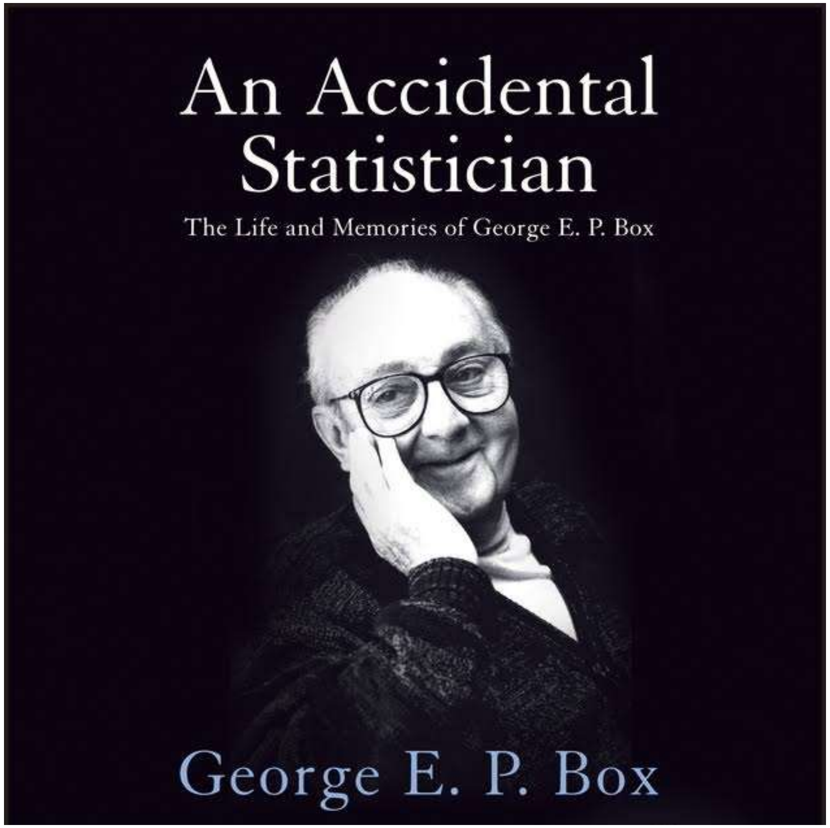
El experimento del helicóptero de papel de G. E. Box. La estadística antes de la experimentación

Victormanuel.casero@uclm.es

Box paper helicopter



CQPI Report No. 76, December 1991

Teaching Engineers Experimental Design With A Paper Helicopter

George Box

How a paper "helicopter" made in a minute or so from a 8 1/2" × 11" sheet of paper can be used to teach principles of experimental design including– conditions for validity of experimentation, randomization, blocking, the use of factorial and fractional factorial designs, and the management of experimentation.

When Søren Bisgaard, Conrad Fung and I teach engineers about designed experiments, we find it very valuable to use a paper helicopter for illustration. We were introduced to this idea some years ago by Kip Rogers of Digital Equipment. Using the generic design shown in Figure 1 a "helicopter" can be made from an 8 1/2 × 11 sheet of paper in a minute or so.

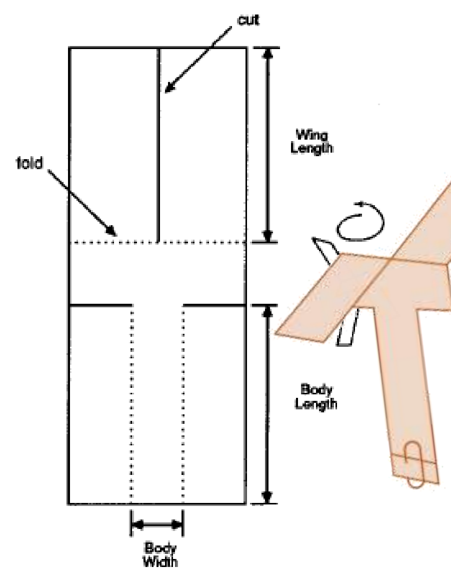


Figure 1. *A paper helicopter*

The scenario I'll describe requires three people whom I'll call Tom, Dick, and Harry. To make an

Dick times its fall with a stopwatch. We explain to the class that we would like to find an improved helicopter design which has a longer flight time. The helicopter can then be used to illustrate a number of important ideas.

VARIATION

We start by Tom dropping a helicopter made from blue paper. He drops it four times and we see that the results vary somewhat. This leads to a discussion of variation and to the introduction of the range and the standard deviation as measures of spread, and of the average as a measure of central tendency.

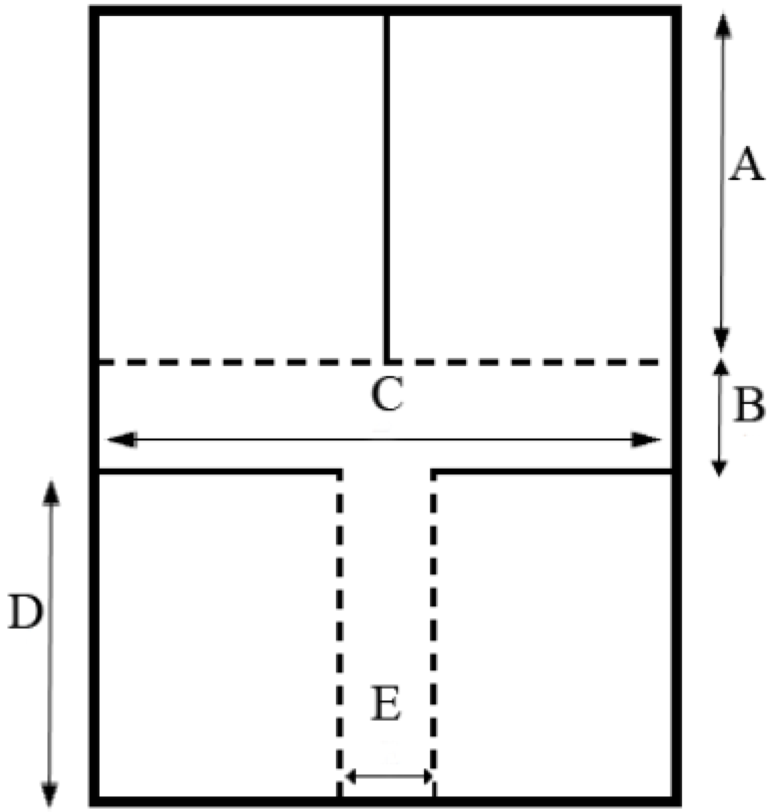
COMPARING MEAN FLIGHT TIMES

At this point Dick says "I don't think much of this helicopter design, I made this red helicopter yesterday and dropped it four times and I got an average flight time which was considerably longer than what we just got with the blue helicopter." So we put up the two sets of data, for the four runs made with the blue helicopter and the four runs made with the red helicopter, on the overhead projector and we show the two sets of averages and standard deviations. Eventually we demonstrate a simple test that shows that there is indeed a statistically significant difference in means, in favor of the runs made with the red helicopter.

