Meta-Analytical Issues in Linear Models

Filiation: Pestana, D.D., CEAUL e U.Lisboa and Martins, J.P., CEAUL and IPLeiria.

Keywords: Prospective cumulative meta-analysis, optimal designs, discriminant designs, mixed designs, quasi-optimal designs.

Abstract

Regression uses one or more covariates to assess the relationship between those covariates and a dependent variable. In meta-regression the same approach is valid with only one difference: the covariates are at the level of the study.

Optimal design theory deals with the choice of the allocation of the observations to accomplish the estimation of the coefficients in a regression model in an optimal way. So far there isn't much literature related with design theory applied to meta-regression. However, when the studies available don't provide enough "statistical evidence" the researcher may conduct a new study to add to its meta-analysis. In this context, it is of great importance no to choose the covariates levels of this new study at random.

The purpose of this work is to provide a framework of this problem. Optimal discriminant, optimal robust and mixed designs are used to provide a solution to this problem. Some examples are given that emphasize the importance of the choice of the design.

References

- Martins, J.P., Mendonça, S. e Pestana, D.D. (2008). Optimal and quasi optimal designs. *RevStat – Statistical Journal*, Vol. 6, 279-307

1 Optimal designs

A linear model is defined as

$$\boldsymbol{Y} = X^T\boldsymbol{\theta} + \varepsilon \tag{1}$$

where $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_k)^T$ is a vector with k unknown parameters and $X = \boldsymbol{f}(x) = (f_1(x), \cdots, f_k(x))^T$ is a $n \times k$ matrix called **design matrix** that depends on the observation points $x_1, \cdots, x_n$. The columns of X are assumed continuous and independent. The vector $\boldsymbol{\varepsilon} = (\varepsilon_1, \cdots, \varepsilon_n)$ corresponds to the random error vector with mean zero and variance σ^2 (homocedasticity). It is usually assumed that $\boldsymbol{Y}$ follows a multivariate normal distribution, $\boldsymbol{Y} \sim N\left(X^T\boldsymbol{\theta}, \sigma^2 I_n\right)$.

We will address the problem of obtaining linear unbiased estimators for $\boldsymbol{\theta}$, that is, an estimator $\widehat{\boldsymbol{\theta}}_L$ that verifies

$$\widehat{\boldsymbol{\theta}}_L = L\boldsymbol{Y} \tag{2}$$

where L is a $k \times n$ matrix.

It is necessary to establish an order relation between two linear unbiased estimators $\widehat{\boldsymbol{\theta}}_L^1$ and $\widehat{\boldsymbol{\theta}}_L^2$ to compare them. It is often used the Loewner order whose definition is given below.

Definition 1.1 *Let $A, B \in \mathrm{Sym}(k)$ be two symmetric matrices with order k:*

1. $A \geq B$, with respect to Loewner order if, and only of, $A - B \in \mathrm{NND}(k)$;

2. $A > B$, with respect to Loewner order if, and only of, $A - B \in \mathrm{PD}(k)$.

It is well known in design theory that the best linear unbiased estimator (BLUE) of $\boldsymbol{\theta}$ is the Gauss-Markov estimator

$$\widehat{\boldsymbol{\theta}}^{GM} = \left(X^TX\right)^- X^TY$$

with respect to the Loewner order, that is,

$$\sigma^2\left(X^TX\right)^{-1} = D\left(\widehat{\boldsymbol{\theta}}^{GM}\right) \leq D\left(\widehat{\boldsymbol{\theta}}_L\right)$$

for all linear unbiased estimators $\widehat{\boldsymbol{\theta}}_L$ of $\boldsymbol{\theta}$ (Dette e Studden, 1997, p. 131) and where $D(A)$ is the dispersion matrix of A.

Sometimes, it is of interest to estimate some linear combinations of the parameters $\theta_1, \cdots, \theta_k$, say $K^T\boldsymbol{\theta}$ with order $k \times s$ $(s \leq k)$. The BLUE for exists $K^T\boldsymbol{\theta}$ only if $\mathrm{Im}(K) \subseteq \mathrm{Im}\left(X^T\right)$ and is given by

$$\widehat{\boldsymbol{\theta}}^K = K^T\left(X^TX\right)^- X^T\boldsymbol{Y} \tag{3}$$

where $\left(X^T X\right)^-$ is the generalized inverse matrix of $X^T X$. The dispersion matrix of $\widehat{\boldsymbol{\theta}}^K$ is equal to

$$D\left(\widehat{\boldsymbol{\theta}}^K\right) = \sigma^2 K^T \left(X^T X\right)^- K. \tag{4}$$

Note that the dispersion matrix $D\left(\widehat{\boldsymbol{\theta}}^K\right)$ doesn't depend on the choice of $\left(X^T X\right)^-$.

