

Optimal Experimental Design: Psychophysics of change point Detection.

Manuel Villarreal and Arturo Bouzas

Mexico City, Mexico

Abstract

Keywords: Optimal Experimental Design, Psychophysics, Change Point Detection

1. Introduction

One of the main challenges in scientific research is the design of an experiment. A good experimental design will make the difference between finding an answer to our research question and wasting valuable resources like time and money.

When designing an experiment, often one starts by making decisions about the number of participants, how many and what values of our independent variable we should test or how many times should we test those values. Most of the time, these choices are made on the basis of previous research on the field. However, it might be the case that there is not enough information to make these decisions with confidence, or that the values that are commonly used, do not allow for strong conclusions. Optimal Experimental Design (OED) offers an alternative to solve this kind of problems through the formalization of the design problem.

15 OED allows us to re-interpret the problem of designing an experiment as
16 a decision problem, in which the purpose is to maximize a utility function.
17 This function is a numeric representation of our preference over the possible
18 consequences of running an experiment. Therefore, the optimization of an
19 experimental design requires us to have a formal interpretation of the purpose
20 of the experiment.

21 The concept of optimizing an experimental design is not new to psycho-
22 logical research. There are already examples of design optimization in the
23 literature (e.g. Myung and Pitt, 2009; Zhang and Lee, 2010). Both of this
24 examples discuss and demonstrate the advantages of OED for model com-
25 parison in psychology.

26 In this paper, we will present a different approach where the problem is
27 not to select between two models but to make inferences about the parameters
28 of a single one.

29 In particular, we will present an example of OED in the context of Gen-
30 eralized Linear Models. These models are widely used in psychophysics. The
31 problem that will be treated here is fairly new, however, the methodologi-
32 cal aspects remain the same even with more straight-forward psychophysical
33 experiments. In order to do this, we will use a particular parametrization of
34 the logistic model which is primarily used in statistics.

35 **2. Optimal experimental design**

36 To apply the concepts of OED to a particular problem, first, we need
37 to define the elements of the design space. This space is conformed by the
38 variables that we can manipulate during the experiment, for example, the

39 values that our independent variable might take or the weight (proportion of
 40 observations) assigned to each of those values. These elements are the ones
 41 that we can modify in order to optimize the design.

42 The second step would be to formalise the objective of the experiment.
 43 For example, in the case of a logistic model, we might want to find the
 44 values of our independent variable that minimize the variance of the model
 45 parameters, or we might be interested in the magnitude of the physical stim-
 46 ulus for which the probability of a response takes on a certain value. The
 47 formalization of the research question will define a utility function.

48 The last step is to specify the prior information that we have about the
 49 problem at hand. This last step can be carried out in two ways, first, we can
 50 try to optimize the experiment for a particular guess about the parameter
 51 values of the model of interest, or we could use a probability distribution
 52 to account for the uncertainty in the values that we are interested in. This
 53 last step is primarily important to the optimization process for generalized
 54 linear models, because the optimal design will depend on the values of the
 55 parameters.

56 Once we have defined the design space η , the objective of the experiment
 57 and the prior information $p(\theta)$, the expected utility of a design is represented
 58 by the following equation:

$$U(\eta) = \int \int U(\eta, y, \theta) p(\theta|y, \eta) p(y|\eta) d\theta dy \quad (1)$$

59 Therefore, finding the experimental design that is optimal given a utility
 60 function, reduces to the problem of finding the values for the variables in η
 61 for which equation 1 takes it's maximum value.

62 Many utility functions have been proposed for both linear and non-linear

design problems, however, as was previously mentioned, the choice of a utility function depends on the research question. For example, Myung and Pitt (2009) utilizes the sum of squared errors between the data generated by a model and the predictions of a competing one given an experimental design. This is because, the objective is to find a design that can discriminate the predictions of the two. On the other hand, Zhang and Lee (2010) choose a utility function based on the Bayes Factor. Both of these functions will have their advantages, nonetheless, the first utility emphasizes discriminability of model's predictions, while the second one aims for a design in which the data generating model is more likely.