VALIDITY OF THE EXPERIMENT

At this point Harry says "So the difference is statistically significant. So what? It doesn't necessarily

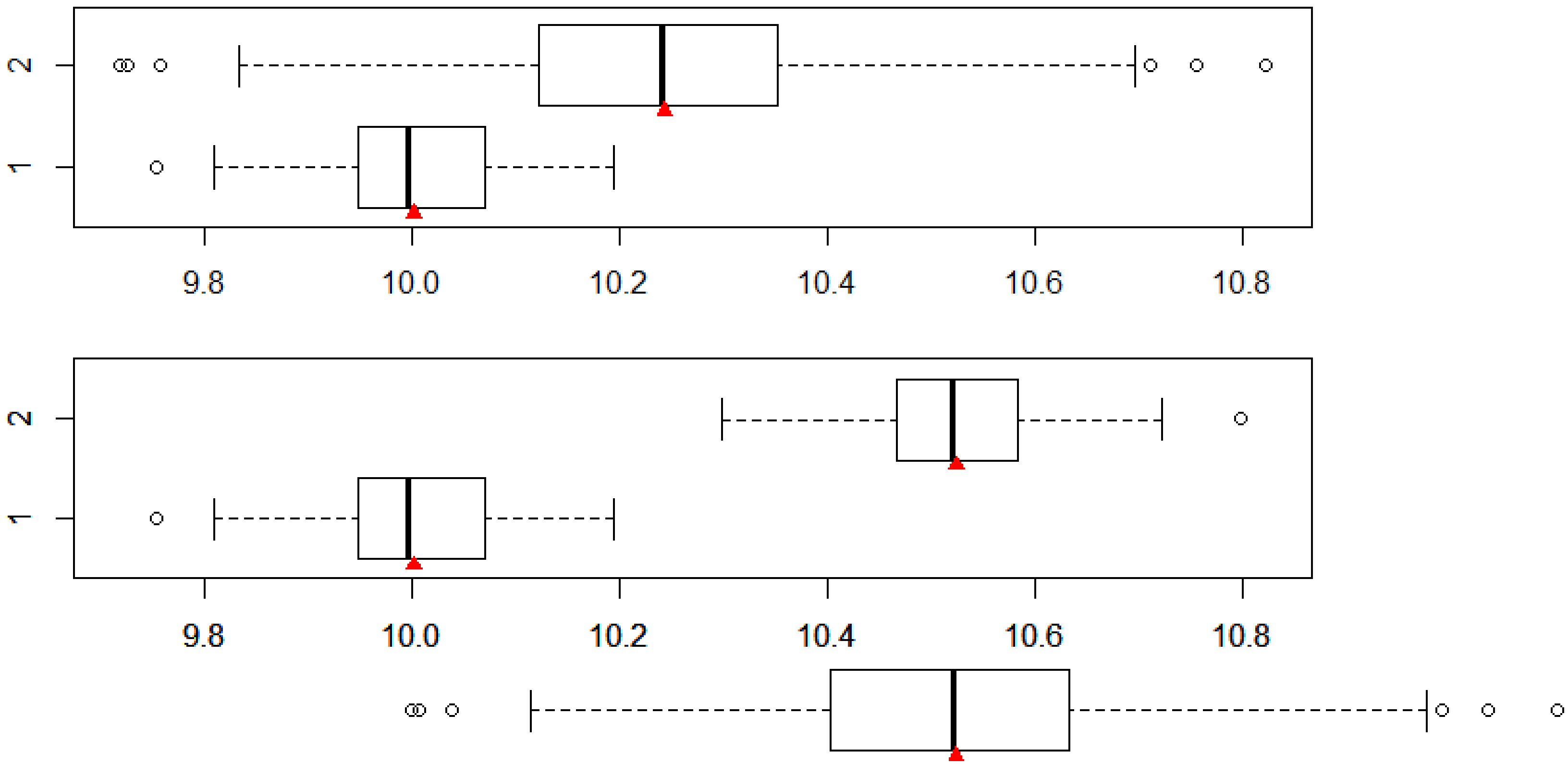
Planteamiento



- A: Largo de Alas
- B: Alto Rotor
- C: Ancho Rotor
- D: Alto Cuerpo
- E: Ancho Cuerpo
- F: Clip
- G: Gramaje Papel
- H: Celo Cuerpo

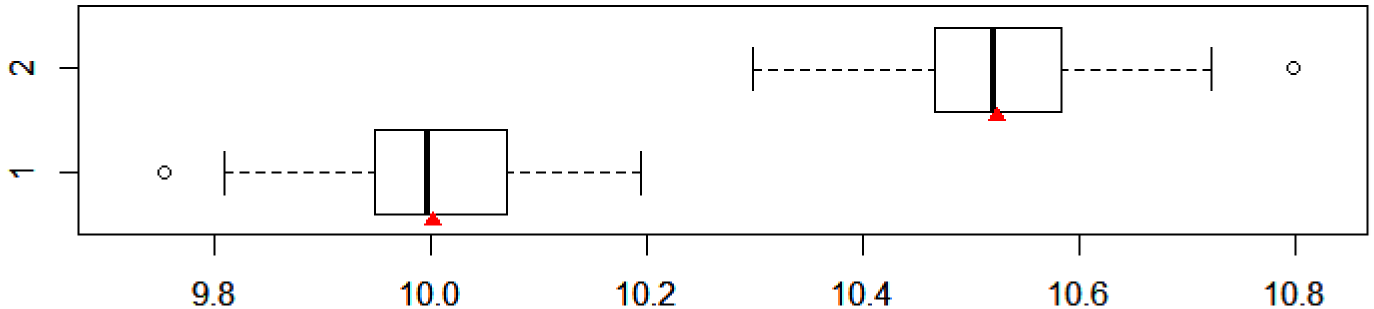
Dimensiones

Comparando tiempos medios de vuelo



Variación

Planificada

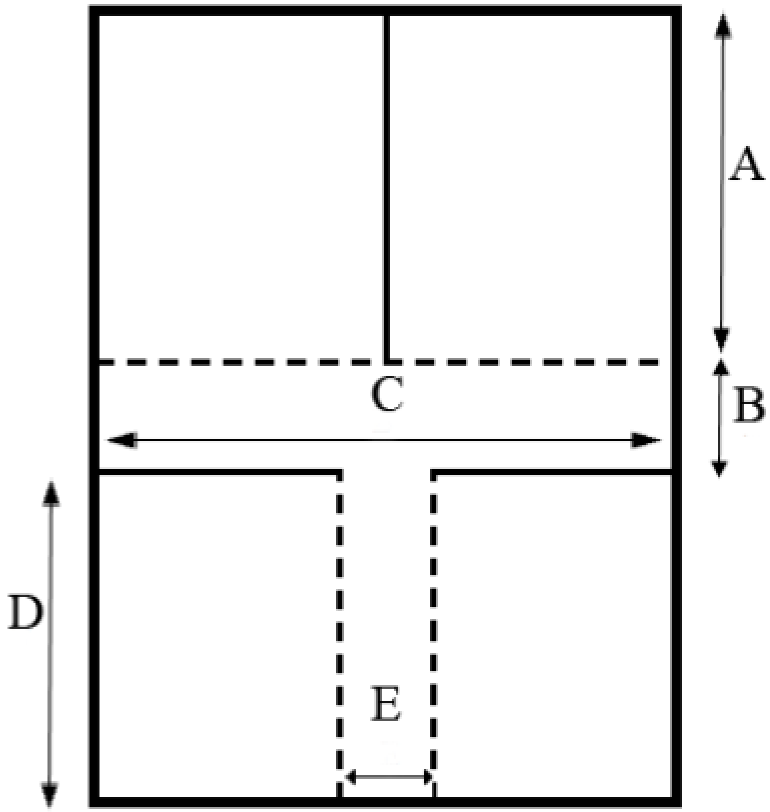


NO planificada

NO planificada

¡AZAR!