Example 1.1 *Consider the linear model (1) and take $\boldsymbol{f}(x) = (1, x, \cdots, x^m)^T$. Then the model (1) may be written as*

$$Y = \sum_{j=0}^{m} \theta_{j+1} x^j + \varepsilon. \tag{5}$$

The matrix X is invertible if there is at least $m+1$ different points of observation. The estimator BLUE of $K^T\boldsymbol{\theta}$ is given by (3) where

$$\frac{1}{n} X^T X = \begin{bmatrix} 1 & c_1 & c_2 & \cdots & c_m \\ c_1 & c_2 & c_2 & \cdots & c_{m+1} \\ \vdots & \vdots & \vdots & & \vdots \\ c_m & c_{m+1} & c_{m+2} & \cdots & c_{2m} \end{bmatrix},$$

$c_j = \frac{1}{n} \sum\limits_{i=1}^{n} x_i^j$, $j = 0, \cdots, 2m$ *and n is the sample size.*

In the context of the linear model (1) it is important to test

$$H_0 : K^T \boldsymbol{\theta} = 0. \tag{6}$$

If the sample size n is higher than the matrix X rank $\mathrm{r}(X)$, $n > \mathrm{r}(X)$, the F-test for testing (6) rejects H_0 when the value of the statistic

$$F = \frac{n - \mathrm{car}(X)}{\mathrm{car}(K)} . \frac{\left(\widehat{\boldsymbol{\theta}}^K\right)^T \left(K^T \left(X^T X\right)^- K\right)^- \widehat{\boldsymbol{\theta}}^K}{Y^T \left(I_n - X\left(X^T X\right)^- X^T\right) Y}$$

is high. The statistic F has a non-central Fisher-Snedecor F distribution with $\mathrm{r}(K)$ and $n - \mathrm{r}(X)$ degrees of freedom. The non-centrality parameter is

$$\frac{1}{\sigma^2} \left(K^T \boldsymbol{\theta}\right)^T \left(K^T \left(X^T X\right)^- K\right)^- \left(K^T \boldsymbol{\theta}\right) \tag{7}$$

that depends on the matrix design. If the parameter (7) increases then the power test also increases. So, the choice of $X^T X$ has a great importance.

Optimal design theory deals with the choice of the allocation of the observations to accomplish the estimation of the coefficients in a regression model in an optimal way.

We will represent the points of observation $x_1, \cdots, x_l$ observed $n_1, \cdots, n_l$ times using the **exact design matrix**

$$\xi_{(n)} = \begin{pmatrix} x_1 & \cdots & x_l \\ \dfrac{n_1}{n} & \cdots & \dfrac{n_l}{n} \end{pmatrix} \tag{8}$$

where $n = n_1 + \ldots + n_l$.

Hence,

$$\begin{aligned} X^T X &= \sum_{j=1}^{n} f(x_j) f^T(x_j) = n \sum_{j=1}^{l} \frac{n_j}{n} f(x_j) f^T(x_j) \\ &= n \int_{\mathcal{X}} f(x) f^T(x) \, d\xi_{(n)}(x) = nM\left(\xi_{(n)}\right), \end{aligned}$$

where $M\left(\xi_{(n)}\right)$ is the **information design matrix**.

Using this notation the BLUE for $K^T\boldsymbol{\theta}$ (3) and its dispersion matrix (**??**) may be re-written as

$$\widehat{\boldsymbol{\theta}}^K = K^T \left(X^T X\right)^{-} X^T \ \mathbf{Y} = \frac{1}{n} K^T M^{-}\left(\xi_{(n)}\right) X^T \mathbf{Y}. \tag{9}$$

and

$$D\left(\widehat{\boldsymbol{\theta}}_K\right) = \frac{\sigma^2}{n} K^T M^{-}\left(\xi_{(n)}\right) K \tag{10}$$

which depends on the choice of $\xi_{(n)}$.

However, it is very difficult to minimize $D\left(\widehat{\boldsymbol{\theta}}_K\right)$ with respect to the Loewner order as if the sample size n is fixed, the weights of the design matrix (8) must be multiple of $1/n$. One solution is to consider **approximate designs** which allow the weights to be non-multiple of $1/n$. Finding a more general solution is an easier task (Fedorov, 1972). In practice, after obtaining the optimal approximate design it is used an exact design with weights similar to the weights of the optimal approximate design.

