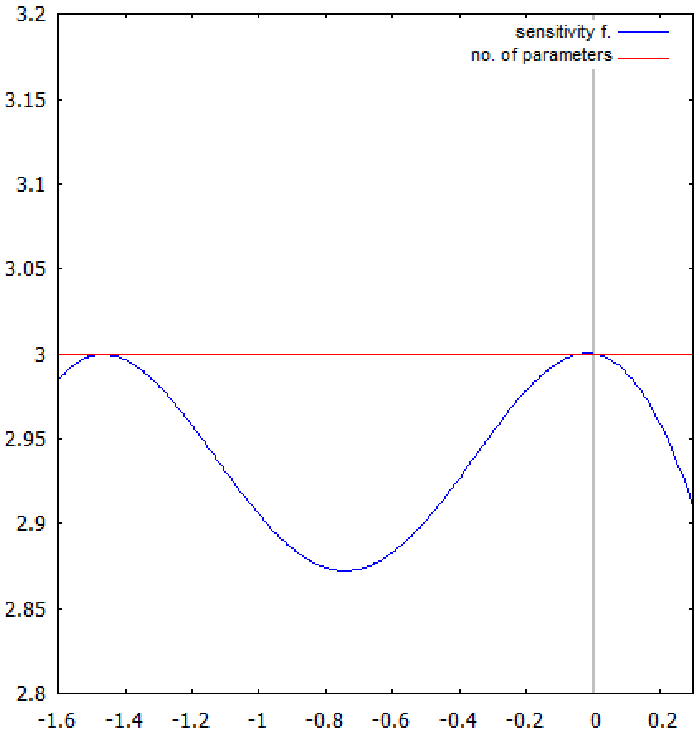
α=1.804, β=1.757, γ=-1		
x_i*	-1.467	-0.018
p_i *	0.5	0.5

- Testing the results:

Using the extension of equivalence theorem:

$$d(x,\xi^*,\theta) = 0.5\, w_M(x,\theta)\, f_M^t(x) M^{-1}(\xi^*,\theta) f_M(x) + 0.5\, w_F(x,\theta)\, f_F^t(x) M^{-1}(\xi^*,\theta) f_F(x) \leq k$$

where $k=3$ is the number of parameters. The equality is produced in the points of the optimal design. Plotting this function to the obtained results, it is checked that a D-optimal design has been obtained.



6.2 THREE POINTS DESIGN

From previous studies (*Atkinson et al. 1995*) it is checked that the optimal weights are symmetrical to the three-point case:

$$\xi = \begin{Bmatrix} z_1 & z_2 & z_3 \\ p & q = 1 - 2p & p \end{Bmatrix}$$

In this case, to calculate the determinant:

$$\left| M_{z_1, z_2, z_3}(\xi) \right| = 0.5 \, \{A \, p^3 + B \, p^2 + Cp\}$$

where A, B and C are the squares of the determinants which result from applying the formula to its fast calculation gathering them conveniently.

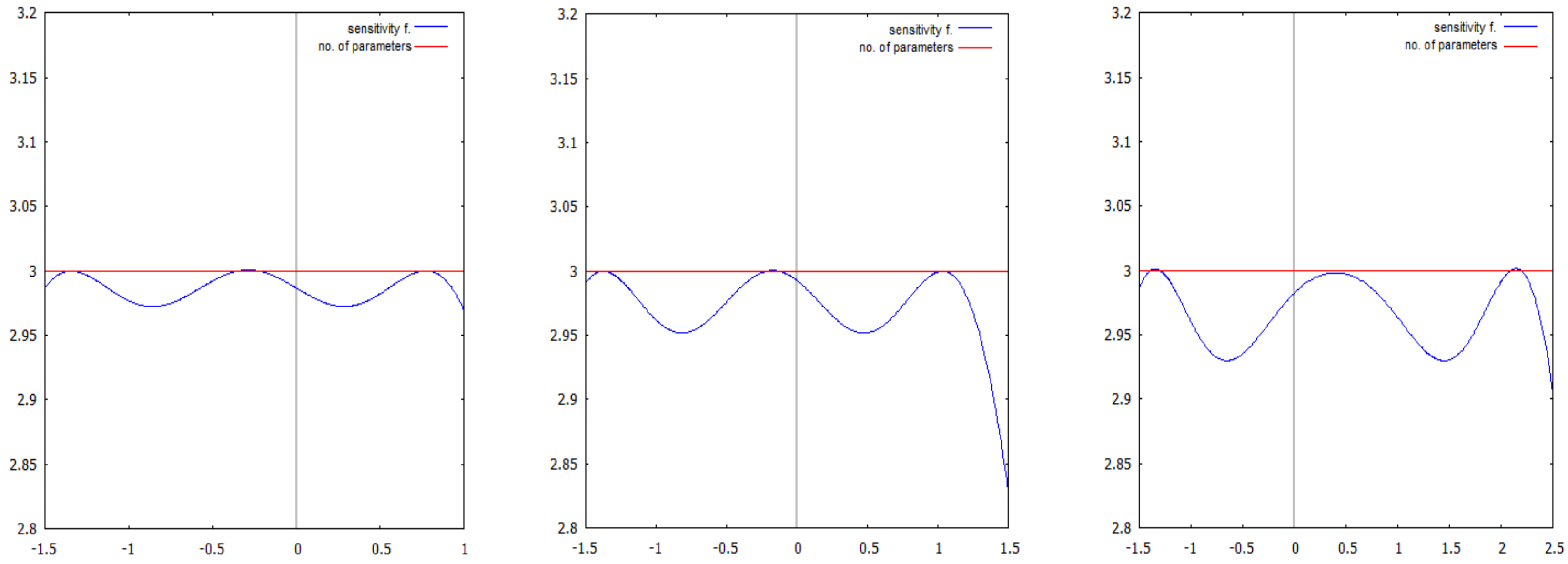
- Analytical Expression to the Optimal Weighs: $p^* = \frac{- B \pm \sqrt{B^2 - 3AC}}{3C}$

Maximizing the determinant expression, it results:

$\alpha=1.804$ $\beta=1.757$	$\gamma=-2.599$			$\gamma=-3$			$\gamma=-5$		
x_i*	-1.347	-0.287	0.772	-1.378	-0.173	1.032	-1.349	0.396	2.141
p_i *	0.375	0.250	0.375	0.339	0.322	0.339	0.316	0.368	0.316

Testing the results:

Through the equivalence theorem, it is checked that the obtained design are D-optimal.



REFERENCES

[1] Atkinson, A.C., Demetrio, C.G.B. and Zochhi, S.S. (1995). *Optimun Dose Levels when Males and Females Differ en Response*. Appl. Statist. 44 (2), 213-226.

[2] Ford, I., Tosney, B. and Wu, C.F.J. (1992). *The use of a canonical form in the construction of locally optimal designs for non-linear problems*. J. R. Statist. Soc. B. 54 (2), 569-583.

Thank you for your attention.