Experimentos/Diseños factoriales



- A: Largo de Alas
- B: Alto Rotor
- C: Ancho Rotor
- D: Alto Cuerpo
- E: Ancho Cuerpo
- F: Clip
- G: Gramaje Papel
- H: Celo Cuerpo

Diseños factoriales 2^k

$2^8 = 256$ experimentos “únicos”

Experimento/Diseño factorial fraccional

FACTORS		–	+	*EFFECT*
Paper Type	*(P)*	Regular	Bond	0.13
Wing Length	*(W)*	3.00"	4.75"	0.77
Body Length	*(L)*	3.00"	4.75"	–0.40
Body Width	*(B)*	1.25"	2.00"	0.02
Paper Clip	*(C)*	No	Yes	0.05
Fold	*(F)*	No	Yes	–0.10
Taped Body	*(T)*	No	Yes	–0.15
Taped Wing	*(M)*	No	Yes	0.17

Random Order	*Standard Order*	*P*	*W*	*L*	*B*	*C*	*F*	*T*	*M*	*Flight Time*
7	1	–	–	–	–	–	–	–	–	2.5
13	2	+	–	–	–	+	–	+	+	2.9
4	3	–	+	–	–	+	+	–	+	3.5
9	4	+	+	–	–	–	+	+	–	2.7
1	5	–	–	+	–	+	+	+	–	2.0
12	6	+	–	+	–	–	+	–	+	2.3
15	7	–	+	+	–	–	–	+	+	2.9
3	8	+	+	+	–	+	–	–	–	3.0
6	9	–	–	–	+	–	+	+	+	2.4
16	10	+	–	–	+	+	+	–	–	2.6
14	11	–	+	–	+	+	–	+	–	3.2
5	12	+	+	–	+	–	–	–	+	3.7
11	13	–	–	+	+	+	–	–	+	1.9
10	14	+	–	+	+	–	–	+	–	2.2
2	15	–	+	+	+	–	+	–	–	3.0
8	16	+	+	+	+	+	+	+	+	3.0

$2^4 = 16$ experimentos

en lugar de
$2^8 = 256$

Figure 2. *Results from sixteen run fractional factorial experiments showing the factor levels and the calculated main effects of the eight factors.*

Conclusiones de los experimentos

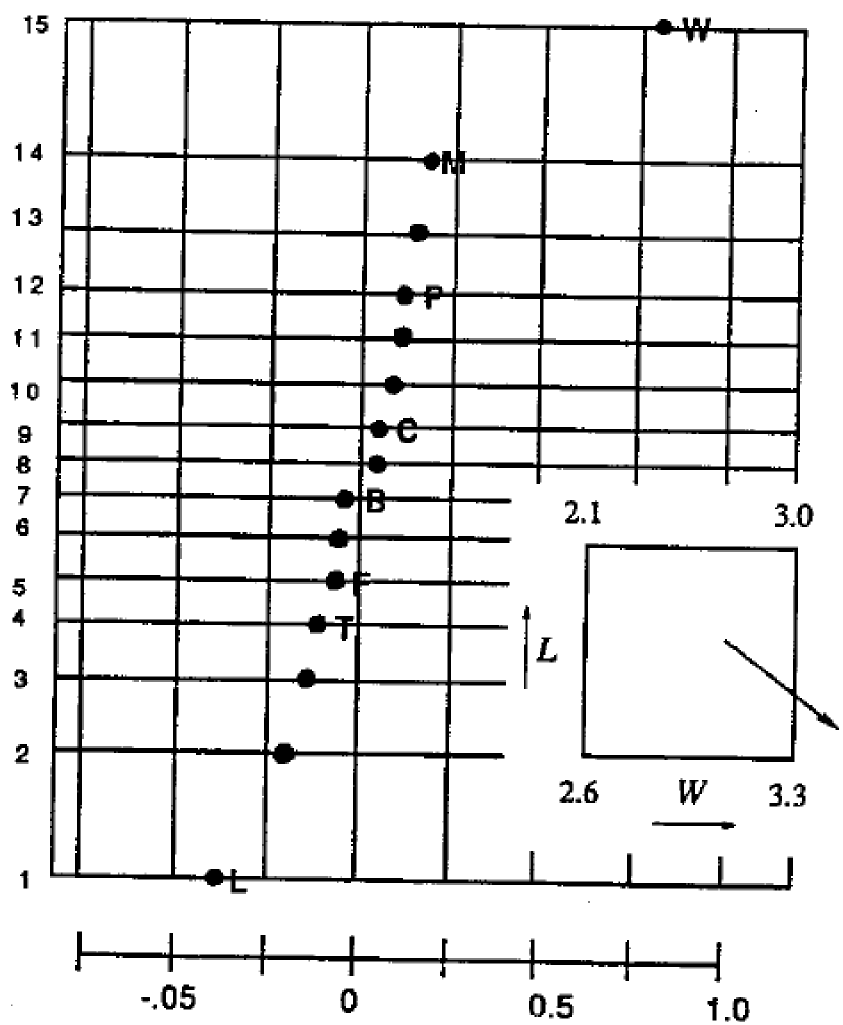
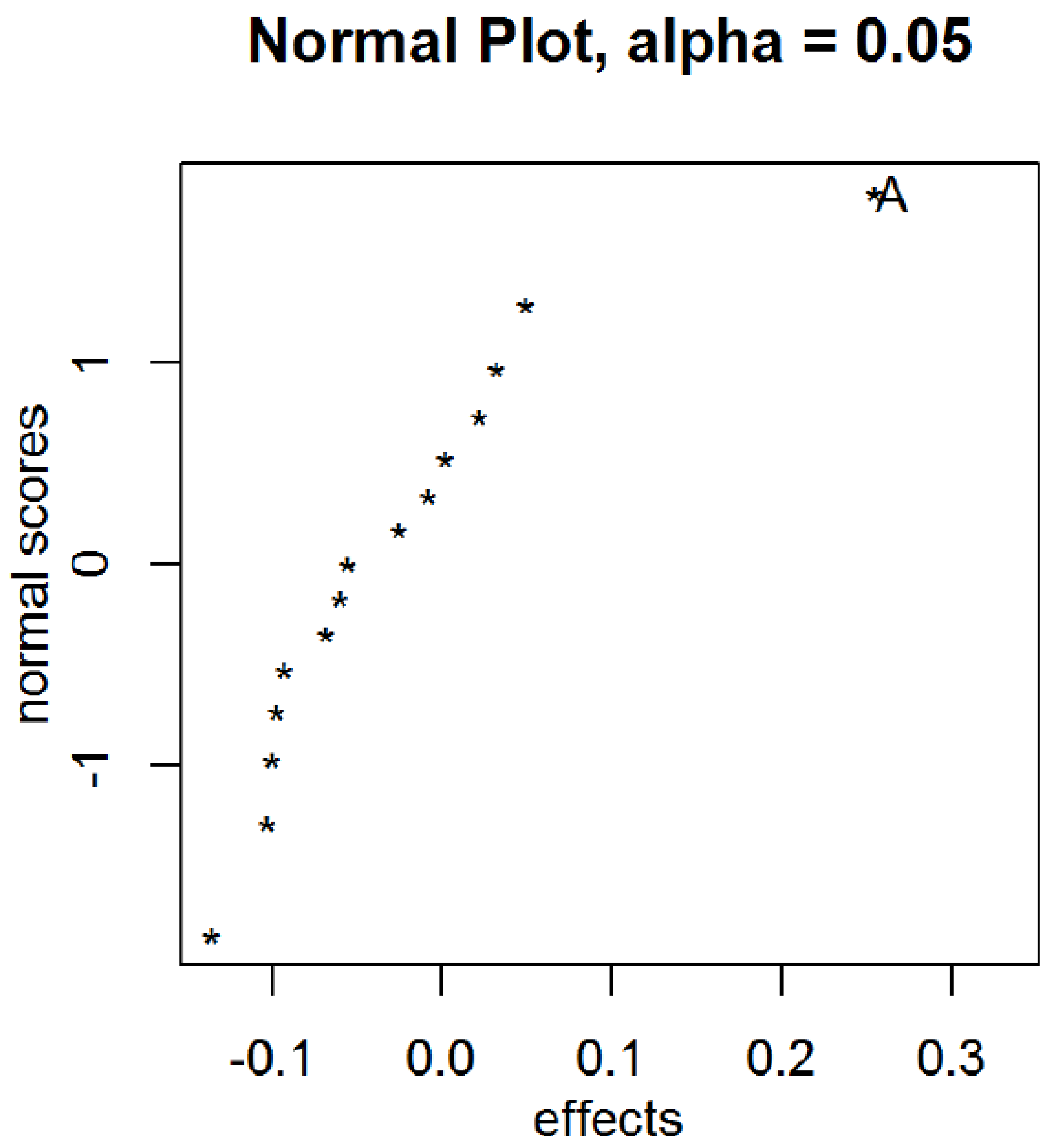
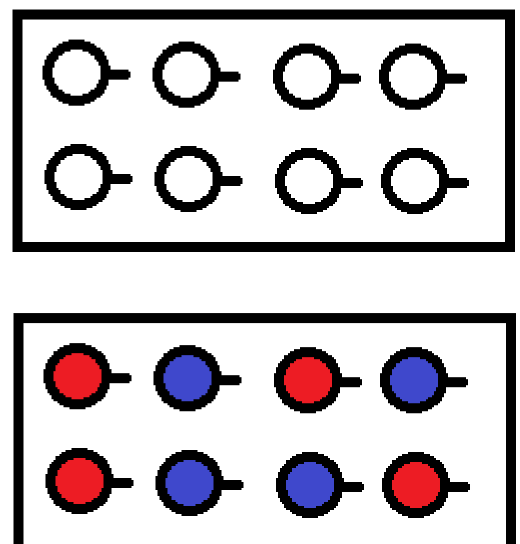