As the minimization of $K^T M^{-}\left(\xi_{(n)}\right) K$ is very hard to do with respect to Loewner order it is preferable to maximize

$$C_K\left(M\left(\xi_{(n)}\right)\right) = \left(K^T M^{-}\left(\xi_{(n)}\right) K\right)^{-1}. \tag{11}$$

with respect to some functional of $C_K\left(M\left(\xi_{(n)}\right)\right)$.

One functional very popular is the family of functions Kiefer-ϕ_p (Kiefer e Studden, 1976).

Definition 1.2 *Let $p \in [-\infty, 1]$ and consider*

$\phi_p : \{C : C \in \mathrm{PD}(s)\} \longrightarrow \mathbb{R}$ *defined by*

$$\phi_p(C) = \begin{cases} \left(\frac{1}{s} trace\left(C^p\right)\right)^{1/p} & \text{if } p \neq -\infty, 0 \\ (\det C)^{1/s} & \text{if } p = 0 \\ \lambda_{\min}(C) & \text{if } p = -\infty \end{cases}$$

e $\phi_p : \{C : C \in \mathrm{NND}(s)\} \longrightarrow \mathbb{R}$ *definied by*

$$\phi_p(C) = \begin{cases} \left(\frac{1}{s} trace\left(C^p\right)\right)^{1/p} & \text{if } p \in (0, 1] \\ 0 & \text{if } p = [-\infty, 0] \end{cases} .$$

A ϕ_p-***optimal design*** *is a design that maximizes* $\phi_p\left(C_K M\left(\xi_{(n)}\right)\right)$.

In particular, when $p = 0$, a design $\xi_{(n)}$ is a **D-optimal design** when it maximizes

$$\phi_0\left(C_K\left(M\left(\xi_{(n)}\right)\right)\right) = \left(\det\left(K^T M^-\left(\xi_{(n)}\right) K\right)\right)^{-1/s}.$$

If $K = I_k$ then the D-optimal design maximizes $\det\left(M\left(\xi_{(n)}\right)\right)^{-1/s}$. If $K = (0, \cdots, 0, 1)^T$ then the optimal design is called $\mathbf{D_1}$**-optimal**.

To compare the efficiency of two designs or of single design in the estimation of $K^T\boldsymbol{\theta}$ the following two concepts are used.

Definition 1.3 *The* ϕ_p-***efficiency of a design*** ξ *for estimating* $K^T\boldsymbol{\theta}$ *is given by*

$$\mathrm{ef}_p(\xi) = \frac{\phi_p\left(C_K\left(M(\xi)\right)\right)}{\sup_\eta\left(\phi_p\left(C_K\left(M(\eta)\right)\right)\right)}.$$

The ϕ_p-***efficiency of a design*** ξ_1 ***compared to a design*** ξ_2 *is given by*

$$r_p(\xi_1, \xi_2) = \frac{\phi_p\left(C_K\left(M(\xi_1)\right)\right)}{\phi_p\left(C_K\left(M(\xi_2)\right)\right)}.$$

2 $\mathbf{D_1}$-optimal and D-optimal designs

We will now consider a particular model of (1) as we will assume $\boldsymbol{f}(x) = (1, x, \cdots, x^m)^T$. So,

$$Y = \sum_{j=0}^{m} \theta_j x^j + \varepsilon. \tag{12}$$

with $x \in [-1, 1]$.

The D_1-optimal design is the best design to use when our focus is in the estimation of the coefficient θ_m. In this case, $K = e_m = (0, \cdots, 0, 1)^T \in \mathbb{R}^{m+1}$ and $K^T \boldsymbol{\theta} = \theta_m$. Hence, (11) is equal to

$$C_{e_m}\left(M\left(\xi_{(n)}\right)\right) = \left(e_m^T M^{-1}\left(\xi_{(n)}\right) e_m\right)^{-1} = \frac{|M_m(\xi)|}{|M_{m-1}(\xi)|} \tag{13}$$

(Dette e Studden, 1997, p. 150) where

$$M_m(\xi) = \int_{-1}^{1} \boldsymbol{f}_m(x) \boldsymbol{f}_m^T(x) \, d\xi_{(n)}(x) = (c_{i+j})_{i,j=0}^{m}$$

Dette e Studden (1997) using canonical moments demonstrate the next theorem.

