APPROVAL SHEET

Title of Dissertation: Statistical Modeling and Hypothesis Testing of Chemical-Chemical Interaction: A Nonparametric Approach

Name of Candidate: Mingyu Xi
Doctor of Philosophy, 2015

Dissertation and Abstract Approved:

Anindya Roy
Professor
Department of Mathematics and Statistics

Date Approved: December 10, 2015

NOTE: *The Approval Sheet with the original signature must accompany the thesis or dissertation. No terminal punctuation is to be used.

Abstract

Statistical modeling and hypothesis testing of chemical-chemical interaction: a non-parametric approach

In environmental studies, people are often interested in understanding how exposures to multiple chemicals affect cell survival. One of the key questions is understanding interaction between the chemicals and often understanding the direction of interaction is important. In the absence of known joint models, we take a nonparametric approach using Bernstein Polynomials to model the probability of cell survivals under multiple chemical effects and propose procedures for testing for interaction in the nonparametric setting.

We propose tests for the two most common forms of interaction, Bliss independence and Loewe additivity. To test for Bliss independence we use a two stage approach. We first choose a best model using model selection and then use the "best" model to construct a likelihood ratio test for interaction. We use resampling methods to approximate the critical region of the test. We illustrate our methodology using a reconstructed designed experiment involving cytotoxicity from exposure to common chemicals in batteries such as Nickel, Cadmium and Chromium.

In the second part we generalize conventional parametric Loewe additive reference models to semiparametric and nonparametric zero interaction models. For the semiparametric model we use a one degree of freedom test for interaction that is analogous to classical one degree of freedom test in ANOVA. In the nonparametric approach we use procedures for likelihood ratio tests in non-nested model and investigate the performance of the test via simulation studies.

The final part of the investigation deals with directional interaction. The Bernstein model is well-suited for testing for directional interaction in terms of the coefficients of the model. We propose a test for synergy/antagonism based on the fitted coefficients. In the Loewe additive model we use a contour based test to investigate directional interaction. We also discuss some future directions for the research.

Statistical Modeling and Hypothesis Testing of Chemical-Chemical Interaction: a Nonparametric Approach

Mingyu Xi

Advisor: Dr. Anindya Roy

Department of Mathematics and Statistics

University of Maryland, Baltimore County

Contents

List of Tables

List of Figures

Chapter 1

Introduction

In this dissertation we investigate the concept of statistical interaction under exposure to a mixture of agents (e.g. chemicals, drugs). The main novelty is that we study interaction in a non-parametric setting. In particular we propose non-parametric test for interaction with possible extensions to tests for one-sided alternatives.

1.1 Motivation

The primary motivation of our investigation of robust forms of statistical interaction comes from a chemical-chemical interaction problem. Throughout the dissertation we will use the words agents, chemicals and drugs interchangeably. We introduce the problem of nonparametric interaction in the context of the motivating example. While the data set in the motivating example is proprietary, a description of the problem will suffice toward providing a general introduction to the problem of interest. Nickel(Ni), Cadmium(Cd) and Chromium(Cr) are common metals found in battery wastes. In environmental studies, people are often interested in how the mixtures of chemicals, for example battery waste containing Ni, Cd and Cr, affect cell survival. In this context, a particular question is whether chemicals in the mixtures acted inde-

pendently or did chemicals interact with each other in the mixtures, toward affecting survival probability of cells which are exposed to the mixture? When chemicals act together interaction between agents can occur mainly in three different scenario. First, the case of no interaction where the joint effect of the chemicals is an aggregate of their individual effects. In the context of the cell survival example a version of no interaction would mean the proportion of cell survival in the mixture is the product of the proportion of cells would survive from exposure to each single chemical. Second option for interaction is that there is a positive interaction or synergy between agents. In synergistic interaction, the proportion of cell survived in the mixture is less than the product of the proportion of cell survived from each single chemical exposure. Synergy means that the mixture does more harm to the cells than when the effects of the single chemicals are simply combined together or that the effect of the combination is worse than the combination of the effects. Third option is that there is a negative interaction or antagonism between agents. In antagonism, the proportion of cells survived in the mixture is greater than the product of the proportion of cells survived from each single chemical exposure. Antagonism implies that the mixture does less harm than when the effects of the individual chemicals are simply combined or that the toxic effects of agents works against each other. Interaction, especially the direction of the interaction is of great interest in environmental studies. At low dose individual chemicals may do little harm to the environment, but together the mixture may cause great harm particularly when the number of chemicals in the mixture increases. On the other hand, the mixtures may do less harm than the simple addition of the individual effects due to negative interaction.