Figure 3. *A normal plot of the effects from the helicopter experiment. The inset diagram summarizes the conclusions.*

En el “taller” también salió “A”



The lady tasting tea experiment (Fisher)



Las cuentas, el p-valor

Tea-Tasting Distribution Assuming the Null Hypothesis

Success count	Permutations of selection	Number of permutations
0	oooo	1 × 1 = 1
1	ooox, ooxo, oxoo, xooo	4 × 4 = 16
2	ooxx, oxox, oxxo, xoxo, xxoo, xoox	6 × 6 = 36
3	oxxx, xoxx, xxox, xxxo	4 × 4 = 16
4	xxxx	1 × 1 = 1
Total		70

$\frac{17}{70} = 0.2428 = 24.28\%$

$\frac{1}{70} = 0.0143 = 1.43\%$

!!

Paradoja de Simpson

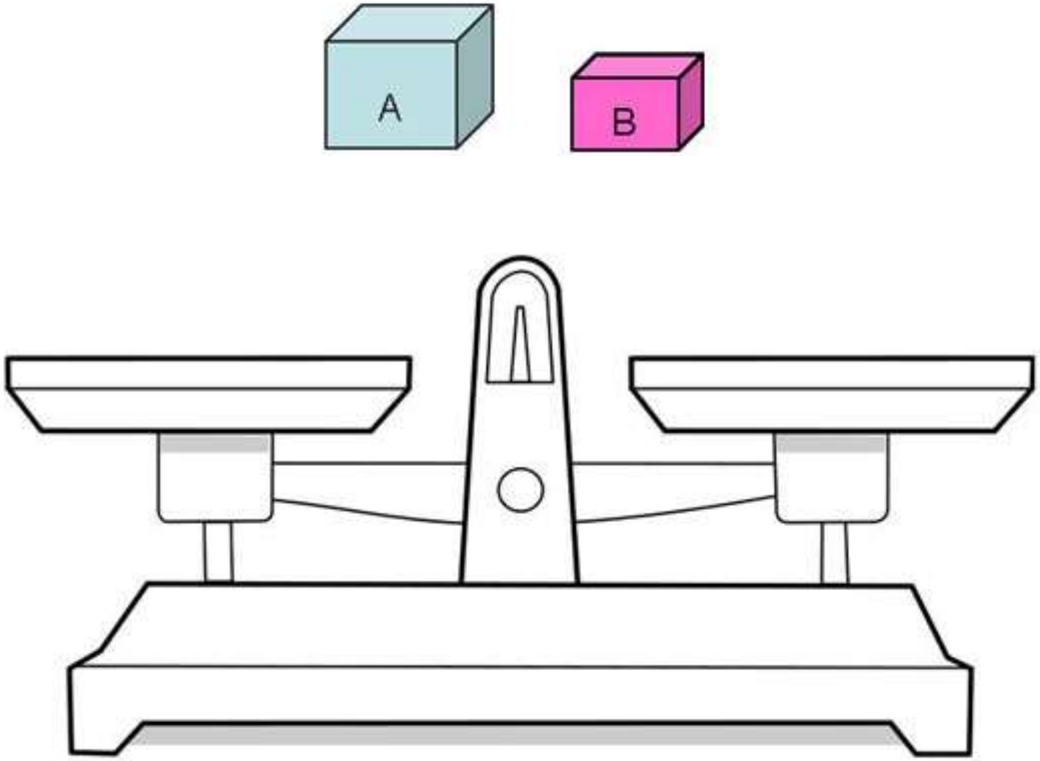
Ensayo (80)	Mejoran	No Mejoran	%
Curamina (40)	20	20	50%
Fraudol (40)	24	16	60%

