EXPERIMENTS FOR DETERMINING

NON-ISOTHERMAL KINETIC RATES

Anthony Atkinson, 30th April 2018

Joint work with Belmiro Duarte, Coimbra

Department of Statistics
London School of Economics
London WC2A 2AE, UK

a.c.atkinson@lse.ac.uk

History and Motivation

- The optimal design of experiments for the nonlinear models arising in chemical kinetics was introduced by Box and Lucas (1959).
- Box had been working in the dyestuffs division of ICI at Blakeley near Manchester.
- At that time dyestuff manufacture was a batch process: mix the ingredients, stir, heat, wait and see what transpires.
- Accordingly, the designs in Box and Lucas specify a set of conditions (sometimes temperature) and a time at which a single observation is taken. This paradigm is firmly entrenched in the statistical literature on optimum design.
- However, in many industrial experiments it is possible to take a series of non-intrusive readings as the reaction proceeds.
- What are the comparative properties of such designs?

Structure

- The kinetic model
- Locally D-optimum design (4 batches)
- Two batches and many observations
- An extended equivalence theorem
- One batch and varying temperature
- Conclusions
- Throughout the focus is on D-optimum designs

Two Consecutive First-Order Reactions

- The reaction scheme is

$$A \to B \to C.$$

  with rates $k_1(T)$ and $k_2(T)$
- With first-order kinetics the concentrations $[A], [B]$ and $[C]$ are given by

$$\frac{\mathrm{d}[A]}{\mathrm{d}t} = -k_1(T)[A], \quad (1)$$

$$\frac{\mathrm{d}[B]}{\mathrm{d}t} = k_1(T)[A] - k_2(T)[B], \quad (2)$$

$$\frac{\mathrm{d}[C]}{\mathrm{d}t} = k_2(T)[B]. \quad (3)$$

  At $t = 0, [A] = 1, [B] = [C] = 0$. Measure $[B]$
- The kinetic rates follow the Arrhenius law $k_i(T) \propto \exp(E_{a,i}/T)$, with T the absolute temperature. This may be written

$$k_i(T) = \theta_{i,1} \exp\left[-\theta_{i,2}\left(T_0/T - 1\right)\right], \quad i \in \{1, 2\}.$$

- The higher the temperature, the faster the reaction.

The Concentration of B

- If T is held constant, the solution for the concentration of B is

$$\mathbb{E}[B(t)] = \frac{k_1(T)}{k_1(T) - k_2(T)} \left[\exp(-k_2(T)t) - \exp(-k_1(T)t)\right] \qquad (4)$$

- Rises from zero to a maximum and then gradually declines to zero.
- Find locally D-optimum designs by measuring $[B(t)]$ assuming i.i.d. errors.
- In (4)

$$k_i(T) = \theta_{i,1} \exp\left[-\theta_{i,2}\left(T_0/T - 1\right)\right], \quad i \in \{1, 2\}.$$

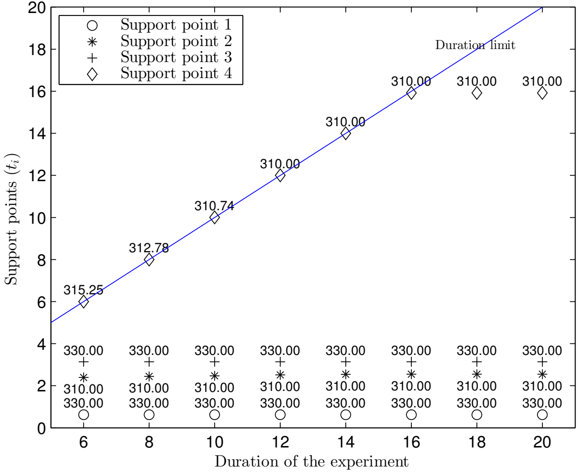
  Take $\theta_{1,1} = 0.7, \theta_{1,2} = 21.875, \theta_{2,1} = 0.2, \theta_{2,2} = 28.175$ with $T_0 = 320$.
- Design Region $t \in H : T \in Z$. Throughout $Z = [310, 330]$. Look at effect of H on the three strategies.