Theorem 2.1 *The D_1-optimal design ξ^{D_1} with respect to model (12) has equal weights $1/m$ at the zeros of the Chebyshev polynomial of second kind $U_{m-1}(x)$ and weight $1/2m$ at the points -1 and 1. If instead of $x \in [-1, 1]$ we have $x \in [a, b]$ the D_1-optimal design is given by*

$$\xi_{[a,b]}^{D_1}(\{x\}) = \xi^{D_1}\left(\left\{\frac{2x - b - a}{b - a}\right\}\right).$$

For $m = 1, \cdots, 4$ the support points of the design ξ^{D_1} are given in the next table.

m	Points
2	$-1, 0, 1$
3	$-1, -0.5, 0.5, 1$
4	$-1, -0.70711, 0, 0.70711, 1$
5	$-1, -0.80902, -0.30901, 0.30901, 0.80902, 1$

Table 1 – Observation points in a D_1-optimal design

Since $C_{e_m}\left(M\left(\xi^{D_1}\right)\right) = (1/2)^{2m-2}$, the **$D_1$-efficiency of a design ξ in the estimation of** θ_m is equal to

$$\text{eff}_m^{D_1}(\xi) = \frac{C_{e_m}(M(\xi))}{C_{e_m}\left(M\left(\xi^{D_1}\right)\right)} = \frac{2^{2m-2}|M_m(\xi)|}{|M_{m-1}(\xi)|}. \tag{14}$$

A D-optimal design is concerned with the estimation of the full vector $\boldsymbol{\theta}$. Dette e Studden (1997, pp. 149 e 150) achieve a solution for the model (12) using canonical moments once again. The result is stated below.

Theorem 2.2 *The D-optimal design ξ^D for the estimation of the full vector $\boldsymbol{\theta}$ has equal weights $1/(m+1)$ at the $m+1$ polynomial $\left(x^2 - 1\right) L_m'(x)$ zeros where L_m' is the first derivate of the m-th Legendre polynomial.*

For $m = 1, \cdots, 4$ the support points of the design ξ^D are given in the next table.

m	Points
1	$-1, 1$
2	$-1, 0, 1$
3	$-1, -0.447214, 0.447214, 1$
4	$-1, -0.654654, 0.654654, 1$

Table 2 – Observation points in a D-optimal design

The **D-efficiency of a design ξ in the estimation of $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_m)$** is given by

$$\text{eff}_m^D(\xi) = \left(\frac{|M_m(\xi)|}{\left|M_m\left(\xi^D\right)\right|} \right)^{1/(m+1)} .$$

where (Dette e Studden, 1997, p. 150)

$$\det\left[M_m\left(\xi^D\right)\right]^{1/(m+1)} = \left(\frac{m}{2m-1}\right)^m \prod_{j=2}^{m} \left(\frac{(m-j+1)^2}{(2(m-j)+1)(2(m-j)+3)} \right)^{m+1-j} .$$

3 Robust model optimal and discriminant optimal designs

When using a D_1-optimal design it is important to determine a value for the parameter m of the model (12). An explanatory analysis may help but it is in general very difficult to assume a specific value for m.

The D_1-optimal design for $m = 4$ could shows a very poor performance if the true value for m is lower than four as the next table puts in evidence.

m	$\text{eff}_m\left(\xi_4^{D_1}\right)$
1	1
2	0.5
3	0.5
4	0.5

Table 3 – Efficiency of the D_1-optimal design for $m = 4$

Unfortunately, the D_1-optimal design for a value m is not the D_1-optimal design for a value $m-1$.

Anderson suggests a test for assessing the significance of the coefficient of the term with the highest degree starting with a reasonable high degree in a decreasing order. This strategy has the problem of estimating global level of significance so we suggest that a prior vector $\boldsymbol{\pi} = (\pi_1, \cdots, \pi_m)$ ($\pi_i > 0$ e $\sum_{i=1}^{m} = 1$) to be taken where the components of the vector π_i reflect our belief that the value of m is i. Hence, a design that maximizes a balanced geometric mean of the values $C_{e_m}\left(M\left(\xi_i^{D_1}\right)\right)$ with weights π_i is called a discriminant $\Psi_0^{\boldsymbol{\pi}}$-optimal design.