When the objective of the experiment is to make an inference about the parameters of a generalized linear model, some authors (e.g. Bernardo, 1979) have proposed to consider the gain in Shannon's Information as a utility function. In this case, the objective is to find the design η that maximizes the expected gain in Shannon Information, or equivalently, maximizes the Kullback-Leibler divergence between the posterior and prior distribution. With this function, the expected utility of an experimental design is represented by the following equation:

$$U(\eta) = \int \int \log \frac{p(\theta|y, \eta)}{p(\theta)} p(y, \theta|\eta) d\theta dy \quad (2)$$

The prior distribution in the denominator of the logarithm can be dropped as it does not depend on the experimental design, therefore, the optimal design will be the one that maximizes:

$$U(\eta) = \int \int \log \{p(\theta|y, \eta)\} p(y, \theta|\eta) d\theta dy \quad (3)$$

Which is the posterior expected Shannon information.

85 When dealing with generalized linear models, the posterior distribution
 86 of the parameter vector θ is not always tractable, however, in the literature
 87 of design optimization it is common to use the following approximation to the
 88 posterior distribution

$$\theta|y, \eta \sim N\left(\hat{\theta}, [nI(\hat{\theta}, \eta)]^{-1}\right) \quad (4)$$

89 Where $I(\hat{\theta}, \eta)$ denotes the observed fisher information matrix given an ex-
 90 perimental design and $\hat{\theta}$ is the maximum likelihood estimate of θ . Even with
 91 this, the marginal distribution of the data ($p(y|\eta)$) in equation 1 also needs
 92 to be approximated, however, when the posterior utility only depends on y
 93 through some consistent estimate of $\hat{\theta}$, a further approximation is to take
 94 the predictive distribution of $\hat{\theta}$ to be the prior distribution Chaloner and
 95 Larntz (1989).

96 With both approximations, the value of $U(\eta)$ is:

$$U(\eta) = -\frac{k}{2}\log(2\pi) - \frac{k}{2} + \frac{1}{2} \int \log\{\det[nI(\theta, \eta)]\}p(\theta)d\theta \quad (5)$$

97 Equation 5 gives the exact expected utility of an experimental design. How-
 98 ever, one could drop the constant and multiplier terms giving the following
 99 form:

$$\phi(\eta) = \int \log\{\det[nI(\theta, \eta)]\}p(\theta)d\theta \quad (6)$$

100 The function $\phi(\eta)$ is known as a design criterion. An optimal design would be
 101 the one maximizing equation 6, nonetheless, it is worth noting that this relies
 102 on the normal approximation to the posterior distribution in 4, therefore, for
 103 small samples there are some constraints that help to assure normality (see
 104 Clyde and Chaloner (2002)).

105 3. OUTLINE

106 4. Optimal Experimental Design: Example

107 Why is detecting changes important for an organism?

108 Change detection in probabilistic series.

109 Arising problems with experimental design.

110 Research question and its statistical interpretation

111 Assumption about the relationship between a subjects response the depen-
112 dent variable under study

113 Design space for this problem and how to reduce the dimensionality of
114 the space by assuming experimental constraints.

115 Utility function and its relationship with the objective of the experiment

116 Arising problems with utility function and the proposed response func-
117 tion. Bayesian solution, assigning a prior distribution to the parameters, the
118 less research in a field the more difficult it is to assign an informative prior,
119 however, we could use other cognitive models in order to propose a prior
120 distribution.

121 4.1. Using a model to generate prior distributions

122 Using the prior distribution, the utility function and the definition of
123 a design space we can optimize the experimental design in this case we are
124 looking for $\delta\theta^*$ that maximizes the following equation:

$$U(\delta\theta^*) = \max_{\delta\theta} \int_{\beta} \log(\det(I(\beta|\delta\theta)))\pi(\beta)d\beta \quad (7)$$

