

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

An Accidental Statistician

The Life and Memories of George E. P. Box



George E. P. Box

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" x 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 x 11 sheet of paper in a minute or so.

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 mean it's because of the different helicopter design

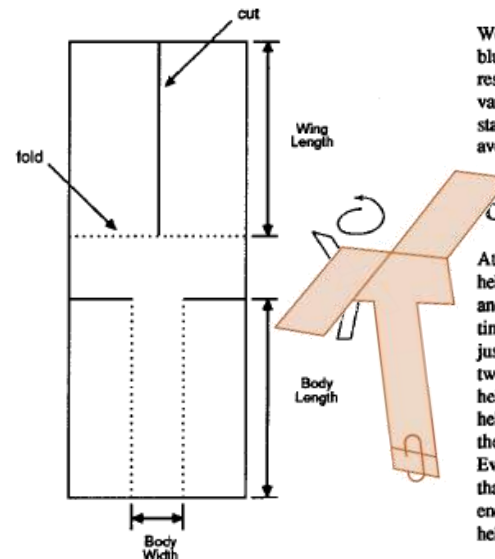
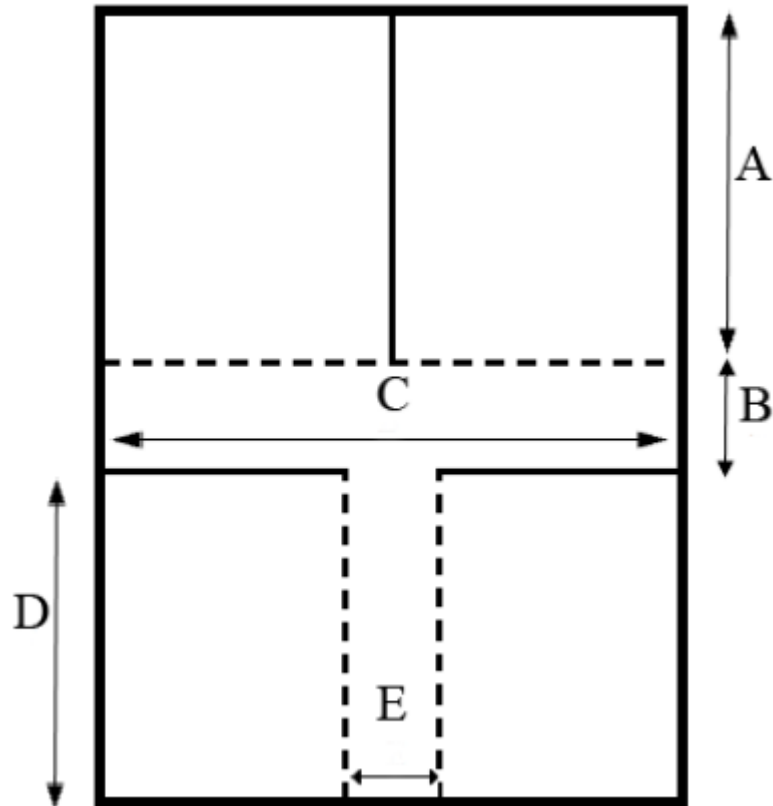


Figure 1. A paper helicopter

The scenario I'll describe requires three people whom I'll call Tom, Dick, and Harry. To make an experiment, Tom starts on a white sheet of paper

Planteamiento



○ A: Largo de Alas

○ B: Alto Rotor

○ C: Ancho Rotor

○ D: Alto Cuerpo

○ E: Ancho Cuerpo

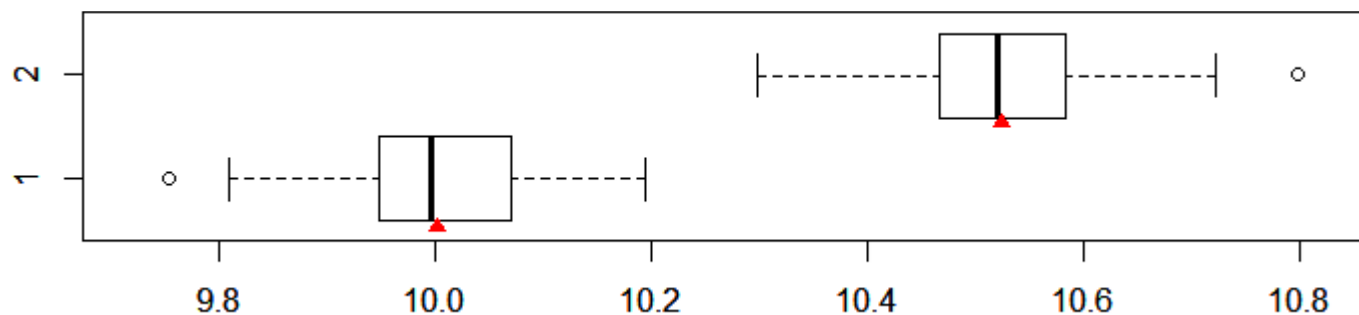
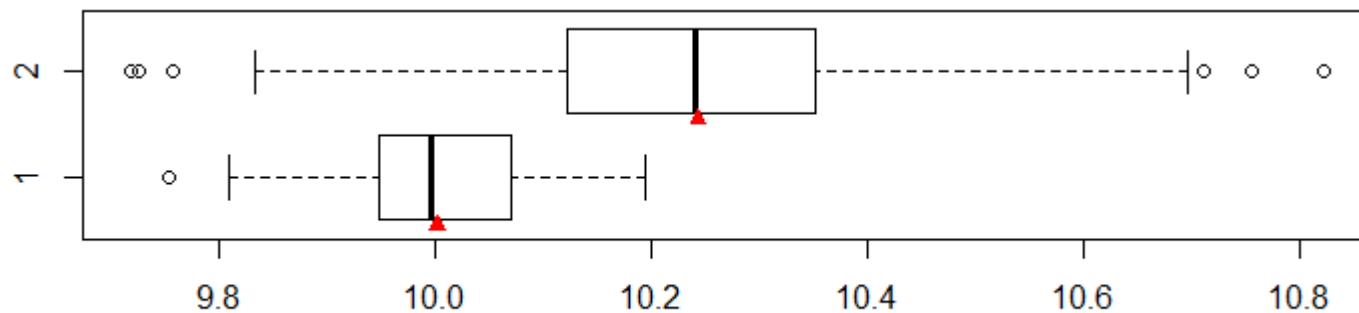
○ F: Clip