Definition 3.1 *The design $\xi_{0,\boldsymbol{\pi}}$ is called a* ***discriminant $\Psi_0^{\boldsymbol{\pi}}$-optimal design for the class $\mathcal{F}_m$*** *with respect to the prior $\boldsymbol{\pi}$ when it maximizes the balanced geometric mean*

$$\Psi_0^{\boldsymbol{\pi}}(\xi) = \prod_{l=1}^{m}\left(\mathrm{eff}_l^{D_1}(\xi)\right)^{\pi_l} = \prod_{l=1}^{m} 2^{2l-2}\left(\frac{|M_l(\xi)|}{|M_{l-1}(\xi)|}\right)^{\pi_l}$$

where

$$\mathcal{F}_m = \{h_l(x)\,|\,l = 1, \cdots, m\}$$

and

$$h_l(x) = \sum_{i=0}^{l}\theta_{li}x^i = \boldsymbol{\theta}_l^T \boldsymbol{f}_l(x), \quad l = 1, \cdots, m$$

The case $\boldsymbol{\pi} = (0, \cdots, 0, 1)$ corresponds to the D_1-optimal design.

Example 3.1 *The discriminant $\Psi_0^{\boldsymbol{\pi}}$-optimal designs for the classes $\mathcal{F}_2, \mathcal{F}_3$ e $\mathcal{F}_4$ with respect to a prior $\boldsymbol{\pi}$ with equal components may be computed using the theory of canonical moments. The results are given in the next table (all the points weights are equal to $1/m$).*

m	$\boldsymbol{\pi}$	*Points*	$\mathrm{eff}_1^{D_1}/\mathrm{eff}_2^{D_1}/\mathrm{eff}_3^{D_1}/\mathrm{eff}_4^{D_1}(\xi)$
2	$\left(\frac{1}{2}, \frac{1}{2}\right)$	$-1, 0, 1$	$0.817/1/-/-$
3	$\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$	$-1, -0.4472, 0, 4472, 1$	$0.600/0.640/0.853/-$
4	$\left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right)$	$-1, -0.6547, 0, 0.6547, 1$	$0.571/0.588/0.627/0.836$

Table 4 – discriminant $\Psi_0^{\boldsymbol{\pi}}$-optimal designs

It is interesting to verify that the discriminant $\Psi_0^{\boldsymbol{\pi}}$-optimal designs for the class $\mathcal{F}_m$ and the D-optimal designs for the model (12) are the same.

The discriminant $\Psi_0^{\boldsymbol{\pi}}$-optimal design for the class $\mathcal{F}_4$ is less efficient than the D_1-optimal design when $m = 4$. But when $m < 4$ discriminant $\Psi_0^{\boldsymbol{\pi}}$-optimal design shows a bigger efficiency.

In a similar way, it is possible to obtain a design that maximizes a balanced geometric mean efficiency for the estimation of the full parameter $\boldsymbol{\theta}$ for the class $\mathcal{F}_m$ given a prior $\boldsymbol{\pi}$. Such a design is called a robust model $\Xi_0^{\boldsymbol{\pi}}$-optimal design.

Definition 3.2 *The design $\xi_{0,\boldsymbol{\pi}}$ is called a **robust model $\Xi_0^{\boldsymbol{\pi}}$-optimal design for the class** $\mathcal{F}_m$ with respect to the prior $\boldsymbol{\pi}$ when it maximizes the balanced geometric mean*

$$\Xi_0^{\boldsymbol{\pi}}(\xi) = \prod_{l=1}^{m} \left(\mathrm{eff}_l^D(\xi)\right)^{\pi_l} = \prod_{l=1}^{m} \left(\frac{|M_l(\xi)|}{\left|M_l\left(\xi^D\right)\right|}\right)^{\pi_l/(l+1)}.$$

Example 3.2 *The robust model $\Xi_0^{\boldsymbol{\pi}}$-optimal designs for the classes $\mathcal{F}_2$, $\mathcal{F}_3$ e $\mathcal{F}_4$, using a prior $\boldsymbol{\pi}$ with equal components are given in the next table.*

m	Pontos	Pesos respectivos	$\mathrm{eff}_1^D/\mathrm{eff}_2^D/\mathrm{eff}_3^D/\mathrm{eff}_4^D(\xi)$
2	$-1, 0, 1$	$0.389, 0.222, 0, 389$	$0.881/0.968/-/-$
3	$-1, -0.401, 0.401, 1$	$0.319, 0.181, 0.181, 0.319$	$0.835/0.914/0.954/-$
4	$-1, -0.605, 0, 0.605, 1$	$0.271, 0.152, 0.153, 0.152, 0.271$	$0.809/0.883/0.927/0.949$