In laboratory experiments on cytotoxicity, scientists study cell survival when cells are exposed to the effects of single chemical or mixture of chemicals under different concentrations. Proportion of cells surviving after chemical exposure can not be mea-

sured directly. To measure how many living cells are left after the chemical exposure, scientists use instruments that measure the density of light going through the liquid containing the cells that have been stained by some dye that adhere to dead cells. Due to imperfections of the experiment, there is usually a lot of variation in the repeated measurements. Calibration of the instruments is often difficult. Thus, typically the data obtained on proportion of cell survival is convoluted with measurement error. Using such data often begs careful postulation of models. Simple parametric models are often used for convenience, ease of interpretation and understanding. However, in experiments where no substantive prior knowledge is available, simple parametric models may not be able to provide adequate fit to data. Using such model based parametric analysis may lead to incorrect inference. For example, the number of studies dealing with cytotoxic effect of mixture of battery waste materials are limited and also empirical analysis of the dose response relationship does not provide conclusive evidence toward any particular parametric model. Using incorrect model adversely affects the test for interaction which in turn will have an adverse effect on inference on main effects. Thus, we propose to use flexible nonparametric model to under stand the dose response of mixture and use the nonparametric set up to test for interaction.

1.1.1 Motivating Data

While the data is proprietary, we provide basic description of the parameters of the data/experiment to describe the type of data that we are going to deal with. Our motivating data are from Dr. Yue Ge's lab in U.S. EPA. The data is from a controlled toxicological experiment where cells in assays are exposed to mixture of different chemicals at different concentrations or to individual chemicals at different concentration. In this data set, there are three metals involved: Nickel, Cadmium and

Chromium. Nickel has 8 different levels (0, 10, 30, 60, 75, 100, 130, 180, 240) [μ**M**], Cadmium has 8 levels (0, 3, 5.6, 7.5, 10, 11.5, 13, 15, 18) [μ**M**] and chromium has 8 levels (0, 0.1, 0.3, 1, 1.8, 3, 4.2, 5.6, 10) [μ**M**]. There are three data sets from experiments where cells are exposed to individual metals, three data sets where cells are exposed to combination of two chemicals and one data set containing all three chemical effects. For each chemical combination, there are multiple concentration levels. Each combination of chemical and concentration level has 6 repeated measurements. The data set remarkably rich as has moderately large sample size given the nature of the experiment. The data measures the ratio of the proportion of live cells after the exposure to that before the chemical exposure. Since the data are obtained via an indirect measurement process, the ratio are not restricted to the expected range of [0, 1]. There are several ratio of cell survival that goes above one or below zero. The fact that the cell survival ratio is not strictly restricted to the range $[0, 1]$ may also be due to cell proliferation, an interesting phenomenon of stimulattion of cell division at low doses that results in an increase in the total number of living cells after the exposure. However, without prior model for cell proliferation or definitive information it is not possible to separate the latent process from the measurement process. Figure 1.1 shows the cell survival ratio as a function of the chemical dose combinations. From the plot of the data, it seems that interactions between the chemicals exist although the general form of the interactions is not immediately clear. The marginal response plots show smooth response functions that are probably adequately captured by conventional parametric models. However, the joint structure is more complicated with some plots showing multiple regions of stabilization and decrease. The question is how can we model these cell survival curves under different chemical effects? How do we statistically test if there is any interaction, especially the direction of the interaction?