The "Box-Lucas" Design

- One reading per batch. $Z = [310, 330]$ and $H = [0, 20]$.
- The Box-Lucas optimum design is

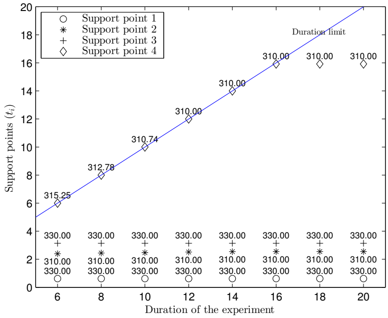
$$\xi_D^{\text{ref}} = \begin{pmatrix} 0.618 & 3.137 & 2.544 & 15.921 \\ 330.00 & 330.00 & 310.00 & 310.00 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{pmatrix}.$$

- Design spans Z, but not H. Readings on two batches at $T = 330$ at relatively short times and two at $T = 310$ at longer times.
- "Reference Design" for exploring effect of changing H.
- Look at $Z = [0, t^{\text{UP}}]$ with $t^{\text{UP}} = \{18, 16, 14, 12, 10, 8, 6\}$.



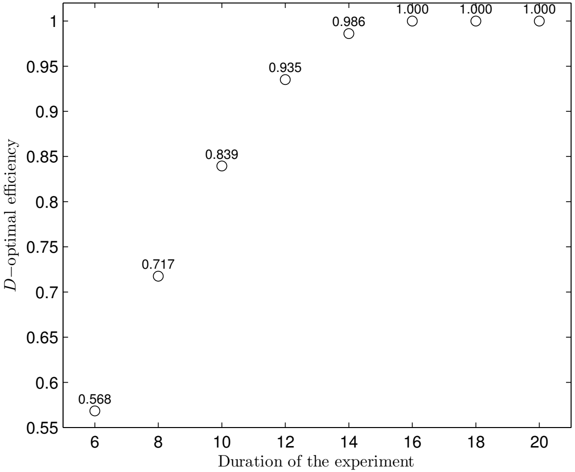
Support points of designs with four batches as t^{UP} increases. The values close to the support points are the respective temperatures

The Effect of Experimental Duration



- The pairs (t_i, T_i) of the first three support points are very close for all t^{UP}.
- $t_4 < t^{\text{UP}}$ when $t^{\text{UP}} \geq 16$ and equals $t^{\text{UP}}, t^{\text{UP}} < 16$.
- For $t^{\text{UP}} \leq 12$ the temperature of the last observation is not the minimum and starts increasing as $\mathbf{H}$ becomes smaller. All other temperatures are unchanged.

Design Efficiency and Duration



D-efficiency of designs for four batches for different interval durations

H

Design Optimality

- Are these designs D-optimum?
- Theory for linear models

$$y_i = \theta^T f(x_i) + \epsilon_i.$$

The parameter vector θ is $p \times 1$, with $f(x_i)$ a known function of the explanatory variables x_i.

- For nonlinear models expand in a Taylor series around some value θ_0
- The information matrix for the design ξ with n support points is

$$M(\xi) = \sum_{i=1}^{n} w_i f(x_i) f(x_i)^T = F^T W F, \tag{5}$$

where F is the $n \times p$ extended design matrix, with ith row $f^T(x_i)$ and W is a diagonal matrix of weights.

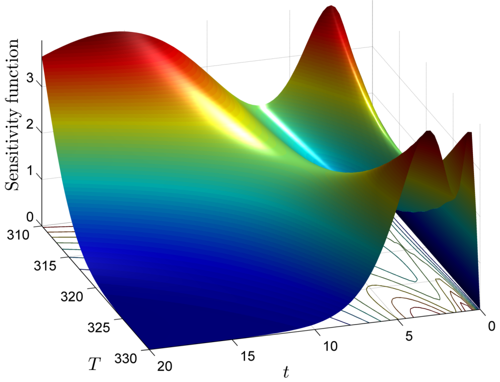
D-optimality

- D-optimum designs, minimizing the generalized variance of the estimates of θ, maximize the determinant $|F^T WF|$ over the design region $\mathcal{X}$ through choice of the optimum design ξ^*. In the locally optimum design above $w_i = 0.25, (i = 1, \ldots, 4)$.
- That this design is D-optimum can be shown by use of the "general equivalence theorem" for D-optimality (Kiefer and Wolfowitz, 1960) which provides conditions for the optimality of a design ξ which depend on the sensitivity function

$$d(x, \xi) = f^T(x) M^{-1}(\xi) f(x). \tag{6}$$

For the optimum design, $\bar{d}(x, \xi^*)$ the maximum value of the sensitivity function over $\mathcal{X}$, equals p, the number of parameters in the linear predictor. These values occur at the points x_i of support of the design.

