Principles of Design of Experiments:

1. Experiments are opportunities to ask nature how she/he works. Nature does not like to answer a lot of questions and frequently gives short answers. Once in a while nature rambles and provides answers to questions not yet asked.

2. Nature responds over a noisy line.

2. Listen to natures answers carefully

3. Before asking the questions, the experimenter should know wha to do with the answers. Different experimenters will come to different conclusions.

Game of 20 questions:

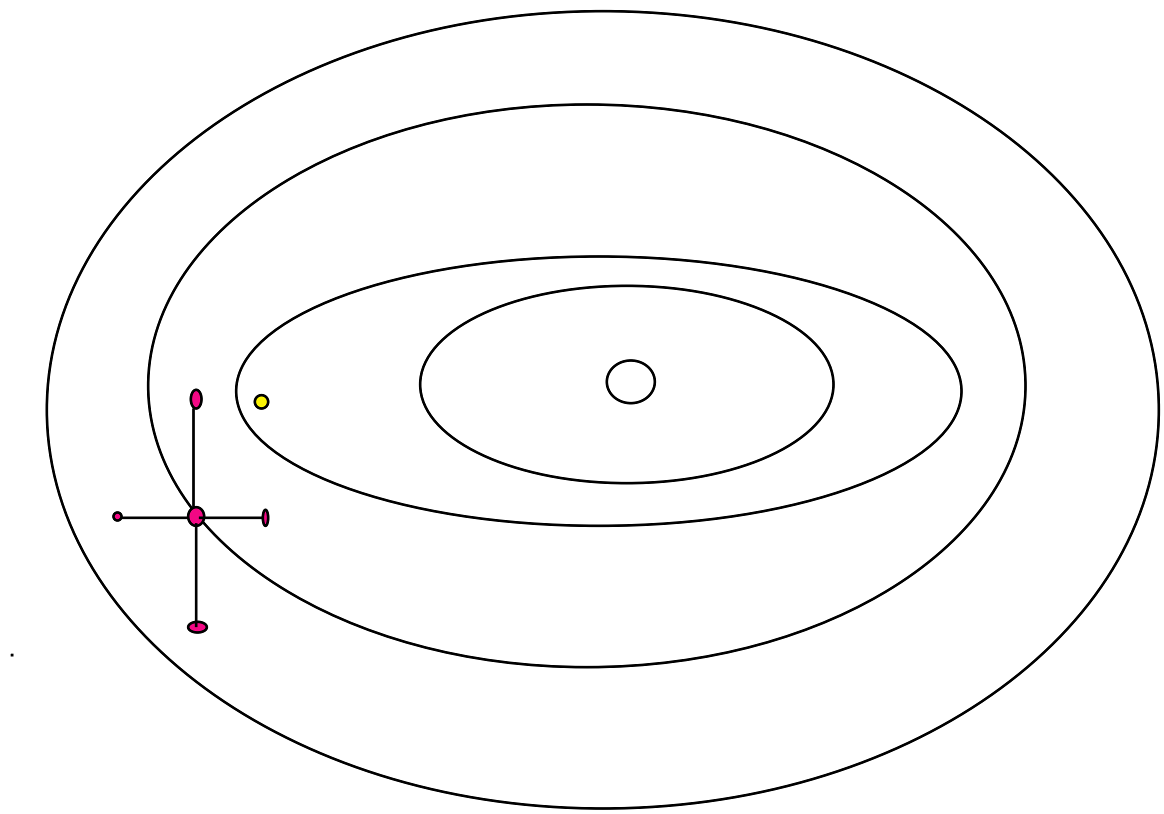
Statistical analysis serves as a noise filter. The very best that it can do is to reduce the noise and interpret the processed answers to the questions that were asked.

One factor at a time experiments:

Habit and tradition.

Continuous processes respond immediately.

Fear of making too big of a change at once.



By changing more than one factor at a time interaction effects can b
estimated.

Interaction effects are equivalent to cross partial derivatives.

Factorial experiments:

Usually each factor is tested at k levels. Often k = 2. If there are p factors then a 2^p facorial designs considers each possible combination of factors.

Such an experiment allows estimation of all possible interactions of an additive model.

$$Y = f(x_1, \quad , x_p) + \varepsilon$$

$$f(x_1, \quad , x_p) = f(x_{01}, \quad , x_{0p}) +$$

$$[(x_1 - x_{01}, \quad , x_p - x_{0p})] \ \ f(x_{01}, \quad , x_{0p})$$

$$+[(x_1 - x_{01}, \quad , x_p - x_{0p})] \ {}^2[(x_1 - x_{01}, \quad , x_p - x_{0p})]'$$

$$+$$

Then f can be decomposed as follows:

$$f(x_1, \quad , x_p) = f(x_{01}, \quad , x_{0p}) + [(x_1 - x_{01}, \quad , x_p - x_{0p})]'\alpha$$

$$[(x_1 - x_{01}, \quad , x_p - x_{0p})] \begin{matrix} \beta_{11} & & \beta_{1p} \\ \vdots & & \vdots \\ \beta_{1p} & & \beta_{pp} \end{matrix} (x_{01}, \quad , x_{0p})$$

+higher order terms.