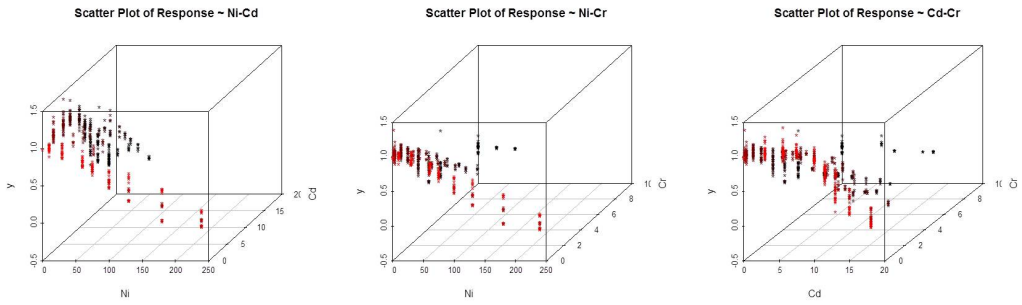


Figure 1.1: Scatter plot of response vs two metals mixture

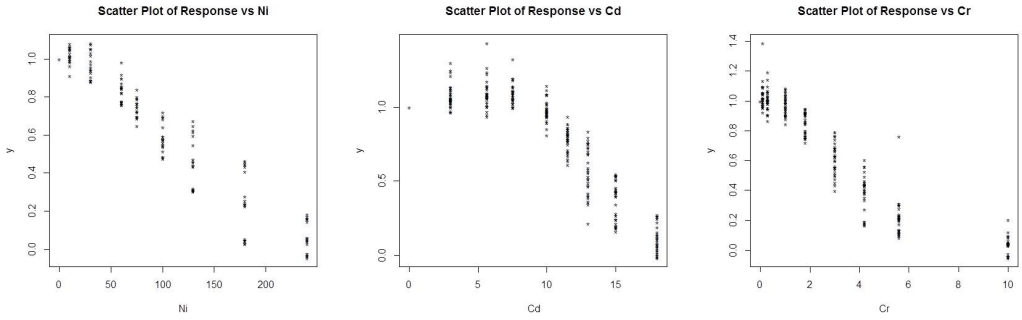


Figure 1.2: Scatter plot of response vs single metal

1.2 Literature Review of Interaction

The definition of interaction is different from discipline to discipline. For example, in biology and genetics, chemistry and biochemistry, or in medicine and pharmacology, the interaction is defined based on the mechanism of the action that is particular to each discipline. In chemistry and biochemistry researchers are interested in chemical-chemical interaction when in medicine or pharmacology scientists are concerned about drug-drug interaction. In principle, chemical-chemical interaction is similar to drug-drug interaction. From a statistical analysis perspective, what is important is not necessarily the mechanism at the molecular levels, but rather the manifestation of the outcome of the action of the agents. Interaction means the effect of two or more agents adding together is not simply what one gets when one combines the effects of the individual agent. However accurate understanding of interaction between agents warrants a more definitive description of the concept.

In order to study interaction, particularly in toxicology experiments dealing with mixtures of agents, one needs well established guidelines about the reference model of no interaction. Only when one understands what constitutes no interaction, one can analyze interaction as a departure from the reference model of zero interaction. There are two most commonly used models for zero interaction: Bliss independence and Loewe additivity.

Bliss independence assumes that the relative effect of an agent at a particular concentration is independent of the presence of the other agent. For example, if drugs A and B individually cause 50% growth inhibition, then Bliss independence predicts that, in combination, drugs A and B decrease growth by $1 - 0.5 * 0.5$, or 75%. Positive or negative deviations from this prediction describe synergistic and antagonistic interactions, respectively. Thus Bliss independence assumes that the agents are mutually non-exclusive, in fact they cooperate and all of them can be

active at the same time. In the simplest mechanical model where agents affect the target through a specific path, under Bliss independence each agent will have its own specific, independent interaction site. The concept was introduced by Bliss (1939) in a seminal paper and later generalized and analyzed in many different context; see for example Berenbaum (1977), Goldoni and Johanssen (2007).