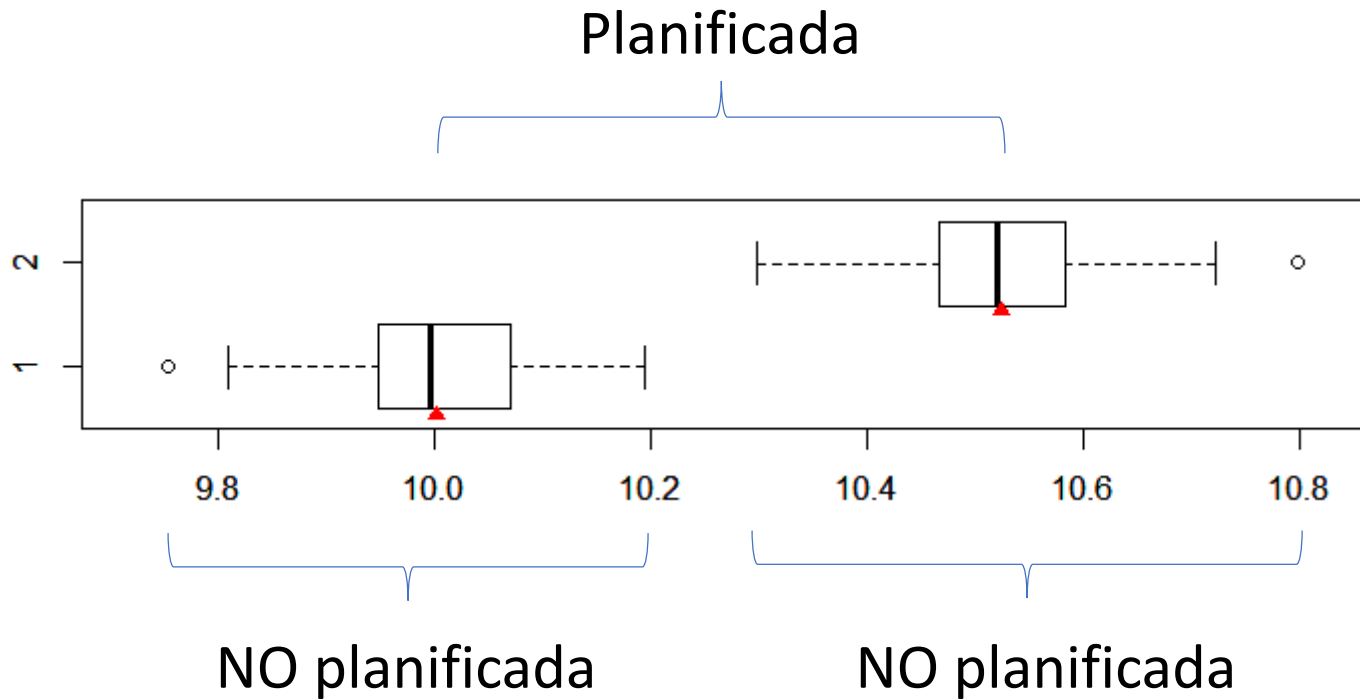
○ G: Gramaje Papel

○ H: Celo Cuerpo

Dimensiones

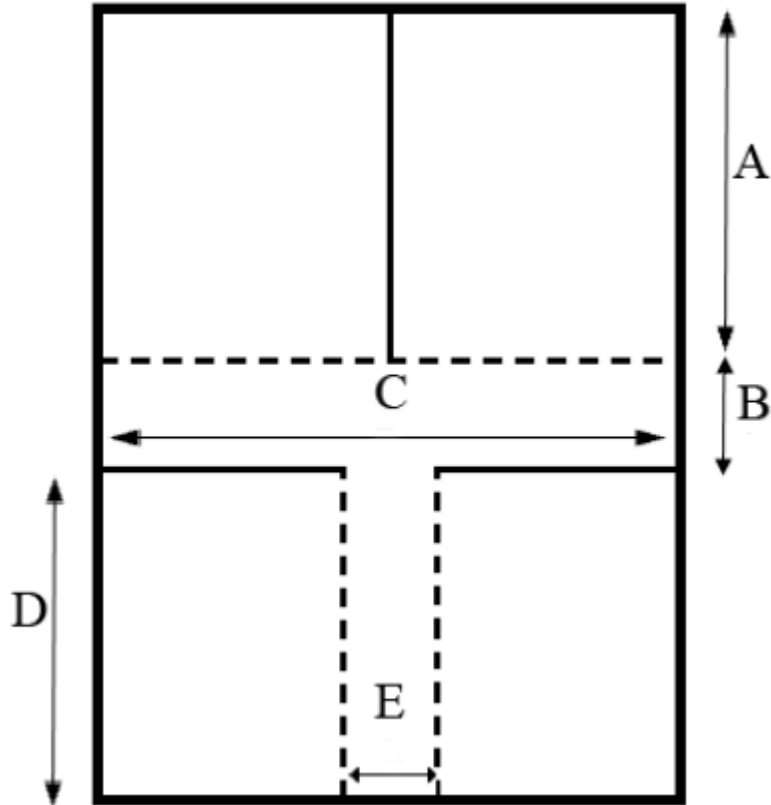
Comparando tiempos medios de vuelo





¡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