125 The previous integral can be approximated via Monte Carlo sampling

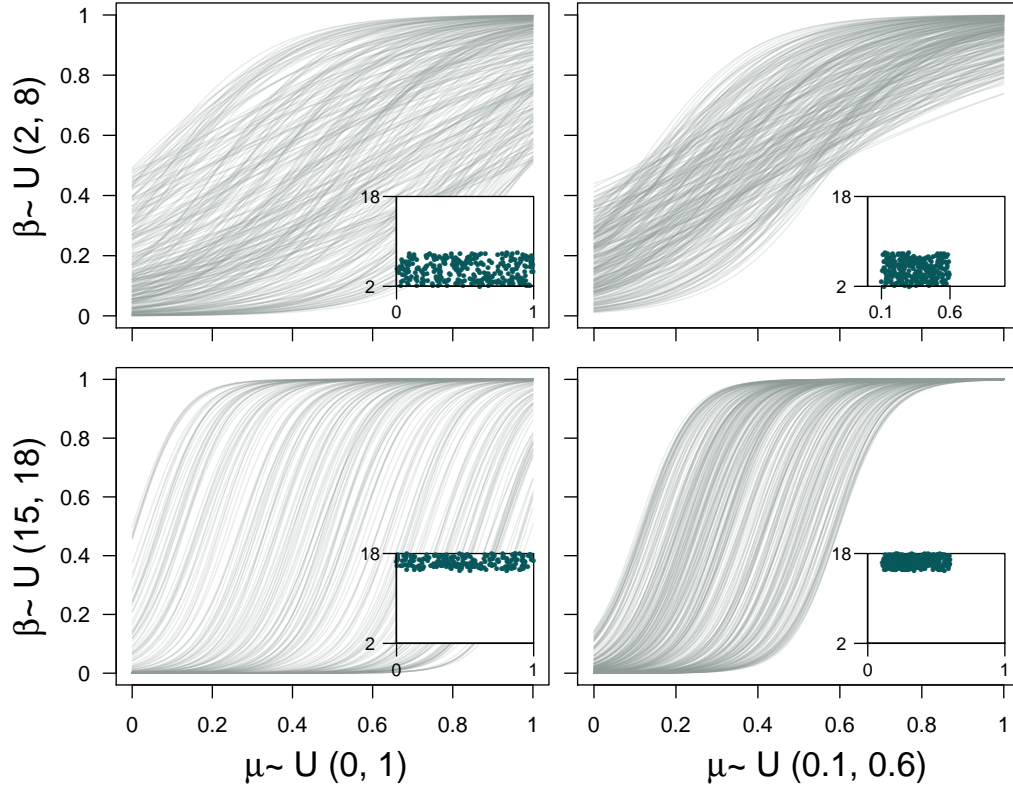


Figure 1: Prior and predictive prior distributions.

126 5. Results

127 5.1. Construction of the prior distribution

128 Prior over model parameters(Gallistel et al 2014) Results Constructing
 129 the prior: we take a multivariate normal distribution with mean and covari-
 130 ance equal to the unbiased estimators for both parameters.

131 5.2. Optimal design

132 Approximating the utility function (integral) through Monte Carlo simu-
 133 lation Utility approximation for 2 Design points

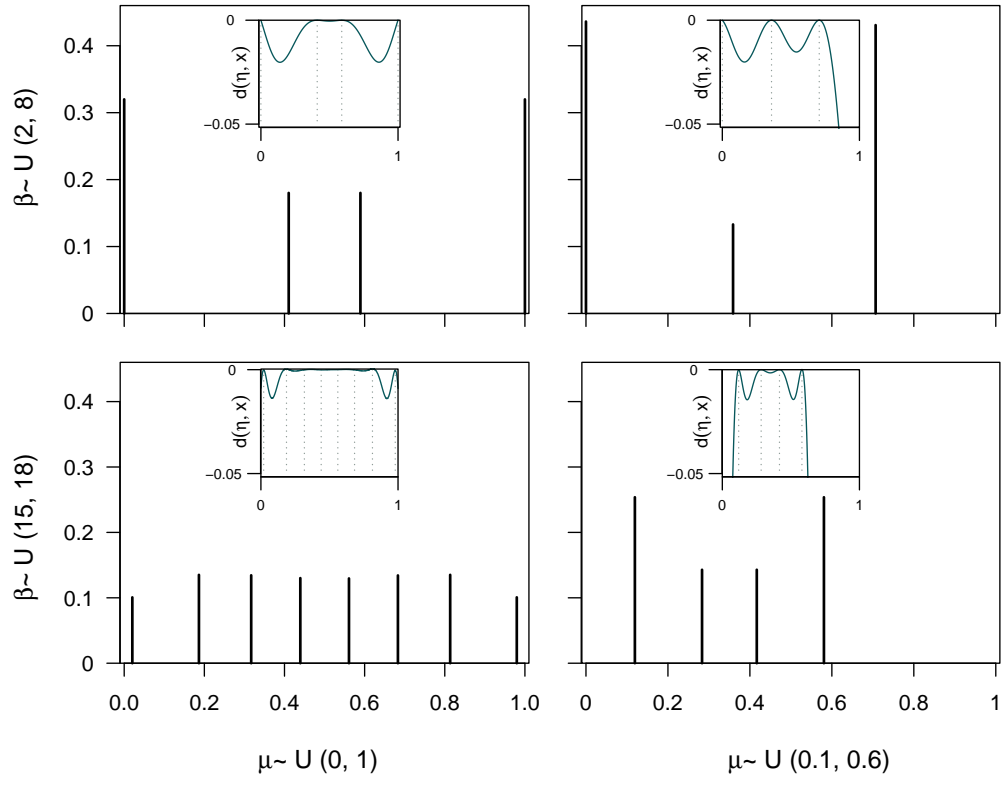


Figure 2: Optimal experimental designs and directional derivatives.