Loewe additivity defines an agent to be non-interacting with itself. For example, two drugs each can kill the cells but they differ the respective doses needed to kill same amount of the cells. The first drug may need 100 mg to kill 50% cells while the second drug may only need 25 mg. These are indicators of drug potency. The drug that requires lower dose is said to have a greater potency than the other. The dose ratio, in this case $100/25 = 4$, called relative potency. This same relative potency may or may not apply to all levels of effect for these two drugs. Let drug A, the lower potency drug, the dose when it acts alone be A; for drug B to achieve the same effect, let the dose needed be B. The relative potency R is A/B, a value greater than 1. For a given mixture having the same effect, let a and b be the dose of the respective constituents. Because these drugs are assumed to have a constant relative potency(R), the mixture (a, b) can be expressed as an equivalent quantity of either drug. If drug A is the reference drug, then the mixture dose satisfies the relation $a + Rb = A$. Equivalently, the same mixture can be expressed in terms of drug B : $a/R + b = B$. These two equations mean that the doses in the mixture contribute to the effect in accordance with the individual drug potency, a situation that is termed additive, or more specifically Loewe additive. Rearrangement of the potency formula gives a more familiar form: $a/A + b/B = 1$. This also describes the isobolograms (Fraser 1870,1872) which are lines of equal activity and often used a simple graphical tool for understanding violation of zero interaction model. Loewe and Muischnek (1926) and Loewe (1953) used the isobolograms to define the reference

model for zero interaction. Under Loewe additivity it is assumed that the agents have mutually exclusive effect on the target, and at most one agent is active at any given time. This corresponds to a common site of interaction where if the site is occupied by one agent then the other has to wait.

There are several other approaches for understanding interaction between agents. The reviews by Berenbaum (1989) Greco et. al (1995) provide excellent reviews of the approaches. Another commonly used form of zero interaction model is the median effect principle (Chou 2010) but we will focus on Bliss independence and Loewe additivity in this research.

1.3 Statistical Concepts

Next we define some of the statistical concepts that are used in the dissertation. They are well known concepts, but we include them for completeness.

1.3.1 Statistical Models for Interaction

Although our motivating data provides continuous response in the form of survival ratio, intrinsically the data is based on binary responses of individual cells. Many of the response variables in the toxicology experiments are binary. The seminal article of Bliss (1939) deals with binary response as well. Thus, to provide some context to our modeling framework we start with a description of binary response models. In statistics literature, specially in biostatistical literature binary response models are as indispensable as normal regression models. In toxicology, Bliss (1939) used a probit model to describe the latent effect of chemicals. Bliss (1939) proposed transforming the percentage of cells killed into a "probability unit" (or probit) which is linearly related to the current understanding of the probit model. The method introduced

by Bliss (1939) was carried forward to probit analysis by Finney and Stevens (1948), Finney (1952) and has been used heavily in toxicology experiments.

Of course, logit is another form of popular analysis for binary response data. There are several other link functions that are used in statistical literature to describe the relationship between a binary response and observed values of agents. Another popular model that appears in toxicology dose response modeling is the Hill model. The USEPA guidelines for benchmark dose rate uses the Hill model as a reference model. Similar functions can be used as mean function in nonlinear regression model for continuous responses whose true value is expected to be between 0 and 1. Our motivating data falls in this category. As stated before, our objective is to move beyond these simple parametric classes as develop a nonparametric framework that is flexible and can adapt to unknown forms for dose responses as well as serve as a platform for analyzing interaction. To this end we describe a popular nonparametric class, the Bernstein class.

For a given function f on [0, 1], Bernstein polynomial is defined as

$$B_n(f;x) = \sum_{r=0}^{n} f(\frac{r}{n})\binom{n}{r}x^r(1-x)^{n-r}$$

for each positive integer n. There is a sequence of Bernstein polynomials corresponding to each function f. A useful property of Bernstein polynomial is, if f is continuous on $[0,1]$, its sequence of Bernstein polynomials converges uniformly to f on $[0,1]$. This property provides basis for modeling non-parametric classes of functions on $[0,1]$ by their Bernstein approximations.

A common model for departure from the null interaction model of additivty is the Tukey's one degree of freedom of test for additivity (Tukey 1949). For developing our alternative models, we use analogus concepts. However, in the nonparametric set up, the alternative space is vast and a single degree of freedom departure from no interaction may not capture all the intersting deviation from the zero interaction

model.

1.3.2 Likelihood Ratio Test

Likelihood ratio test is a statistical test used to compare the fit of two models, one of which (the null model) is a special case of the other (the alternative model). The test is based on the likelihood ratio, which expresses how many times more likely the data are under one model than the other. The likelihood ratio, or equivalently its logarithm, can be used to compute a p-value, or can be compared to a critical value to decide whether to reject the null model in favor of the alternative model.

