Sequential Experimental Designs for Generalized Linear Models

Hovav A. Dror and David M. Steinberg, JASA (2008)

Bob A. Salim

June 4, 2013

Background (Motivation)

- Designing an experiment or a study
- Restrictions on sample size (due to cost, time, etc)
- How do we design an efficient experiment with valid estimate of the parameter?

Previous Works

- Chaudhuri and Mykland (1993): Sequential designs in GLMs could lead to fully efficient designs and asymptotically efficient MLEs
- Dixon and Mood (1948); Neyer (1994); Haines, Perevozskaya, and Rosenberger (2003); Ivanova and Wang (2004); Biedermann, Dette, and Zhu (2006); Karvanen, Vartiainen, Timofeev, and Pekola (2007): Sequential designs for binary data, focused on single-factor experiments.

Efficient Design

- Efficient experimental designs for GLMs depend on the unknown coefficients.
- Sequential design: next design point chosen based on current data.

What is Efficient?

- Optimality criterion: D-optimal
- D-optimality criterion: maximize the determinant of the information matrix $\mathbf{I}(\boldsymbol{\beta};\mathrm{d})$
  $\boldsymbol{\beta}$ = parameters in the model; d = design
- Bayesian D-optimality criterion (Chaloner and Larntz, 1989):

$$\phi(d) = \int log(\mathbf{I}(\boldsymbol{\beta}; d)) d\pi(\boldsymbol{\beta}) \tag{1}$$

where $\pi(\boldsymbol{\beta})$ = prior distribution on $\boldsymbol{\beta}$.

Overview of Proposed Method: Approximate Design Criteria

- Discretized posterior

$$\phi_1(d) = \sum_{u=1}^{n} r_u log(\mathbf{I}(\boldsymbol{\beta}_u; d)) \tag{2}$$

where $r_u = L(\boldsymbol{\beta}_u) / \sum_{u=1}^{n} L(\boldsymbol{\beta}_u)$

- A faster approximation

$$\phi_2(d) = log(\mathbf{I}(\tilde{\beta}; d)) \tag{3}$$

where $\tilde{\beta} = \beta_{(g)}$ such that $\sum_{u=1}^{g} r_{(u)} \geq 0.5$ and $\sum_{u=g}^{n} r_{(u)} \geq 0.5$.

- Another approximation for determining augmentation horizon

$$\phi_3(d) = (1/p)\phi_2(d) - log(n) \tag{4}$$

Overview of Proposed Method: Algorithm

- In the beginning of the experiment, define the augmentation horizon (m).
- Find a locally D-optimal m-run augmentation to the current design, maximizing ϕ_2 at the current parameter median.
- Generate a candidate set for the augmentation consisting of the m points found in the previous step and their coordinatewise median.
- If the design points run thus far provide a nonsingular information matrix, choose the next design point as the candidate that gives the best ϕ_1 when added to the current design.
- If the design thus far does not provide a nonsingular information matrix, then choose the next design point from among the candidates by comparing the values of ϕ_1 for designs that consist of the points run thus far, the m-run augmentation, and the candidate.

Overview of Proposed Method: Augmentation Horizon

- Proposed to avoid problems, such as singular information matrix
- It is the number of observations (m) needed for highly efficient D-optimal design at prior median
- Determined at the start of the experiment

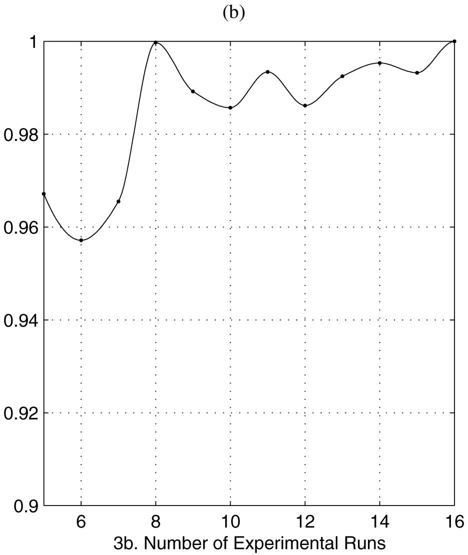
Overview of Proposed Method: Augmentation Horizon

- Find locally D-optimal designs at prior median for $n = p, ..., P$
- Define the efficiency of $d_n = \exp[\phi_3(d_n) - \phi_3(d*)]$
  where $d*$ is the design among $d_p, ..., d_P$ that maximizes $\phi_3(d)$
- Augmentation horizon m is the smallest value of n for which the efficiency is at least 99%