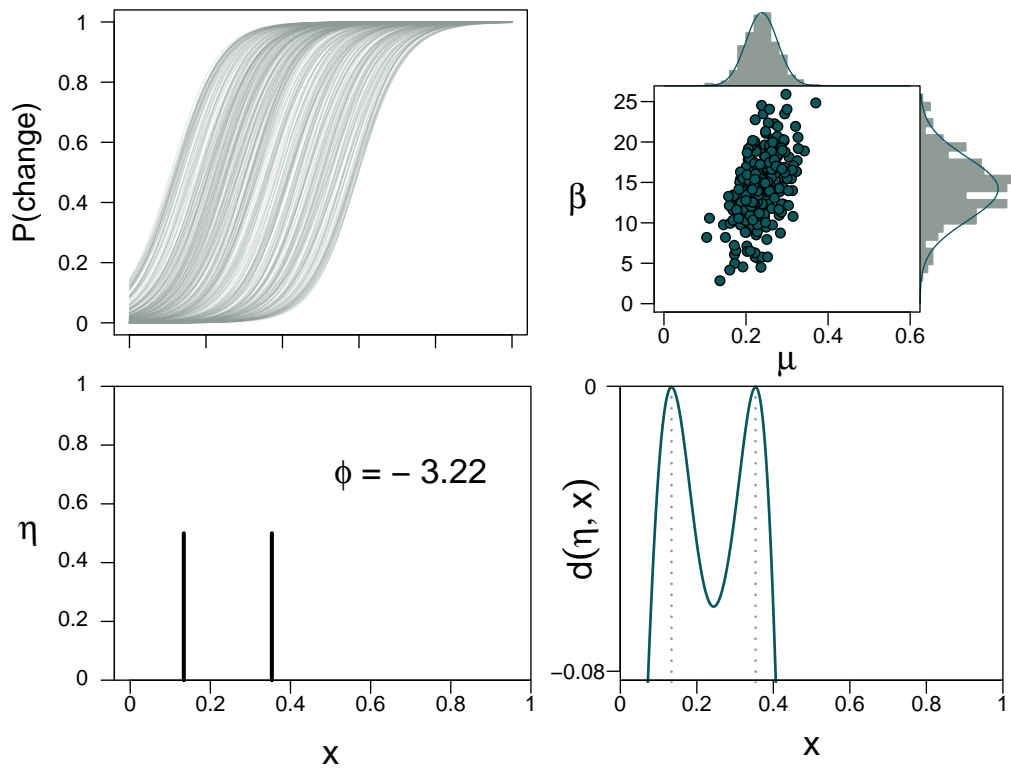


Figure 3: Informative prior distribution, prior predictive, optimal design and directional derivative.

134 the approximation returns a smooth curve over the 2 point design space.

135 **6. Discussion**

136 Optimal design for the example Properties of the most useful points (they
137 land on the points of the curve where the steepness changes most dramati-
138 cally)

139 Advantages of Optimal Design

140 Using models to generate prior distributions.

141 **References**

142 Bernardo, J., 1979. Expected information as expected utility. *The Annals of*
143 *Statistics* 7 (3), 686–690.

144 Chaloner, K., Larntz, K., February 1989. Optimal bayesian design applied
145 to logistic regression experiments. *Journal of Statistical Planning and In-*
146 *ference* 21 (2), 191–208.

147 Clyde, M., Chaloner, K., 2002. Constrained design strategies for improving
148 normal approximations in nonlinear regression problems. *Journal of Sta-*
149 *tistical Planning and Inference* 104, 175–196.

150 Myung, J., Pitt, M., 2009. Optimal experimental design for model discrimi-
151 nation. *Psychological Review* 116 (3), 499–518.

152 Zhang, S., Lee, M., 2010. Optimal experimental design for a class of bandit
153 problems. *Journal of Mathematical Psychology* 54, 499–508.