3

George Box

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

4

Teaching Engineers Experimental Design With a Paper Helicopter

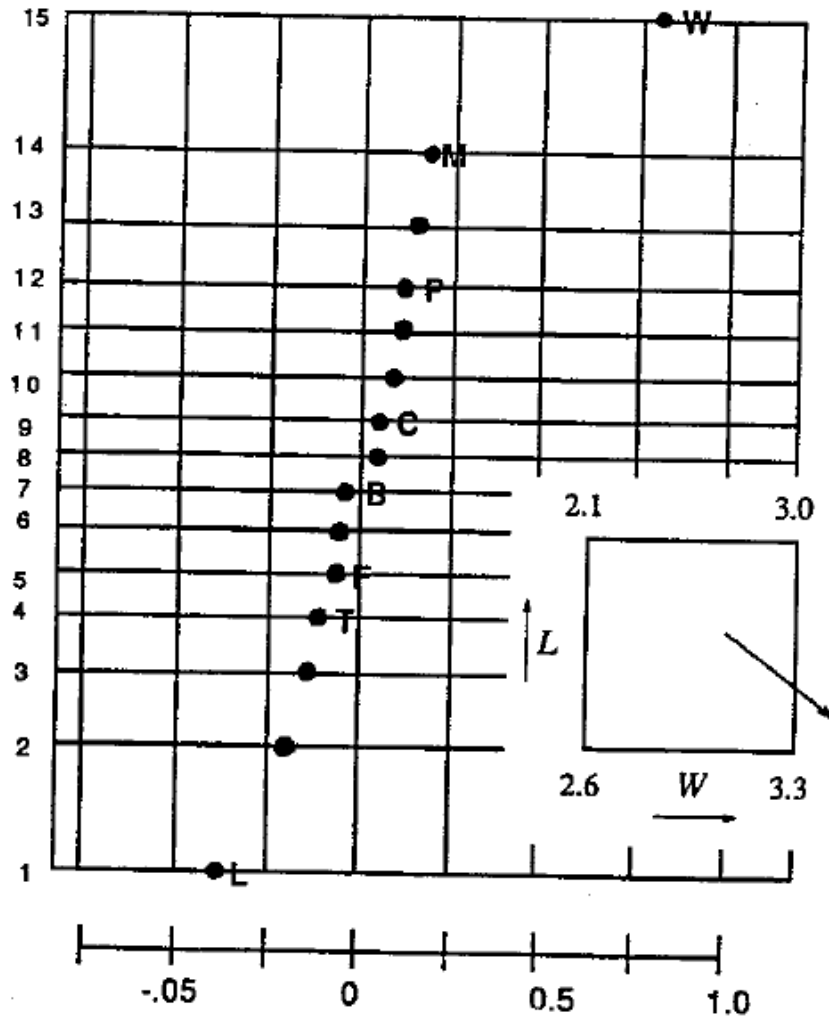
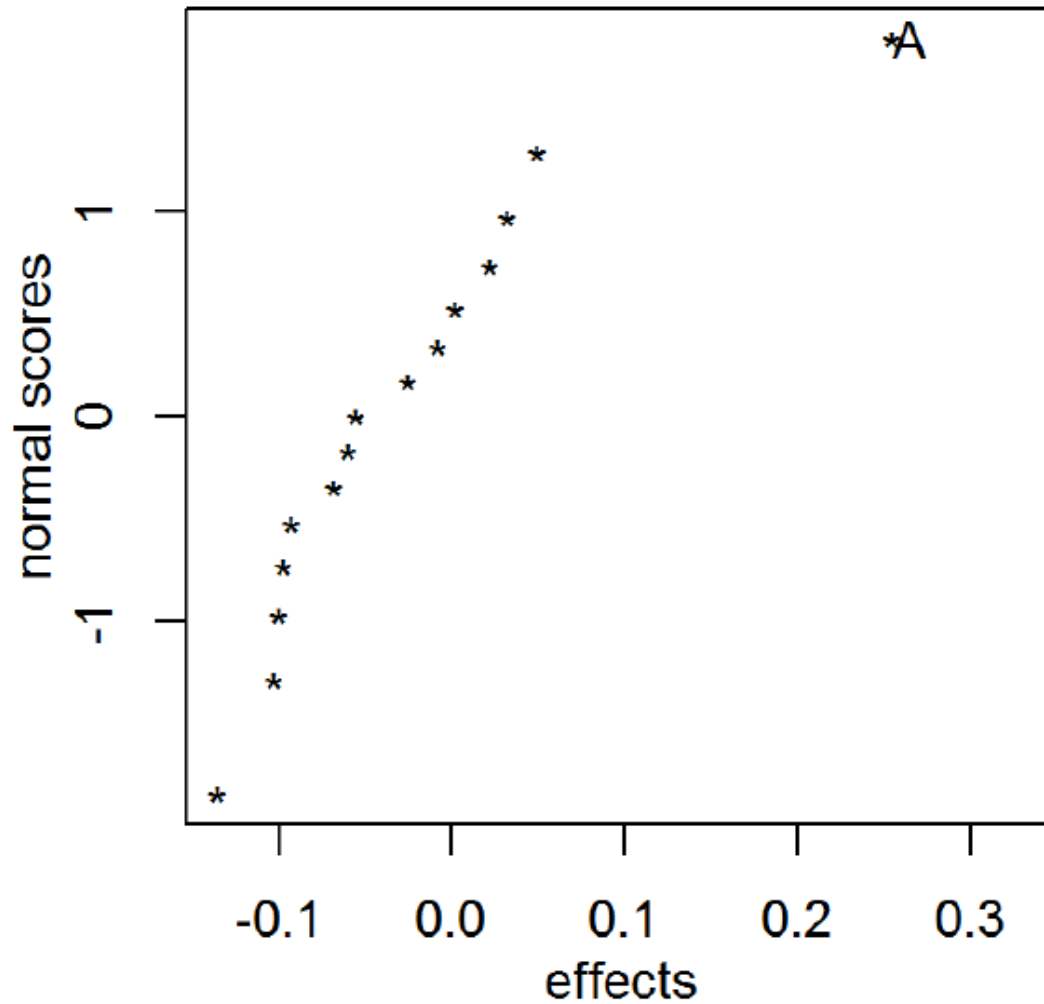