Let $\mathbf{X}$ be data generated from a model $f(X, \theta)$, if we want to test the hypothesis $H_o : \Theta$, one could use the test $\Lambda(\theta) = \frac{\sup_{\theta \in \Theta_0} L(\theta|data)}{\sup_{\theta \in \Theta} L(\theta|data)}$, where $L(\theta|data)$ is the likelihood of θ at a given sample $data$ of size n. The test rejects H_0 for small values of $\Lambda(\theta)$ or equivalently large values of $-2 \log \Lambda_\theta$. Typically the critical values are obtained from limiting distribution of the likelihood ratio statistic under an increasing n asymptotic framework. However, when Θ is large dimensional but Θ_0 is finite dimensional (e.g. in semiparametric problems) or Θ_0 is also large dimensional (e.g. nonparametric problems) the standard asymptotic distributions may not hold. In semiparametric problems, under mild assumptions Murphy and Van der Vaart (1997) and Banerjee (2005) establish asymptotic distribution of $-2 \log \Lambda_\theta$. In fully nonparametric set up, the tools for finding the LRT and its distribution is problem specific. There are several different modifications including penalization and sieve methods that are used in the literature.

1.3.3 Residual Bootstrap

Bootstrapping is a statistical method for estimating the sampling distribution of an estimator by sampling with replacement from the original sample, most often with

the purpose of deriving robust estimates of standard errors and confidence intervals of a population parameter like a mean, median, proportion, odds ratio, correlation coefficient or regression coefficient. The sampling distribution we are interested is $-2log\Lambda_\theta$.

Residual bootstrap is convenient way of applying resampling procedures in the regression context. The method is straightforward and involves estimating the residuals from original sample and then sampling with replacement from these residuals to create bootstrap samples. Properties of residual bootstrap, particularly under a high number of parameters as in the present set up has been studied by Freedman (1981), Bickel and Freedman (1983).

Chapter 2

Nonparametric Test of Bliss Independence

In this chapter we build a flexible statistical test for checking if there is interaction between agents when the response is obtained from exposure of bioassays to mixtures of several agents. The particular form of interaction that we are interested in is Bliss Independence. We start by describing a general model and the relevant aspects of the nonparametric testing paradigm.

2.1 Nonparametric Modeling Chemical-Chemical Interaction

We illustrate the methodology and the model, for example, with a two-agent interaction when an assay of living cells is exposed to a mixture of two of the common chemicals found in battery waste, Cadmium (Cd) and Chromium (Cr) and where the interest is in finding out whether Cd and Cr act independently. However, the methods are easily extended to multi-agent models. Let x_1 and x_2 denote the dose levels for

the two chemical agents. Hereafter we will scale the dose space of the two agents to the unit square. This is possible because typically there is an upper bound on the possible dose beyond which the survival rate is practically zero. Scaling makes the doses unit free, an aspect that will help in generalizing conventional parametric models where the basic building blocks maybe unit free. Thus, unless otherwise specified we assume throughout $(x_1, x_2) \in [0, 1]^2$.

Let $y(x_1, x_2)$ denote the response, i.e. the proportion of living cells, as observed by indirect measurements in the laboratories when the cells are exposed to mixture of chemicals at a dose combination (x_1, x_2). The response is inherently a continuous variable with ideally restricted in the range [0, 1]. However, due to the indirect measurement processes and due to error accrued during the measuring process, the actual value of y could be outside the range. There are several established methods for measuring cytotoxicity. The methods usually rely on staining of dead cell and the living cells are counted by fluorescence and luminescence plate readers. Each method has its advantage but none are without error. The errors may occur due to instrument error (although that part is generally minimal), human error, other artifacts of assays, or due to other natural phenomena such as cell proliferation or the dyes not able to distinguish between dead cells and cells rapidly losing functionality. Some are more prone to measurement error than other. While there maybe more advanced methods available, compromise between accuracy and cost may lead to more error prone methods being used. In essence, the measurement process does have a potentially significant impact on the observed cytotoxicity and the variability due to the measurement process may or may not be independent of the unobserved cytotoxicity. Of course the case when the measurement error depends on the underlying true cytotoxicity is not an identifiable problem without further information about the nature of relationship and is not discussed here. Hereafter we assume that the measurement

error is independent of the true cytotoxicity and appears in an additive form.