Hombres (40)	Mejoran	No Mejoran	%
Curamina (30)	12	18	40%
Fraudol (10)	3	7	30%

Mujeres (40)	Mejoran	No Mejoran	%
Curamina (10)	8	2	80%
Fraudol (30)	21	9	70%

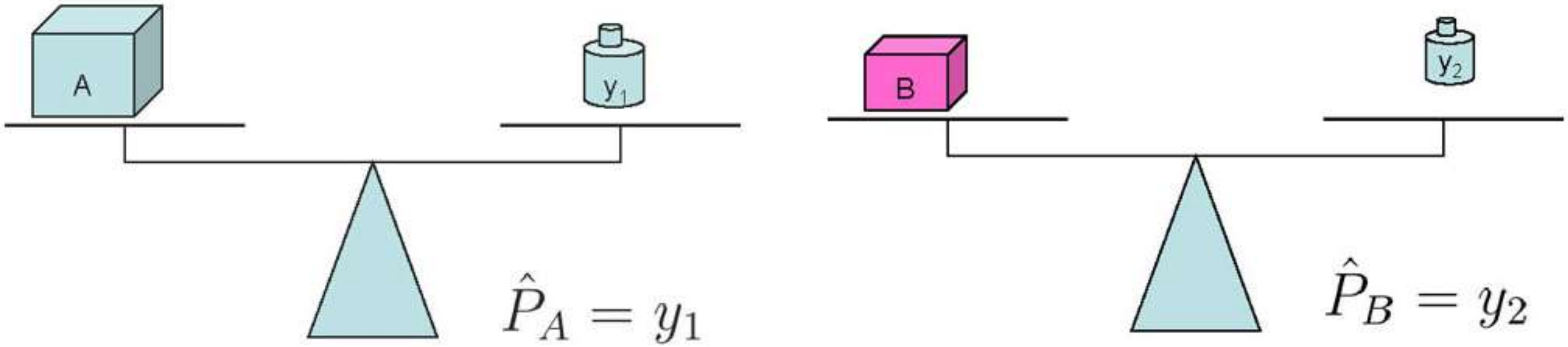
!!

Mejor forma de pesar 2 objetos



Restricción:
¡ Sólo 2 pesadas!

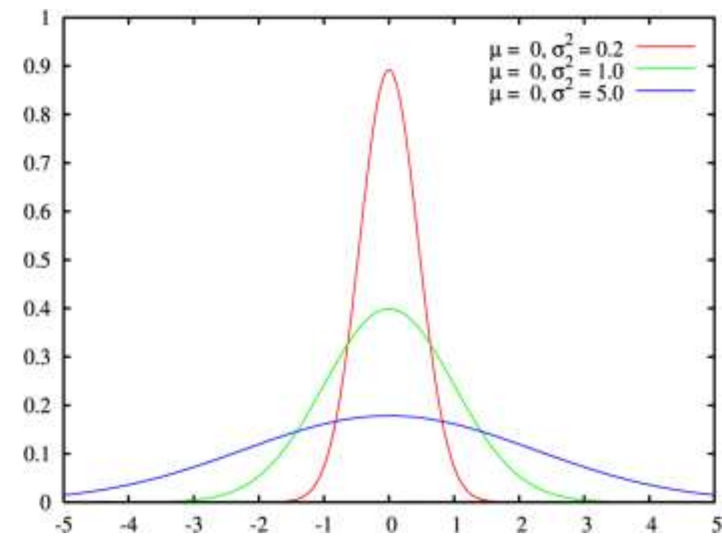
Forma obvia



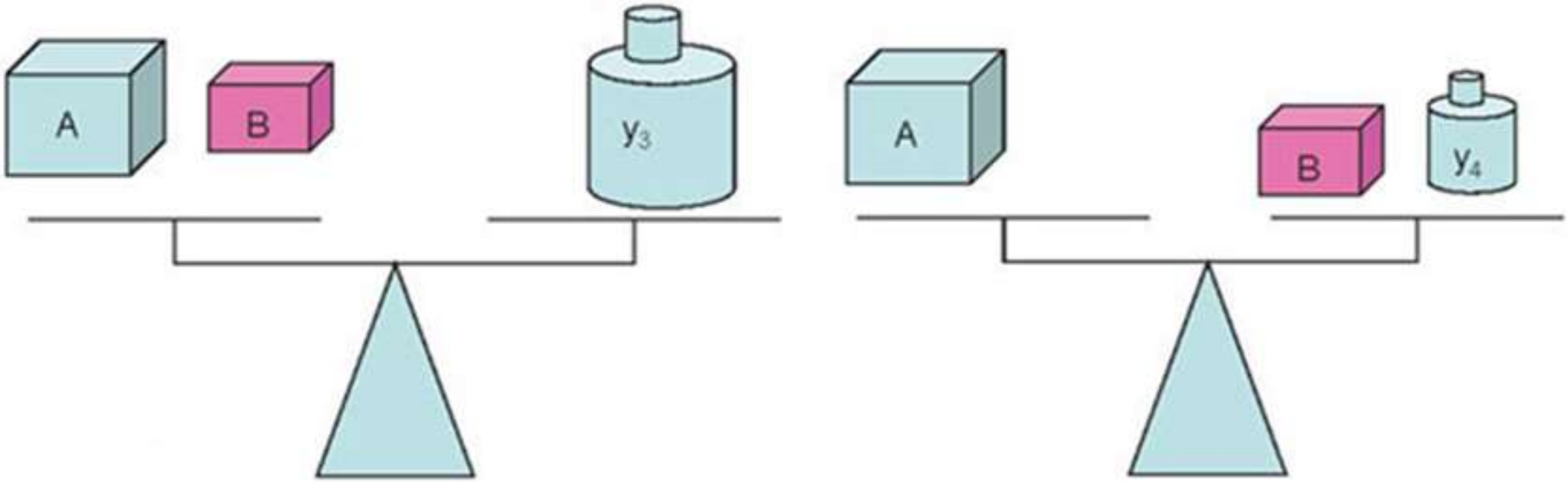
$\hat{P}_A = y_1$ $\hat{P}_B = y_2$

$$y = P + \epsilon, \qquad \epsilon \equiv \mathcal{N}(0, \sigma^2)$$

$$\mathrm{Var}(y) = \sigma^2 \longrightarrow \mathrm{Var}(\hat{P}_A) = \sigma^2$$
$$\searrow \mathrm{Var}(\hat{P}_B) = \sigma^2$$



Forma (diseño) óptimo



$$y_3 = P_A + P_B + \epsilon_3 \qquad y_4 = P_A - P_B + \epsilon_4$$

$$\hat{P}_A = \frac{y_3 + y_4}{2} \quad \hat{P}_B = \frac{y_3 - y_4}{2} \Rightarrow \mathrm{Var}(\hat{P}_A) = \mathrm{Var}(\hat{P}_B) = \frac{\sigma^2}{2}$$

!!