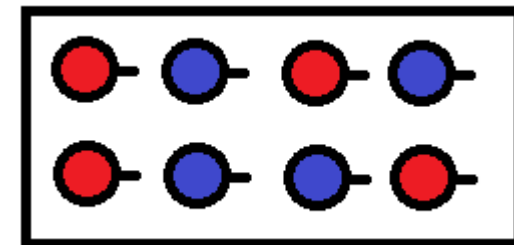
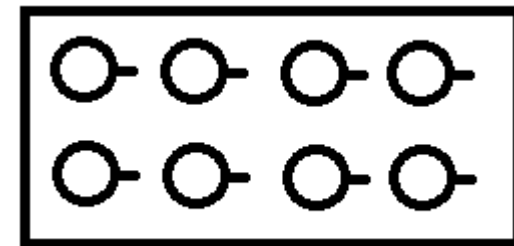
Figure 3. A normal plot of the effects from the helicopter experiment. The inset diagram summarizes the conclusions.

En el “taller” también salió “A”

Normal Plot, alpha = 0.05




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	$1 \times 1 = 1$
1	OOOX, OOXO, OXOO, XOOO	$4 \times 4 = 16$
2	OOXX, OXOX, OXXO, XOXO, XXOO, XOOX	$6 \times 6 = 36$
3	OXXX, XOOX, XXOX, XXXO	$4 \times 4 = 16$
4	XXXX	$1 \times 1 = 1$
Total		70