Quadro 1: *Table 5 – Robust model $\Xi_0^{\boldsymbol{\pi}}$-optimal designs*

Martins, Mendonça e Pestana (2008) suggest the use of designs that combine the D-efficiencies with the D_1-efficiencies. These designs maximize the geometric mean

$$\Theta_0^{\boldsymbol{\pi}} = \prod_{l=1}^{m} \left(\mathrm{eff}_l^{D_1}(\xi)\right)^{\pi_l} \prod_{j=1}^{m} \left(\mathrm{eff}_j^D(\xi)\right)^{\pi_j} \tag{15}$$

$$= \prod_{l=1}^{m} \left(2^{2l-2}\frac{|M_l(\xi)|}{|M_{l-1}(\xi)|}\right)^{\pi_l} \prod_{j=1}^{m} \left(\frac{M_j(\xi)}{M_j\left(\xi_j^D\right)}\right)^{\pi_j/(j+1)}$$

Definition 3.3 *The design $\xi_{0,\boldsymbol{\pi}}$ is called a mixed **$\Theta_0^{\boldsymbol{\pi}}$-optimal design for the class** F_m with respect to the prior $\boldsymbol{\pi}$ when it maximizes the geometric mean $\Theta_0^{\boldsymbol{\pi}}$ defined at (15).*

4 Prospective cumulative meta-analysis

A cumulative meta-analysis is a meta-analysis that is performed with a number of steps equal to the number n of available studies. In some sistematic way, a meta-analysis is done first with one study, then with two studies and so on. Presenting the results of the n meta-analysis could be a way of identifying patterns in the data associated to the sistematic choice of the studies. If time is chosen to order the studies a pattern may be found and the use of meta-regression may be a very useful tool for understanding the relationship between time and the variable response.

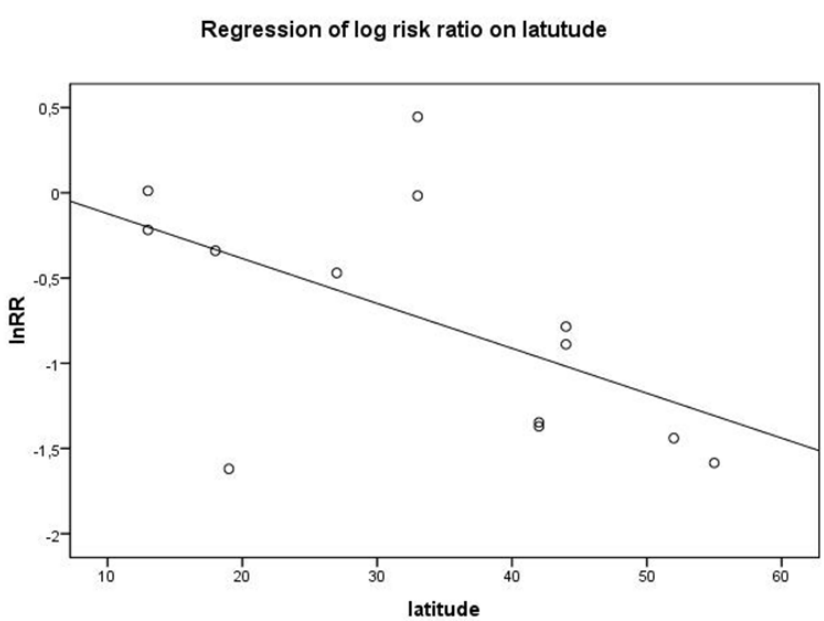
A prospective cumulative meta-analysis is similar to the cumulative meta-analysis at least inits first step. Studies are gathered and then k meta-analysis are produced. If the results obtained are not satisfatory the investigator could wait for the next study or do it by itsself. If the last situation happens it is important design a study that will have the "greatest chance" of improving the results.

If the results obtained using meta-regression need to be improved, a discriminant Ψ_0^π-optimal, a robust model Ξ_0^π-optimal or a mixed Θ_0^π-optimal design may be used to provide the best allocation for the new study to be done.

As an example consider the meta-analysis of 13 studies of the impact of a vaccine, BCG, to prevent the development of tuberculosis (TB). The risk ratio (RR) obtained in each study is presented in the next table.