Optimal Experimental Design

In a regression framework

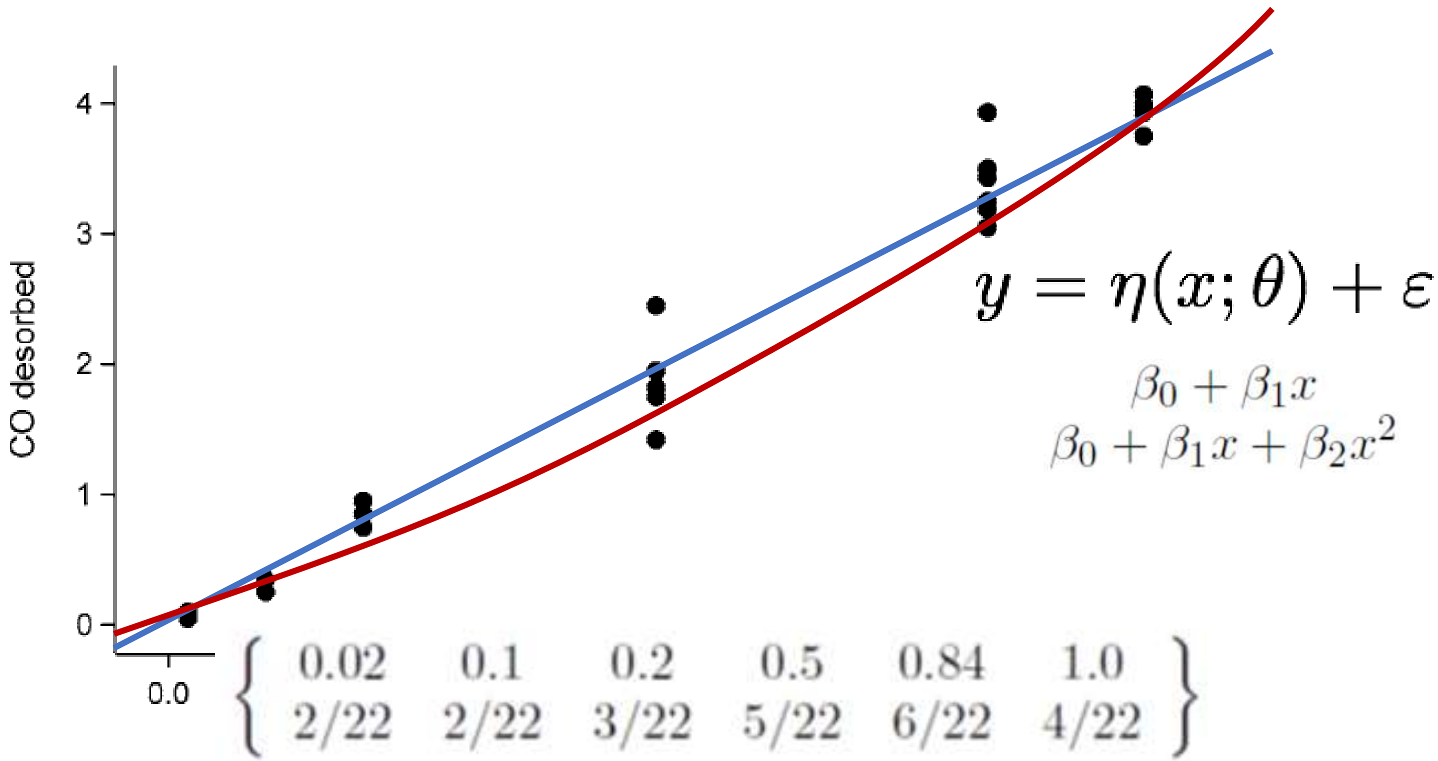


FIG. 1.1. Example 1.1 the desorption of carbon monoxide. Yield (carbon monoxide desorbed) against K/C ratio.

Optimal designs and efficiencies

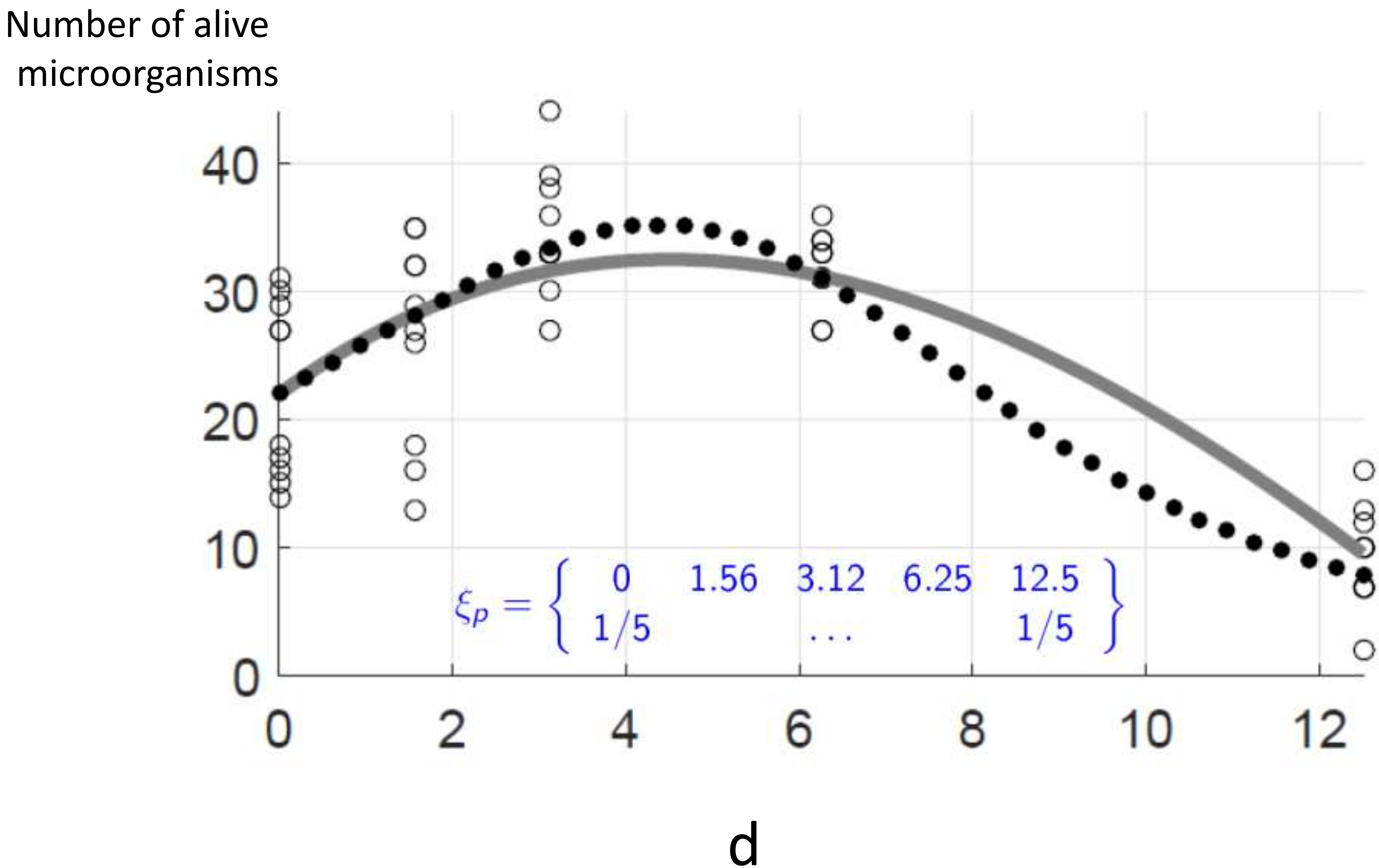
TABLE 11.2. Example 1.1. Efficiency for a variety of purposes of the design* of Table 1.1 for measuring the desorption of carbon monoxide