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



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

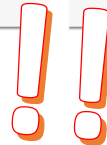


Paradoja de Simpson

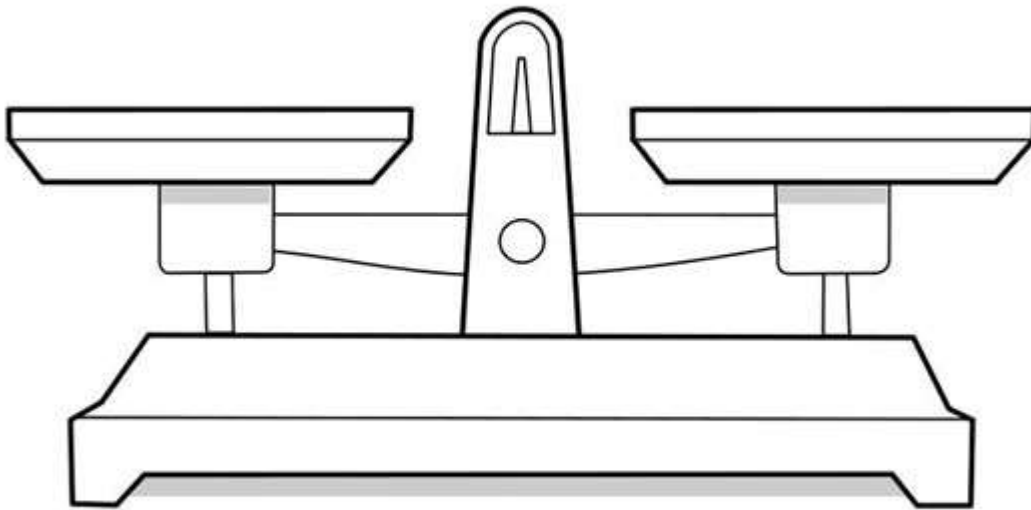
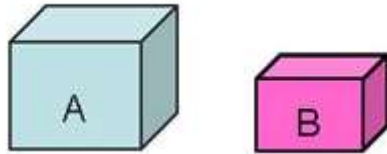
Ensayo (80)	Mejoran	No Mejoran	%
Curamina (40)	20	20	50%
Fraudol (40)	24	16	 60%

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%

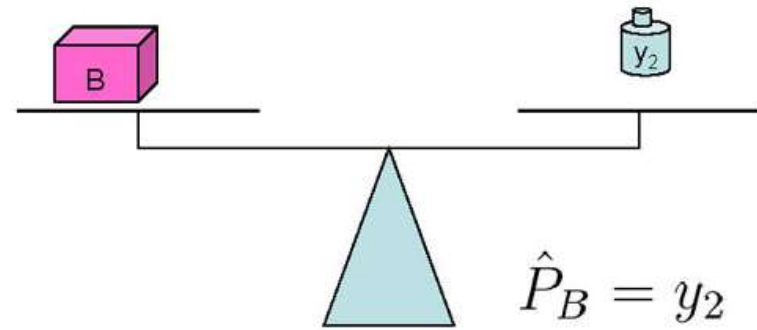
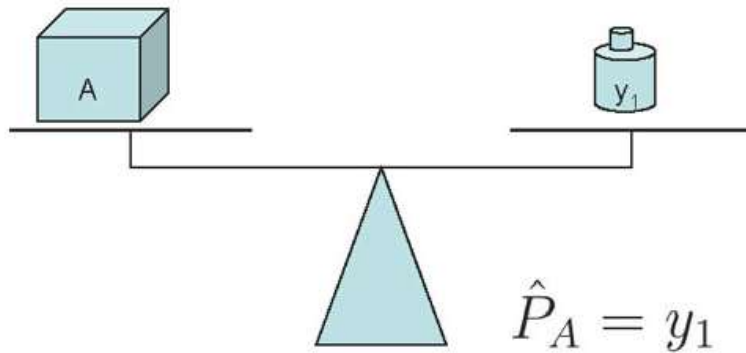


Mejor forma de pesar 2 objetos



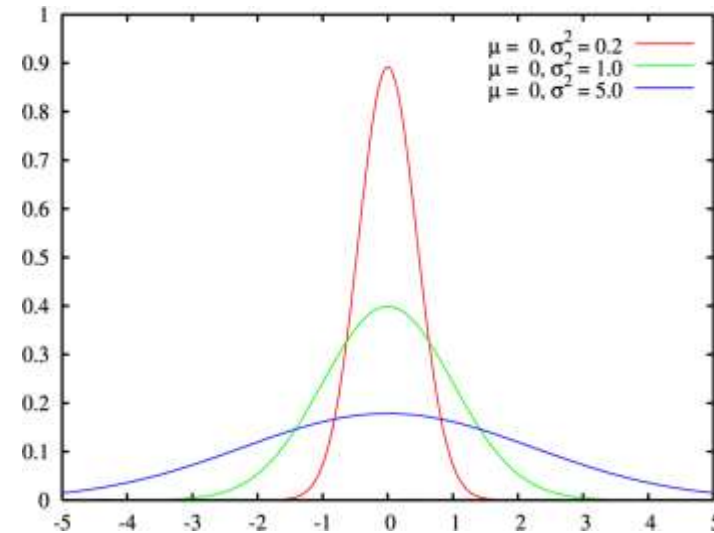
Restricción:
¡ Sólo 2 pesadas!