Fractional factorial designs:

Uses fewer obsevations and not all interactions can be estimated.

Methods of comparing regression experiments:

Let $D(E) = (X^t \Sigma X)^{-1}$

Experiment E_1 is preferred to experiment E_2 if:

1. $D(E_1) < D(E_2)$

2. $|D(E_1)| < |D(E_2)|$

3. $\text{TRACE } D(E_1) < \text{TRACE } D(E_2)$

4. $\text{Max } D_{\alpha\alpha}(E_1) < \text{Max } D_{\alpha\alpha}(E_2)$

5. $L^t D(E_1) L < L^t D(E_2) L$ (L may be a matrix).

6. $\underset{\text{design space}}{\text{Max}} \; x^t D(E_1) x < \underset{\text{design space}}{\text{Max}} \; x^t D(E_2) x$

7. $\oint_{\text{design space}} x(\omega)^t D(E_1) x(\omega) \, d\omega < \oint_{\text{design space}} x(\omega)^t D(E_2) x(\omega) \, d\omega$

Picture:

Let $d(\omega, E) = x(\omega)^t D(E) x(\omega)$.

Concept of a continuous design.

Take $\sum = I$ (for now).

$$(X^tX) = \sum_{i=1}^{n} n_i \; x(\omega_i)x(\omega_i)^t$$

where $\sum_{i=1}^{n} n_i = n$.

We use the approximation $\sum_{i=1}^{n} \frac{n_i}{n} \; x(\omega_i)x(\omega_i)^t \approx$

$$\oint_{\Omega} \xi(\omega) \; x(\omega)x(\omega)^t \; d\omega \,. = M(\xi)$$

It is easier to work with continuous designs.

Caratheodory's Theorem:

Each point s* in the convex hull S* of any subset S, of R^n, can be represented in the form

$$s^* = \sum_{i=1}^{n+1} \alpha_i s_i \ ,$$

where $,\alpha_i > 0, \ \sum_{i=1}^{n+1} \alpha_i = 1$, and s_i is in S_i.

If s* is a boundary point then α_{n+1} can be set to zero.

Theorem:

1.For any design E the matrix $M(\xi)$ is a symmetric positive semi-definite matrix.

2.The matrix $M(\xi)$ is degenerate ($|M(\xi)| = 0$) if the measure ξ contains less than m points. (The number of parameters = m)

3.The family $M(\xi)$ of all possible normalized designs is convex. If the functions $x(\omega)$ are continuous and W is compact then the family $M(\xi)$ is compact.

4. For any design matrix $M(\xi)$ there is a representation

$$\sum_{i=1}^{n} p_i \ x(\omega_i) x(\omega_i)^t$$

where $n \le [(m+1)m/2]$ the $0 \le p_i \le 1$ and $\sum_{i=1}^{n} p_i = 1$.

Equivalence of design criteria:

Lemma:

1. $\sum_{i=1}^{n} p_i \; d(\omega_i, E) = m$ (m = the number of columns of X^tX)

2. $\underset{\omega}{\text{Max}} \; d(\omega,E)) \ge m$.

Define convex function.

Lemma: The function log $|M(\xi)|$ is a strictly convex function.

Big Theorem(Keifer & Wolfowitz(1960)):

The following assertions are equivalent:

1. The design ξ^* maximizes $|M(\xi)|$.

2. The design ξ^* minimizes $\underset{\omega}{\text{Max}}\ d(\omega, \xi))$.

3. $\underset{\omega}{\text{Max}}\ d(\omega, \xi) = m$. Any linear combination of designs satisfying these criteria all satisfy them.

DETMAX

Suppose that the design space Ω consists of a finite number N of elements and we wish to choose only n of them (replicates are possible) how should we proceed?

We can choose $\binom{N^n}{n}$ possible designs and compare the resulting M matrices. Such an exhaustive search is not practical. Instead we use an algorithm. Recall that the function log $|M(\xi)|$ is a strictly convex function and that convex hull of the design space is convex. So we have a convex programming problem on the space of probability measures ξ.

How to find the optimal continuous design.

1. For a full rank design (measure) ξ_0 calculate $M(\xi_0)$.

2. Find the point in ω_0 in Ω such that $d(\omega_0, \xi) = \underset{\omega}{Max}\ d(\omega, \xi)$.

3.The design $\xi_1 = (1-\alpha_0)\ \omega_0 + \alpha_0\ \xi_0$ is constructed where

$$\alpha_0 = \frac{d(\omega_0, \xi)}{[d(\omega_0, \xi) + (m-1)]m}.$$

4.The matrix $M(\xi_1)$ is constructed and steps 2-4 are repeated until convergence.

Short cut formulas: Each new update does not have to go through a complete recalculation.

The finite design space problem.