Finding Locally D-optimal Designs

- Given in their earlier paper (Dror and Steinberg, 2006)
- For linear regression, the information matrix is given by $\mathbf{X}^T\mathbf{X}$
- For GLMs, the information matrix is given by $\mathbf{X}^T\mathbf{W}\mathbf{X}$, where $W = V^{-1}(\boldsymbol{\mu})(d\boldsymbol{\mu}/d\boldsymbol{\eta})^2$. $\boldsymbol{\mu}$ is the vector of expected response, and $\boldsymbol{\eta}$ is the linear predictor $\mathbf{X}\beta$.
- For example, for logistic regression, $\mathbf{W}$ is a diagonal matrix with diagonal elements $w_{ii} = exp(\mathbf{X}_i\beta)/(1 + exp(\mathbf{X}_i\beta))^2$.
- Applies a row-exchange algorithm (Federov, 1972)

Back to the Augmentation Horizon



Method Comparison: Bruceton

Algorithm:

- Described in Dixon and Mood (1948)
- Determine starting value and step size (s)
- If current design point (x) produces positive outcome, the next design point is chosen to be the design point one step size below the current design point ($x - s$)
- If current design point (x) produces negative outcome, the next design point is chosen to be the design point one step size above the current design point ($x + s$)

Method Comparison: Neyer

Overview

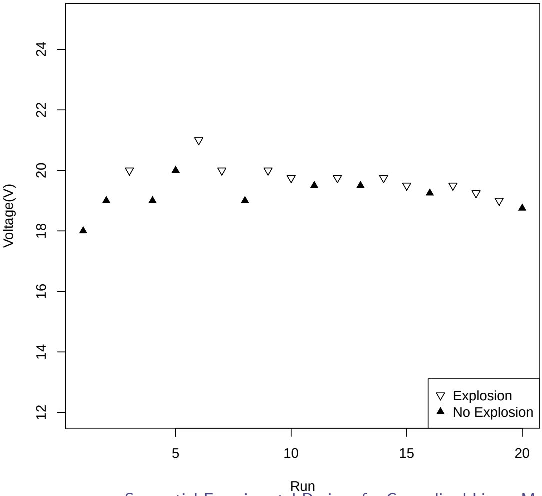
- Described in Neyer (1994)
- Three-part procedure
- First part: "Close in" on the region of interest
- Second part: Determine unique estimates of the parameters
- Third part: Use local D-optimal design based on the MLE of the parameters

Explosives Testing Example

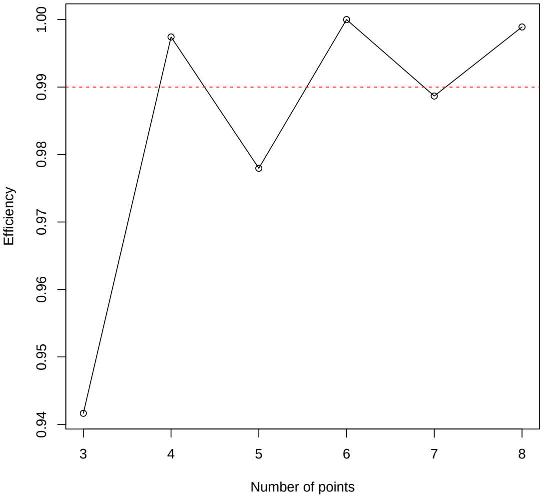
- Sensitivity Experiment in June 2006 at an industrial plant.
- Compare the performance of this method to Bruceton up-and-down method (default)
- Requirement: probability of detonation at 12V (or below) is $< 5\%$ and the probability of detonation at 25V (or above) is $> 95\%$.
- Parameterization: $P(y = 1|...) = F(\frac{x-\mu}{\sigma})$, where F is the inverse of the link function. For example, in doing logistic regression, F is the *expit* function.
- The authors used the following priors: $\mu \sim$ lognormal(log(17),0.5^2) and $\sigma \sim$ lognormal(log(0.7),0.5^2)
- True parameters: $\mu = 19, \sigma = 0.7$