Sensitivity Function



- Sensitivity function of the D-optimum design with four batches obtained for $\mathbf{H} = [0, 20]$ and $\mathbf{Z} = [310.0, 330.0]$.
- The maxima have a value of 4.

Two Batches and Many Observations

- Two batches. Measure $[B(t)]$ frequently; measurements are cheap and non-intrusive.
- Measurements are taken with a constant frequency at a set of pre-defined time instants, $t_i,\ i \in \{1, \cdots, N\}$. The interval between consecutive measurements is $\Delta t = t^{\text{UP}}/(N-1)$.
- Design problem. Choose temperatures T_1 and T_2.
- Same grid of observational values for the two temperatures. (The grid for the upper temperature will be too coarse and that for the lower too fine).

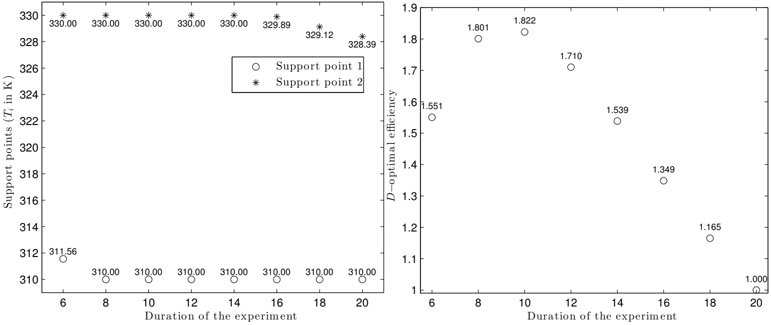
Optimum Two Batch Design

- In all numerical tests $N = 21$, although the duration of the experiment depends on t^{UP}.
- The time points were equally spaced $\Delta t = t^{\text{UP}}/(N-1)$ units distant from each other.
- For $\mathbf{H} = [0, 20]$ the optimum design is

$$\xi_D^{\text{batch}} = \begin{pmatrix} 0 & 1 & \cdots & 20 \\ 310.00 & 310.00 & \cdots & 310.00 \\ 330.00 & 330.00 & \cdots & 330.00 \\ 1/21 & 1/21 & \cdots & 1/21 \end{pmatrix}.$$

- The optimum temperatures for the experiments coincide with the lower and upper bounds.
- The efficiency of ξ_D^{batch} relative to ξ_D^{ref} for the four batch design is 0.1130.
- These are efficiencies calculated from measures ξ. The two batch experiment requires about 10 times more observations than that for four batches (surprisingly little difference).

Design Efficiency and Duration



- Effect of changing $\mathbf{H}$: optimum designs and efficiencies relative to ξ_D^{batch}.
- Except for extreme $\mathbf{H}$, temperatures are 310 and 330.
- For $\mathbf{H} = [0, 6]$, $T = 311$ provides greater information about k_2 at the end of the trial.

Extended Equivalence Theorem

- For given $\mathbf{H}$ the design region is $\mathbf{Z} = [310, 330]$
- At the point $i \in \mathcal{X}$ obtain $N = 21$ readings
- Let S_i denote the ith set of observations, taken at times $t_{i1}, t_{i2}, \ldots, t_{iN}$ and let

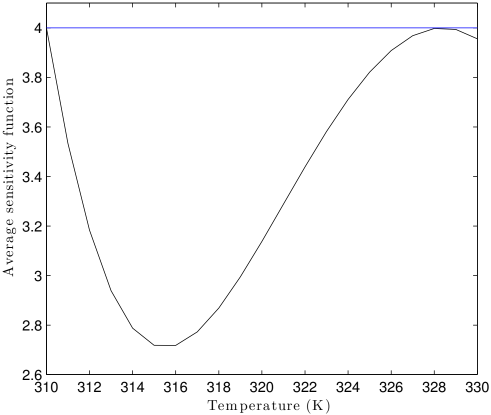
$$d_{\mathsf{AVE}}(i, \xi) = \sum_{j \in S_i} d(t_{ij}, \xi)/N.$$

  Further, let $\bar{d}_{\mathsf{AVE}}(\xi)$ be the maximum over $\mathcal{X}$ of $d_{\mathsf{AVE}}(i, \xi)$.
- The **Equivalence Theorem** states the equivalence of the following three conditions on ξ^*:
  1. The design ξ^* maximizes $|M(\xi)|$;
  2. The design ξ^* minimizes $\bar{d}_{\mathsf{AVE}}(\xi)$;
  3. The value of $\bar{d}_{\mathsf{AVE}}(\xi^*) = p$, this maximum occurring at the points of support of the design.