Forma obvia

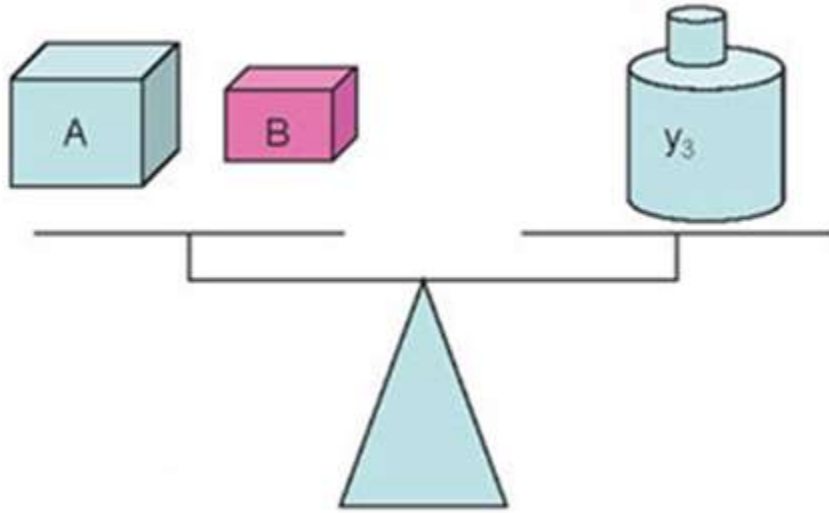


$$y = P + \epsilon, \quad \epsilon \equiv \mathcal{N}(0, \sigma^2)$$

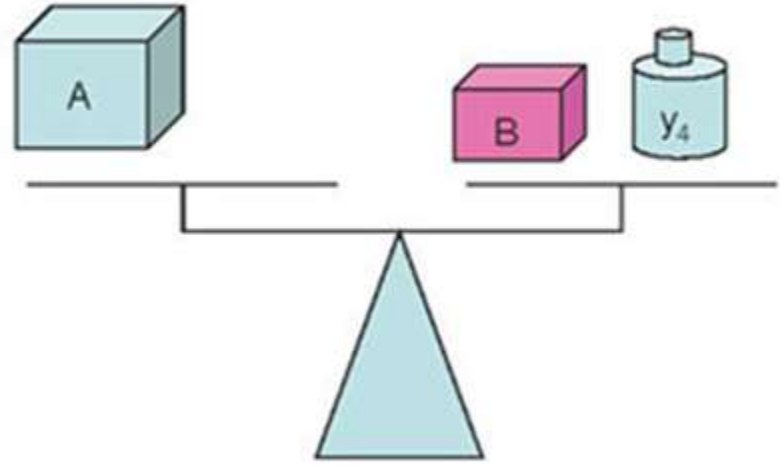
$$\text{Var}(y) = \sigma^2 \begin{cases} \longrightarrow \text{Var}(\hat{P}_A) = \sigma^2 \\ \searrow \text{Var}(\hat{P}_B) = \sigma^2 \end{cases}$$



Forma (diseño) óptimo



$$y_3 = P_A + P_B + \epsilon_3$$



$$y_4 = P_A - P_B + \epsilon_4$$

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

!!