	Vaccinated		Control			
Study	TB	Total	TB	Total	RR	Latitude
1	8	2545	10	629	0.198	19
2	6	306	29	303	0.205	55
3	62	13598	248	12867	0.237	52
4	17	1716	65	1665	0.254	42
5	3	231	11	220	0.260	42
6	4	123	11	139	0.411	44
7	180	1541	372	1451	0.456	44
8	29	7499	45	7277	0.625	27
9	186	50634	141	27338	0.712	18
10	33	5069	47	5808	0.804	13
11	27	16913	29	17854	0.983	33
12	505	88391	499	88391	1.012	13
13	5	2498	3	2341	1.562	33

Looking at the previous table it is clear that the studies don't report the same effects. Representing the logarithmic of the risk ratio *versus* latitude (L) a pattern is observed.



The adjusted line to the scatter is defined by

$$\ln(RR) = 0.2595 - 0.0292L.$$

The p-value for testing if the slope is null is about 0.001.

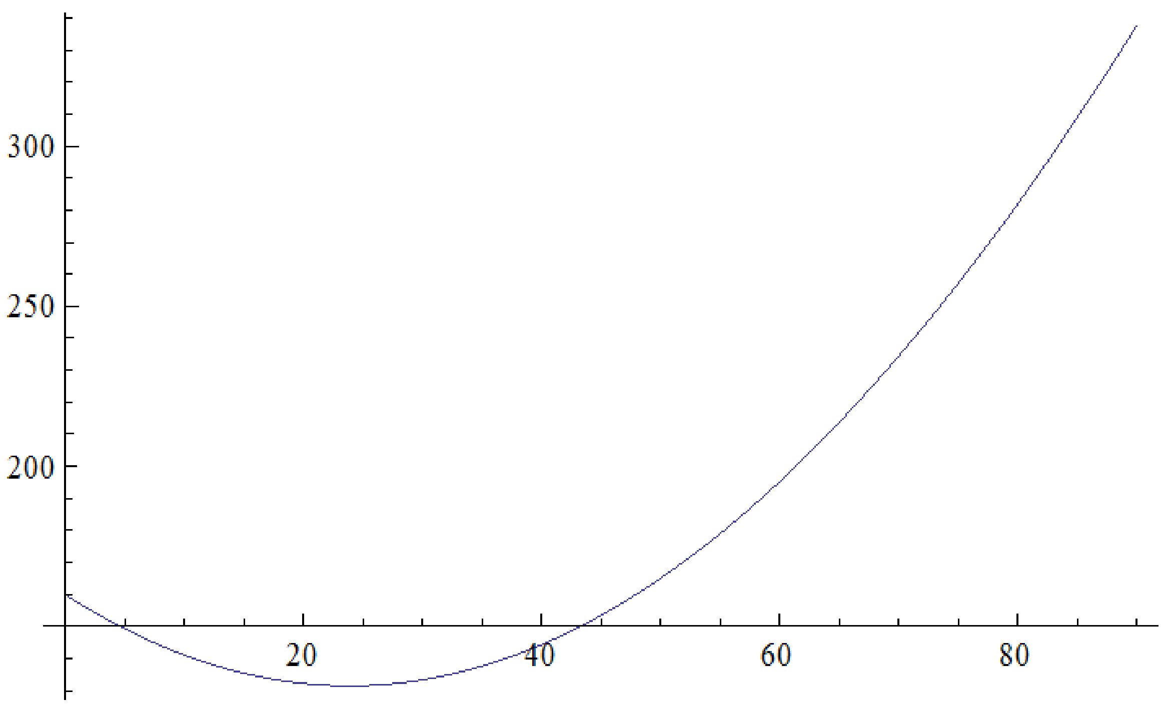
We could try to fit a model of the type

$$\ln(RR) = \theta_0 + \theta_1 L + \theta_2 L^2.$$

....