  As a consequence of 3, we obtain the further condition:

  4. For any non-optimum design the value of $\bar{d}_{\mathsf{AVE}}(\xi) > p$.
- See Atkinson (2016)

Sensitivity Function



- Average of the sensitivity function for the two-batch D-optimum design ξ_D^{batch} ($N = 21$, $\mathbf{H} = [0, 20]$ and $\mathbf{Z} = [310, 330]$).
- The design is D-optimum ($p = 4$).
- The figure confirms that the upper temperature in the optimum design is < 330K.

One Batch with Varying Temperature

- All information is extracted from a single batch experiment of length $\mathbf{H} = [0, t^{\mathsf{UP}}]$.
- The temperature can be manipulated at regular time points and is constant between changes.
- We assume that the changes in temperature for implementing the optimal profile are instantaneous. Contrarily, the measured response $[B(t)]$ varies continuously and its dynamics depend upon the time as well as on current values of control and state variables.
- Because of the changes in temperature, a system of ODEs is required to be integrated numerically together with the equations for the parameter sensitivities.
- An example of classical optimal control. It would be more realistic to use B-splines (Uciński and Bogacka, 2004), rather than a piecewise linear temperature profile

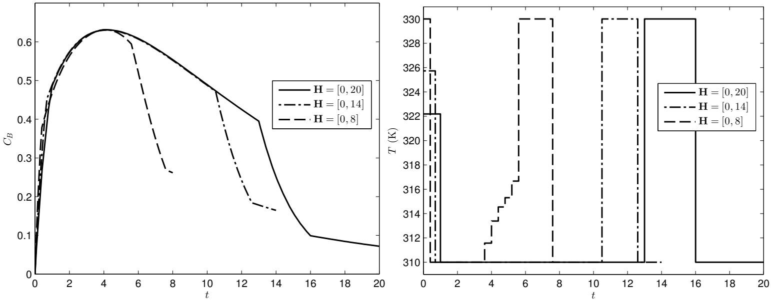
One Batch, Varying Temperature, Coarse Grid

The optimum design for the coarse grid $\mathbf{H} = [0, 20]$

$$\xi_D^{\text{varT}} = \left(\begin{array}{ccccccc} 0 & 1 & 2 & 3 & 4 & 5 & 6 \\ 322.18 & 310.00 & 310.00 & 310.00 & 310.00 & 310.00 & 310.00 \\ 1/21 & 1/21 & 1/21 & 1/21 & 1/21 & 1/21 & 1/21 \\ \hline 7 & 8 & 9 & 10 & 11 & 12 & 13 \\ 310.00 & 310.00 & 310.00 & 310.00 & 310.00 & 310.00 & 330.00 \\ 1/21 & 1/21 & 1/21 & 1/21 & 1/21 & 1/21 & 1/21 \\ \hline 14 & 15 & 16 & 17 & 19 & 20 & \\ 330.00 & 330.00 & 310.00 & 310.00 & 310.00 & 310.00 & \\ 1/21 & 1/21 & 1/21 & 1/21 & 1/21 & 1/21 & \end{array} \right).$$

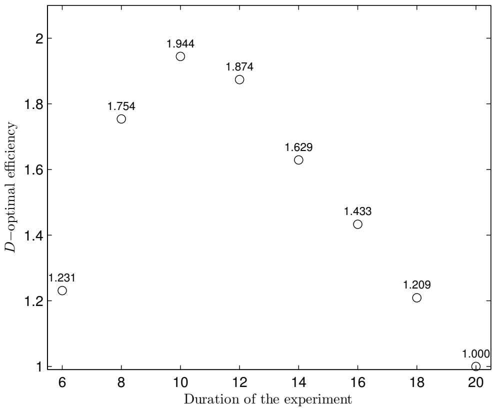
- Hi, lo, hi, lo: 322, 310, 330, 310.
- Not quite “bang-bang” control (not 330 initially).

Concentration and Temperature Profiles



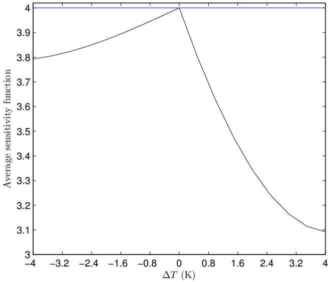
- Dynamic D-optimum designs for different $\mathbf{H}$ by changing t^{UP} ($N = 21$, $\mathbf{Z} = [310, 330]$)
- (a) Concentration $[B(t)]$ and (b) temperature profiles.
- All temperature profiles have a low final period.
- The high temperature about 2/3 of the way through reduces $[A]$ and provides information on the second reaction B $\rightarrow C$

Efficiencies



- D-optimum efficiency of the dynamic designs for different $\mathbf{H}$ with the reference being the design ξ_D^{varT}.