Optimal Experimental Design

In a regression framework

Regression

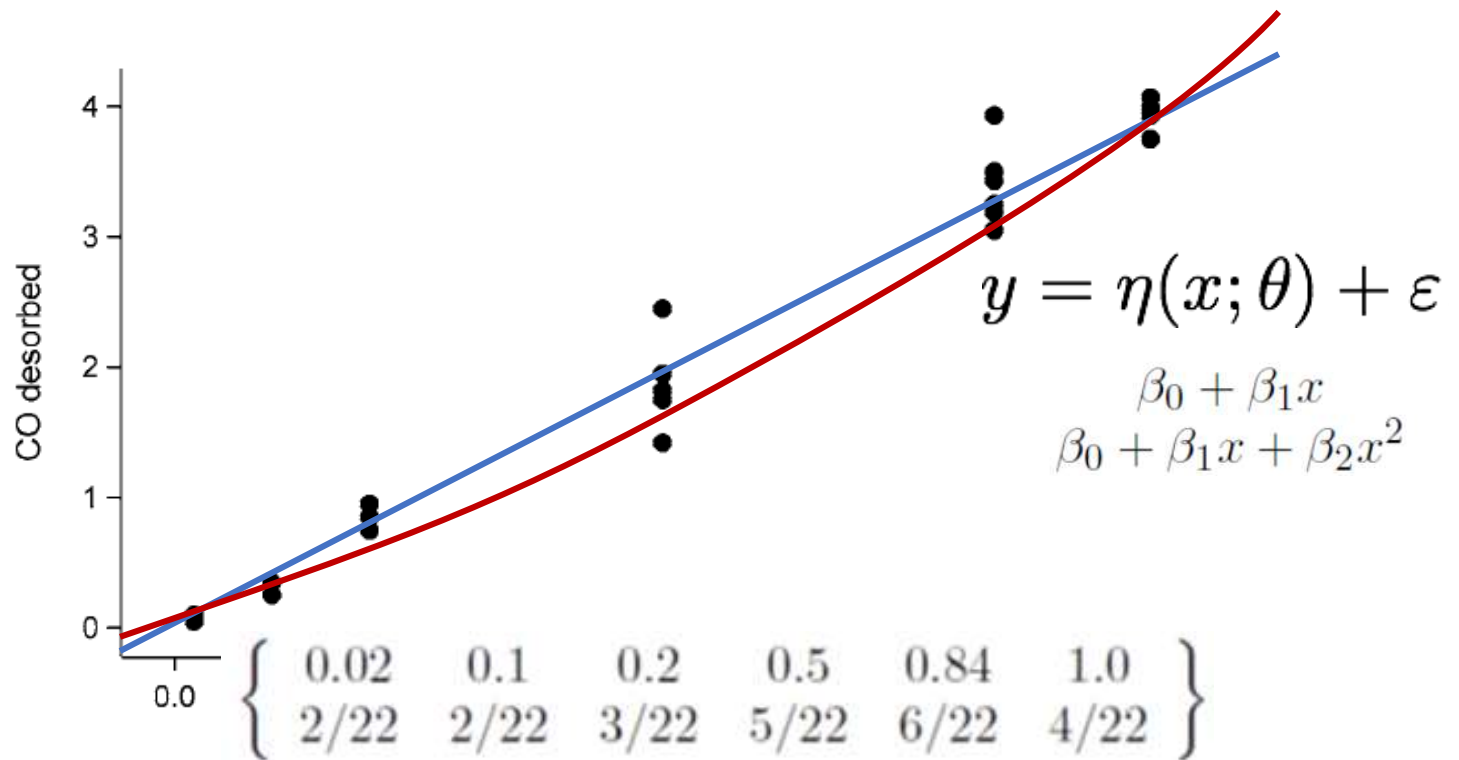


FIG. 1.1. Example 1.1 the desorption of carbon monoxide. Yield (carbon monoxide desorbed) against K/C ratio.

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

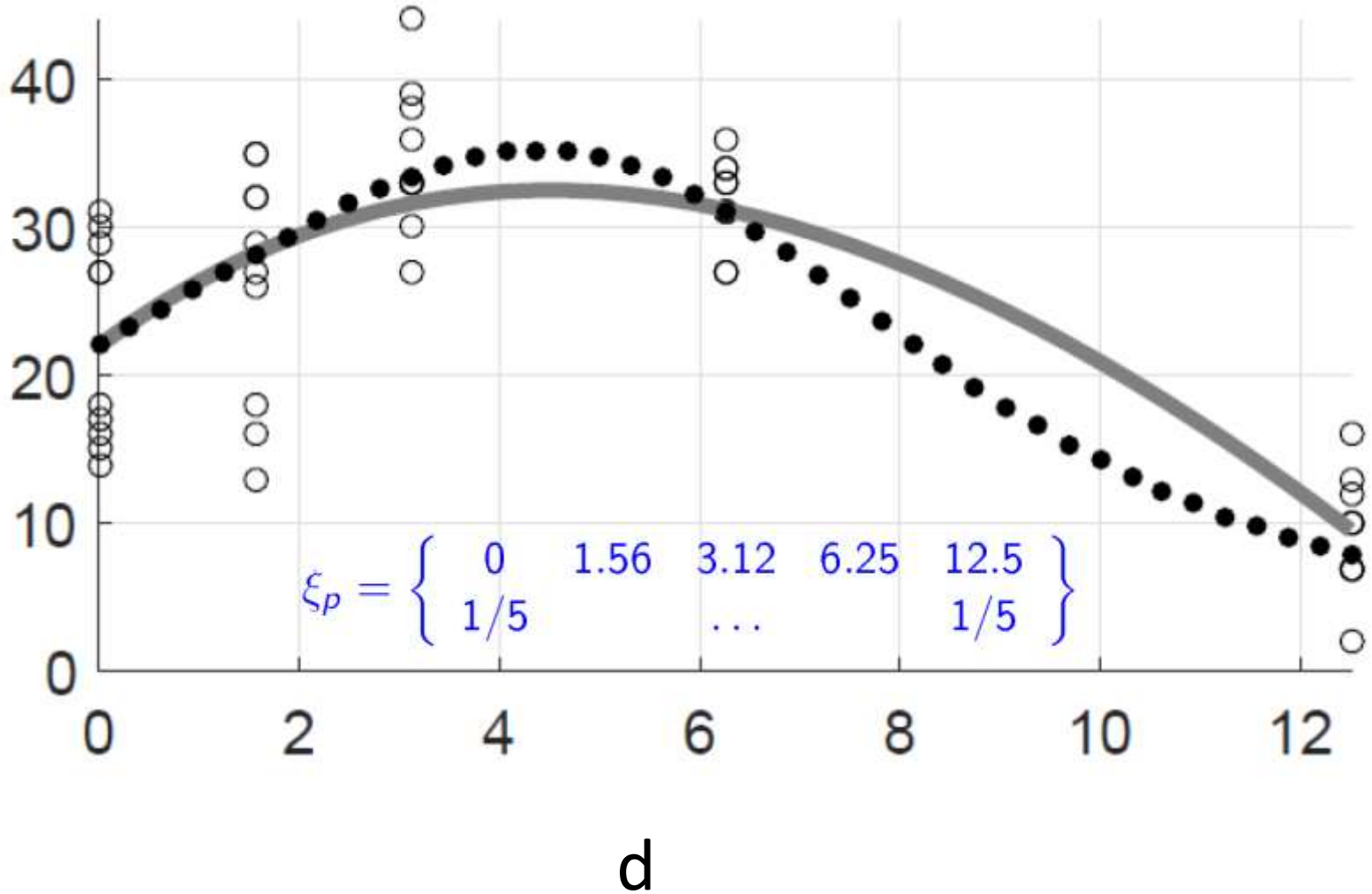
Model $\eta(x)$	Optimality criterion	Weight at design points				Efficiency (%)
		0	$\sqrt{2}/2$	0.5	1	
$\beta_0 + \beta_1x$	D	1/2	—	—	1/2	69.5
$\beta_0 + \beta_1x + \beta_2x^2$	D	1/3	—	1/3	1/3	81.7
$\beta_0 + \beta_1x + \beta_2x^2$	D _S for β_2	1/4	—	1/2	1/4	47.4
β_1x	D	—	—	—	1	43.7
$\beta_1x + \beta_2x^2$	D	—	—	1/2	1/2	62.4
$\beta_1x + \beta_2x^2$	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