Model	Optimality	Weight at design points				Efficiency
$\eta(x)$	criterion	0	$\sqrt{(2)}-1$	0.5	1	(%)
$\beta_0+\beta_1 x$	D	1/2	—	—	1/2	69.5
$\beta_0+\beta_1 x+\beta_2 x^2$	D	1/3	—	1/3	1/3	81.7
$\beta_0+\beta_1 x+\beta_2 x^2$	D_S for β_2	1/4	—	1/2	1/4	47.4
$\beta_1 x$	D	—	—	—	1	43.7
$\beta_1 x+\beta_2 x^2$	D	—	—	1/2	1/2	62.4
$\beta_1 x+\beta_2 x^2$	D_S for β_2	—	$\sqrt{2}/2$	—	$1-\sqrt{2}/2$	47.2

* The design region is scaled to be $\mathcal{X}=[0,1]$. The design of Table 1.1 is then

$$\xi_{22}=\left\{\begin{array}{cccccc} 0.02 & 0.1 & 0.2 & 0.5 & 0.84 & 1.0 \\ 2/22 & 2/22 & 3/22 & 5/22 & 6/22 & 4/22 \end{array}\right\}.$$

Hormesis related problems



OED Tools

- Model :

$$\mu(d,\theta)\begin{cases} = \theta_1 - \theta_2 d + \frac{\theta_3}{\theta_4}(1 - e^{-\theta_4 d}) \\ = \frac{\theta_1 + \theta_2 d}{1 + e^{-\theta_3} d^{\theta_4}} \end{cases}$$

- Designs:

$$\xi = \left\{ \begin{array}{cccc} d_1 & d_2 & \dots & d_k \\ w_1 & w_2 & \dots & w_k \end{array} \right\} \quad d_i \in \Omega$$

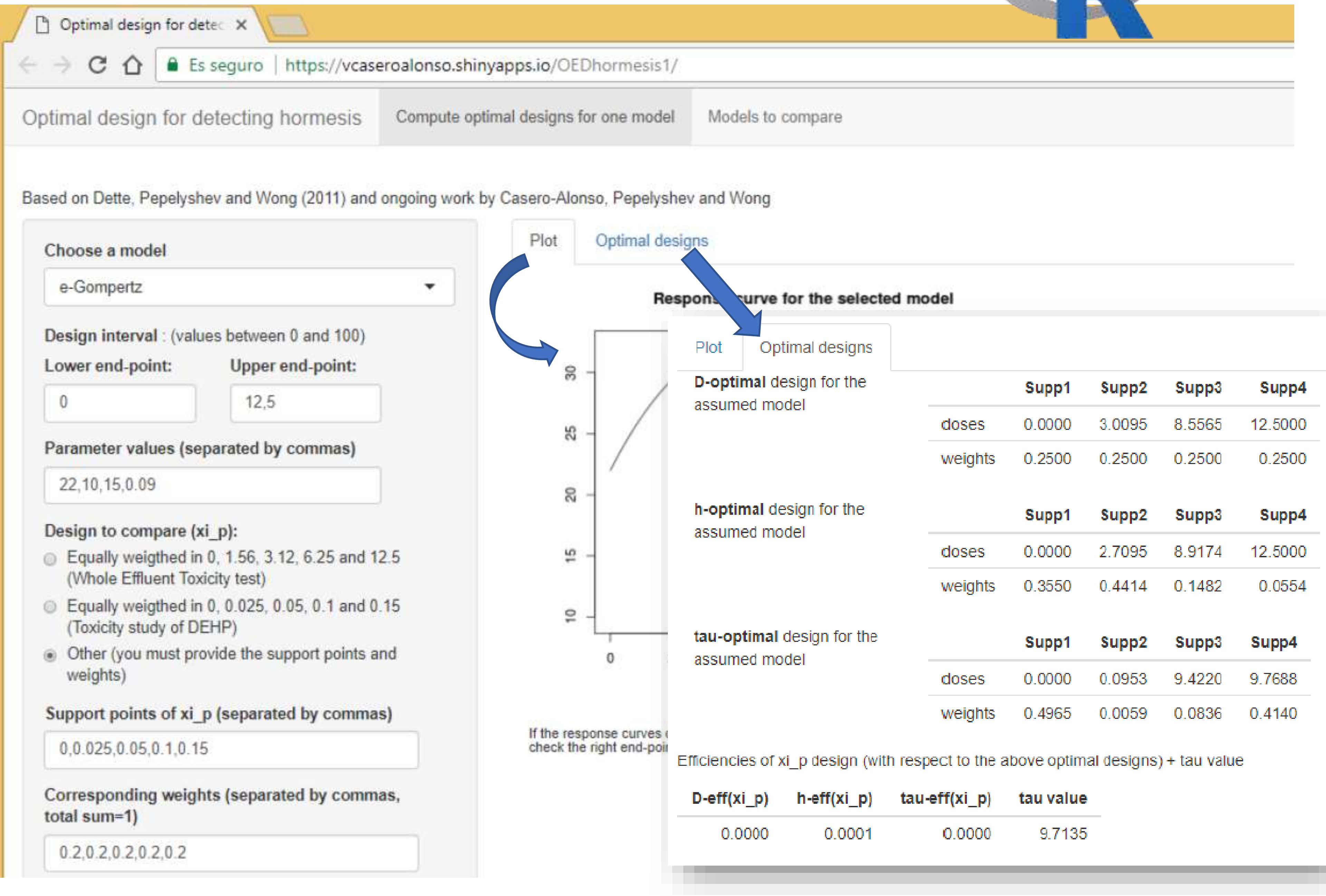
- Information matrix:

$$M(\xi,\theta) = \sum_{i=1}^{k} w_i f(d_i,\theta) f^T(d_i,\theta),$$

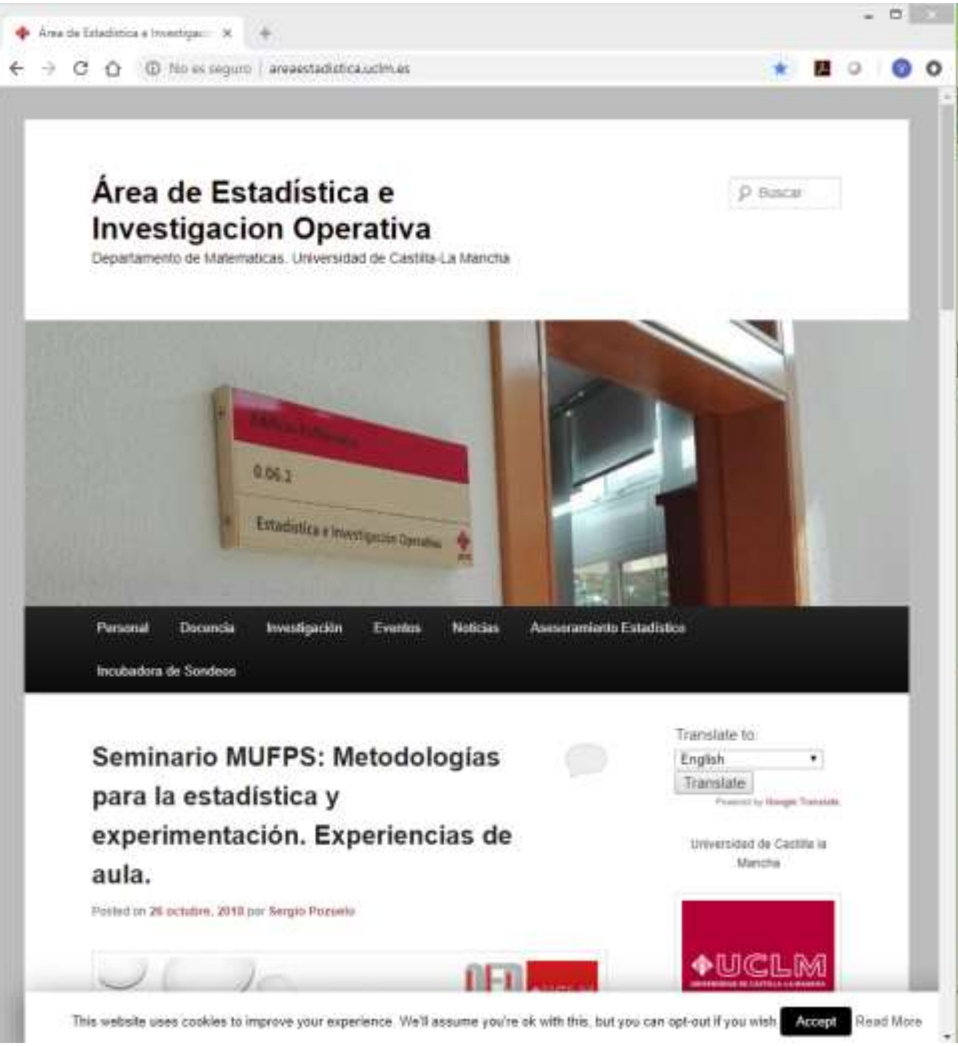
$$\text{where } f(d,\theta) = \frac{\partial \mu(d,\theta)}{\partial \theta}.$$

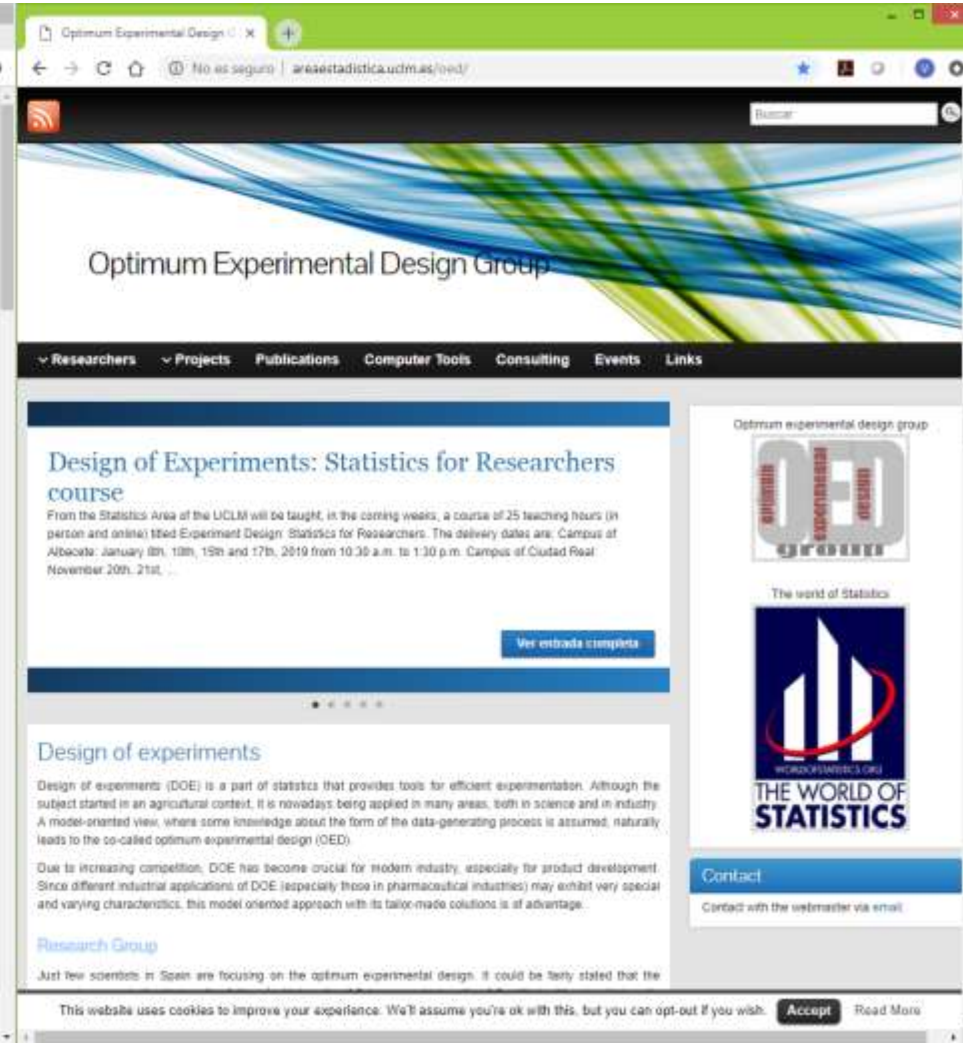
- Criteria: …

Online R tools



areaestadistica.uclm.es /oed





El experimento del helicóptero de papel de G. E. Box. La estadística antes de la experimentación

¡Muchas gracias por su atención!

Victormanuel.casero@uclm.es

UCLM
UNIVERSIDAD DE CASTILLA-LA MANCHA

OED group
optimum experimental design

BIOSTATNET

imaci
Instituto de Matemática Aplicada a la Ciencia y la Ingeniería