Explosives Testing Example: Bruceton

Simulation study. True parameter: $\mu = 19$ and $\sigma = 0.7$

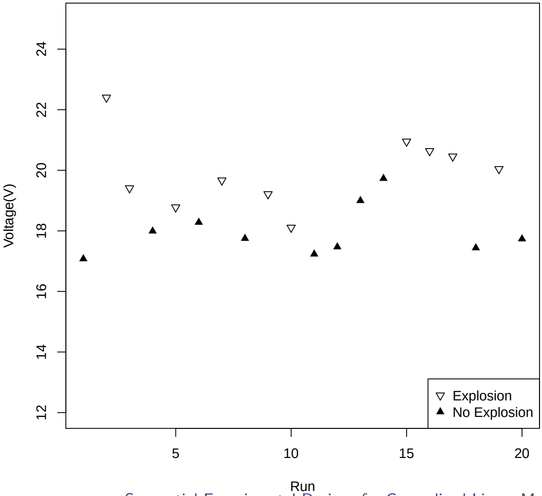


Augmentation Horizon for Sensitivity Test



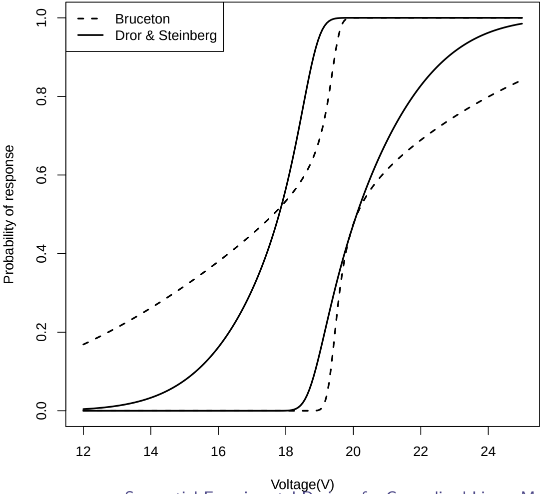
Explosives Testing Example: Dror & Steinberg

Simulation study. True parameter: $\mu = 19$ and $\sigma = 0.7$



Explosives Testing Example: Comparison of results

Simulation study. True parameter: $\mu = 19$ and $\sigma = 0.7$



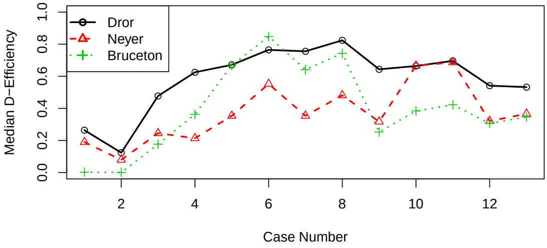
More Simulations: Efficiency

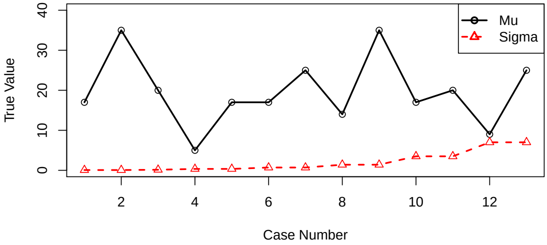
- Bruceton: Initial point = 17, step size = 1.28
- Neyer: $\mu_{min} = 12$, $\mu_{max} = 25$, $\sigma_{guess} = 0.7$
- Dror: $\mu \sim$ lognormal(log(17),0.5^2) and $\sigma \sim$ lognormal(log(0.7),0.5^2)
- Efficency = $\exp(\phi_3(D))/\exp(\phi_3(Dopt))$
  where D is the obtained design and Dopt is the optimal design points given that we know the true parameter values.

More Simulations: Median D-Efficiency

Case	True μ	True σ	Bruceton	Neyer	Dror
1	17.00	0.07	0.002	0.19	0.26
2	35.00	0.07	0.0004	0.08	0.12
3	20.00	0.14	0.18	0.25	0.48
4	5.00	0.35	0.36	0.21	0.62
5	17.00	0.35	0.67	0.35	0.67
6	17.00	0.70	0.85	0.55	0.76
7	25.00	0.70	0.64	0.35	0.76
8	14.00	1.40	0.74	0.48	0.82
9	35.00	1.40	0.25	0.32	0.64
10	17.00	3.50	0.38	0.67	0.66
11	20.00	3.50	0.42	0.69	0.70
12	9.00	7.00	0.31	0.32	0.54
13	25.00	7.00	0.35	0.37	0.53

More Simulations: Median D-Efficiency





Summary

Advantages:

- Allows for design of multifactor experiments
- Allows for design of experiments for GLMs (not just binary data)
- Less singularity problem, or cases where MLEs don't exist
- Uses D-optimality criterion from the beginning of the experiment.

Disadvantages:

- A bit more complicated to do, more computationally heavy
- Augmentation horizon sensitive to parameters

Thank You!

- Jon and Patrick
- Everyone in the class

References

Dror, H.A. and Steinberg, D.M. (2008). Sequential Experimental Designs for Generalized Linear Models. Journal of American Statistical Association 103:481, 288-298

Abdelbasit, K.M. and Plackett, R.L. (1983). Experimental Designs for Binary Data. Journal of American Statistical Association 101, 747-759

Chaloner, K. and Larntz, K. (1989). Optimal Bayesian Design Applied to Logistic Regression Experiments. Journal of Statistical Planning and Inference, 21, 191-208

Dixon, J.W. and Mood, A.M. (1948) A Method for Obtaining and Analyzing Sensitivity Data. *Journal of American Statistical Association* 43, 109-126

Robbins, H. and Monro, S. (1951). A Stochastic Approximation Method. Annals of Mathematical Statistics 29, 400-407