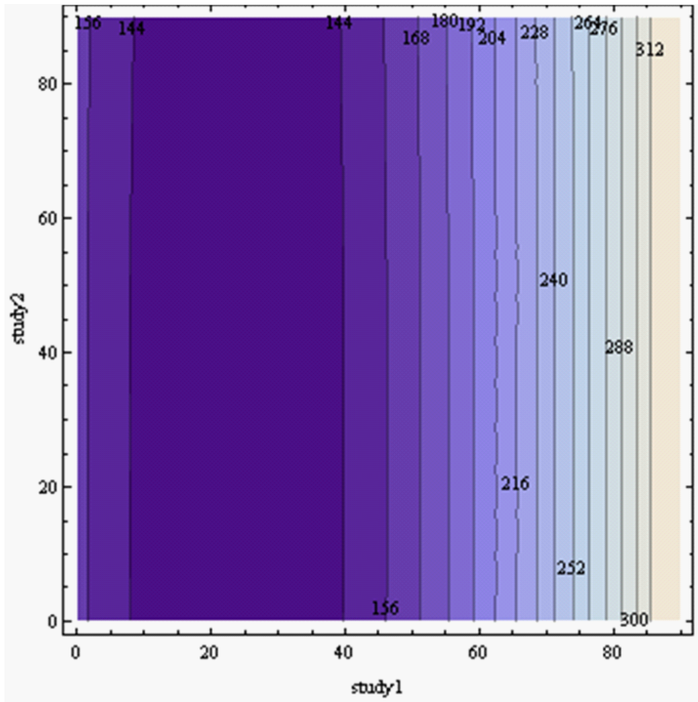
So we will just consider a polynomial regression till degree equal to two. Using a prior $(0.9, 0.1)$ we may determine the discriminant Ψ_0^{π}-optimal design that maximizes $C_{e_m}\left(M\left(\xi_{(n)}\right)\right)$ defined in (13). We will assume that the new study has the same sample size as study 8. The problem we are concerned is to estimate the best latitude to do this new study.

Using software Mathematica 6.0 we obtain a graph representing $C_{e_m}\left(M\left(\xi_{(n)}\right)\right)$ on the latitude of the new study.



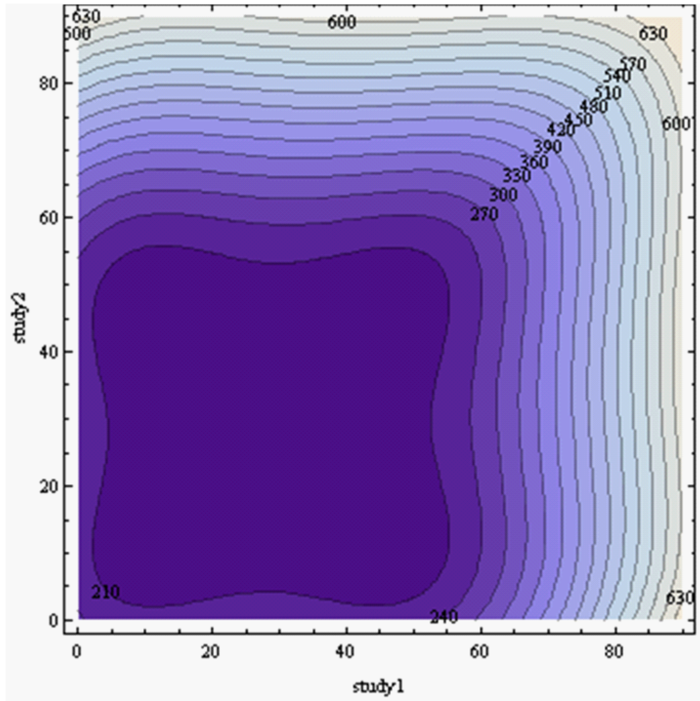
Clearly, the new study should be done approximately at latitude 24^o. The minimum is around 131.

If instead of one study we wish to add two new studies each one having a sample size equal to half of the sample size of study 8 one would get the following contour plot.



The minimum value of $C_{e_m}\left(M\left(\xi_{(n)}\right)\right)$ is close to 131 obtained using latitudes 23.87^o and 44.70^o.

The use of a different prior, for example $\pi = (0.5, 0.5)$ would yield a different contour plot.



As there is no significant differences between the values of $C_{e_m}\left(M\left(\xi_{(n)}\right)\right)$ when using one or two studies it's possible to choose the solution that helps more in avoiding problems with bias publication...

Bibliography

Anderson, T. (1962). The choice of the degree of a polynomial regression as a multiple decision problem. *Annals of Mathematical Statistics*, Vol. 33, 255-265.

Dette, H., and Studden, W. (1997). *The Theory of Canonical Moments with Applications in Statistics, Probability and Analysis.* Wiley-Interscience Publication, New York.

Fedorov, V. (1972). *Theory of Optimal Experiments*, Academic Press, New York.

Kiefer, J., and Studden, W. (1976). Optimal designs for large degree polynomial regression. *The Annals of Statistics*, Vol. 4, 1113-1123.

Martins, J.P., Mendonça, S. e Pestana, D.D. (2008). Optimal and quasi optimal designs. *RevStat – Statistical Journal*, Vol. 6, 279-307