Number of alive
microorganisms



- 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: ...

Based on Dette, Pepelyshev and Wong (2011) and ongoing work by Casero-Alonso, Pepelyshev and Wong

Choose a model
 e-Gompertz

Design interval : (values between 0 and 100)
Lower end-point: 0 **Upper end-point:** 12.5

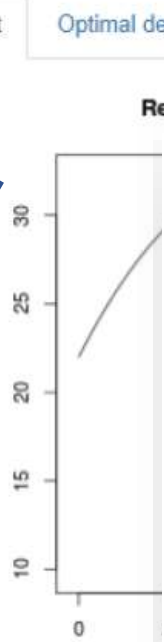
Parameter values (separated by commas)
 22,10,15,0.09

Design to compare (xi_p):

- Equally weighed in 0, 1.56, 3.12, 6.25 and 12.5 (Whole Effluent Toxicity test)
- Equally weighed in 0, 0.025, 0.05, 0.1 and 0.15 (Toxicity study of DEHP)
- Other (you must provide the support points and weights)

Support points of xi_p (separated by commas)
 0,0.025,0.05,0.1,0.15

Corresponding weights (separated by commas, total sum=1)
 0.2,0.2,0.2,0.2,0.2



If the response curves check the right end-point

Response curve for the selected model

Plot **Optimal designs**

D-optimal design for the assumed model

	Supp1	Supp2	Supp3	Supp4
doses	0.0000	3.0095	8.5565	12.5000
weights	0.2500	0.2500	0.2500	0.2500

h-optimal design for the assumed model

	Supp1	Supp2	Supp3	Supp4
doses	0.0000	2.7095	8.9174	12.5000
weights	0.3550	0.4414	0.1482	0.0554

tau-optimal design for the assumed model

	Supp1	Supp2	Supp3	Supp4
doses	0.0000	0.0953	9.4220	9.7688
weights	0.4965	0.0059	0.0836	0.4140

Efficiencies of xi_p design (with respect to the above optimal designs) + tau value

D-eff(xi_p)	h-eff(xi_p)	tau-eff(xi_p)	tau value
0.0000	0.0001	0.0000	9.7135

Área de Estadística e Investigación Operativa
 Departamento de Matemáticas. Universidad de Castilla-La Mancha

0.06.1
 Estadística e Investigación Operativa

Personal Docencia Investigación Eventos Noticias Asesoramiento Estadístico
 Incubadora de Sondeos

Seminario MUFPS: Metodologías para la estadística y experimentación. Experiencias de aula.

Posted on 26 octubre, 2018 por Sergio Pozuelo

UCLM
 UNIVERSIDAD DE CASTILLA-LA MANCHA

Optimum Experimental Design Group

Design of Experiments: Statistics for Researchers course

From the Statistics Area of the UCLM will be taught, in the coming weeks, a course of 25 teaching hours (in person and online) titled Experiment Design: Statistics for Researchers. The delivery dates are: Campus of Albacete: January 8th, 10th, 15th and 17th, 2019 from 10.30 a.m. to 1.30 p.m. Campus of Ciudad Real: November 20th, 21st, ...

Design of experiments (DOE) is a part of statistics that provides tools for efficient experimentation. Although the subject started in an agricultural context, it is nowadays being applied in many areas, both in science and in industry. A model-oriented view, where some knowledge about the form of the data-generating process is assumed, naturally leads to the so-called optimum experimental design (OED).

Due to increasing competition, DOE has become crucial for modern industry, especially for product development. Since different industrial applications of DOE (especially those in pharmaceutical industries) may exhibit very special and varying characteristics, this model oriented approach with its tailor-made solutions is of advantage.

Research Group

Just few scientists in Spain are focusing on the optimum experimental design, it could be fairly stated that the

Optimum experimental design group

THE WORLD OF STATISTICS

Contact

Contact with the webmaster via email.

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