Sensitivity Function



- Analyze the optimality of this design with the extended equivalence theorem.
- The set of optimum temperatures $T_i,\ i \in [N]$ is $\mathbb{T}^{\text{opt}}$. Each $T_i \in \mathbf{Z}$, so the design region has dimension $N = 21$.
- We look at the subset of profiles, for the same grid of times, from increasing the observations in $\mathbb{T}^{\text{opt}}$ by an amount ΔT providing they belong to $\mathbf{Z}$.
- Plot shows average sensitivity function for ξ_D^{varT} ($N = 21$, $\mathbf{H} = [0, 20]$ and $\Delta T \in [-4, 4]$).

Conclusion 1

Table: Relative D-optimum efficiency for four batch designs under the three strategies for $\mathbf{H} = [0, t^{\text{UP}}]$.

Rule	Min. No. Batches	t^{UP} 6	8	10	12	14	16	18	20
Locally optimum	4	2.27	2.87	3.36	3.74	3.94	4.00	4.00	4.00
Two batches	2	7.36	8.55	8.65	8.12	7.30	6.40	5.53	4.75
Temp. profile	1	4.76	6.78	7.51	7.24	6.29	5.54	4.67	3.86

- The table gives a comparison of the relative efficiencies of the three designs for experiments with four batches (the minimum for the locally optimum design).
- Despite the second and third designs producing 21 measurements per batch, rather than one, at their most efficient, that is for $t^{\text{UP}} = 10$, these designs are only about twice as informative as the locally optimal design. If measurement were the major cost, an unlikely scenario, the locally optimal design would be preferred.

Conclusion 2

Table: Relative D-efficiency for four batch designs under the three strategies for $\mathbf{H} = [0, t^{\text{UP}}]$.

	Min. No.	t^{UP}							
Rule	Batches	6	8	10	12	14	16	18	20
Locally optimal	4	2.27	2.87	3.36	3.74	3.94	4.00	4.00	4.00
Two batches	2	7.36	8.55	8.65	8.12	7.30	6.40	5.53	4.75
Temp. profile	1	4.76	6.78	7.51	7.24	6.29	5.54	4.67	3.86

- Two disadvantages of the locally optimum design are that it requires more batches and takes longer ($t^{\text{UP}} = 16$, rather than 10). A further disadvantage is that the efficiency of the locally optimum design depends upon the parameter values being close to the value θ_0 used to find the optimum design. With a series of readings, dependence of efficiency on the prior value is reduced.
- An advantage of the second and third designs is the reduced number of batches required. If sufficient accuracy could be obtained from readings on a single batch, then the experiment with varying temperature profile would be preferred. However, the design with two batches is simpler to run.

Conclusion 3

- There are many other reasonable measurement strategies
- Experiments with two batches could have different sets of measurement points at the upper and lower temperatures.
- Sampling could be more frequent at the beginning of each run of a batch.
- We have focused on the two strategies that appear most frequently in the literature.
- Chapter 7 of Fedorov and Leonov (2014) describes cost optimal designs for repeated dose scheduling in clinical trials.

References

Atkinson, A. C. (2016). Optimum experiments with sets of treatment combinations. In C. H. Müller, J. Kunert, and A. C. Atkinson, editors, *mODa 11 – Advances in Model-Oriented Design and Analysis*, pages 19–26. Springer, Heidelberg.

Box, G. E. P. and Lucas, H. L. (1959). Design of experiments in nonlinear situations. *Biometrika*, **46**, 77–90.

Fedorov, V. V. and Leonov, S. L. (2014). *Optimal Design for Nonlinear Response Models*. Chapman and Hall/ CRC Press, Boca Raton.

Kiefer, J. and Wolfowitz, J. (1960). The equivalence of two extremum problems. *Canadian Journal of Mathematics*, **12**, 363–366.

Uciński, D. and Bogacka, B. (2004). T-optimum designs for multiresponse dynamic heteroscedastic models. In A. Di Bucchianico, H. Läuter, and H. P. Wynn, editors, *MODA 7 – Advances in Model-Oriented Design and Analysis*, pages 191–199. Physica-Verlag, Heidelberg.