A reasonable and convenient model for the response is

$$y(x_1, x_2) = \mu(x_1, x_2) + \epsilon(x_1, x_2) \tag{2.1}$$

where $\mu(x_1, x_2)$ denotes the true survival rate, a quantity that is assumed to be monotonically decreasing in either x_1 or x_2 and also bounded between zero and one. While μ is range restricted, due to the measurement error model the response y is not. Values of y can be below zero is borne out of the fact that the response may be derived as the difference of two measuring process, each of which maybe subject to error. Using the present context,

$$\mu(x_1, x_2) = p(\text{cell survival}|Cd = x_1, Cr = x_2).$$

Similarly one can define the marginal survival rates as

$$\mu_1(x_1) = p(\text{cell survival}|Cd = x_1)$$
$$\mu_2(x_2) = p(\text{cell survival}|Cr = x_2)$$

Under the *Bliss Independence* assumption that states that there is no multiplicative interaction between the agents, one expects the factorization

$$\mu(x_1, x_2) = \mu_1(x_1)\mu_2(x_2). \tag{2.2}$$

The possible alternatives to Bliss independence is either a two-sided alternative $\mu \neq \mu_1\mu_2$ or one-sided alternatives such as $\mu > \mu_1\mu_2$ pertaining to antagonism or $\mu < \mu_1\mu_2$ pertaining to synergy between the agents. Moreover, scientists maybe interested in violation of the null hypothesis only over a smaller subset of the dose-space, e.g. lower doses. In this chapter we only consider the broad two-sided global alternative to build our test for interaction.

For testing the null hypothesis of bliss independence one could postulate a parametric model for μ and reduce the hypothesis to a null hypothesis about the value

of the parameters. However, as in the example of battery chemical interaction, little or nothing maybe known a priori about the joint behavior of the agents and using a prespecified parametric form maybe misleading. In such situations, a more prudent approach is to build a broad flexible model for the response and build the test under the flexible model. To build a nonparametric model, one still has to account for qualitative features of the survival rate function that maybe known beforehand. For example, even though the general shape and form of the response curve maybe unknown, other qualitative features maybe known: e.g., decreasing in either argument, range restricted. In addition there maybe further special features, such as smoothness of the model, that scientists maybe interested to incorporate.

As mentioned before, the function μ would generally be continuous. Thus broadly, the parameter space can be considered a subset of the function space $C([0,1]^2$, an infinite dimensional space. There are some constraints that are natural constraints on the parameter imposed by the model and there are certain constraints that will be imposed by the null hypothesis of no multiplicative interaction. The parameter space for the problem is then

$$\Theta = \{\mu(x_1, x_2) \in C[0,1]^2 : \mu \text{ is decreasing in } x_1 \text{ and } x_2,\ 0 \leq \mu \leq 1\} \tag{2.3}$$

is the infinite dimensional subset of $C[0,1]$ that satisfy the natural constraints that are imposed on the model μ. Also, let

$$\mathcal{F} = \{\mu(x_1, x_2) : \exists \mu_1, \mu_2 \ni \mu(x_1, x_2) = \mu_1(x_1)\mu_1(x_2)\ \forall x_1, x_2\}$$

be the set of functions that satisfy the constraint of factorization under the Bliss independence unll model. The null set is the set of functions

$$\Theta_0 = \Theta \bigcap \mathcal{F}. \tag{2.4}$$

Note that the null set is an infinite dimensional set as well. Thus, both the null and the alternative sets are infinite dimensional and care must be exercised in order to

test the null aganist any alternative and conventional finite dimensional results might not be applicable to this case.

Before we describe the model and the testing procedure we want to discuss the general method that we will be employing for testing the nonparametric null hypothesis that the true function $\mu \in \Theta_0$. From a finite sample of size n it is not reasonable to expect that one can differentiate between the classes of functions Θ_0 and $\Theta - \Theta_0$ effectively. Either space is too big. To build a feasible procedure we use a *sieve* method in which the null and the alternative sets are replaced by a sequence of approximating sets that increase (the sequences area not necessarily nested) to desired null and alternative as $n \to \infty$, but for each n provide a possible test that distinguishes the approximating null from the approximating alternative with reasonable power. In general, let $\Theta^{(n)}$ be a *sieve* for the entire parameter set Θ. Then let $\Theta_{0,n} = \Theta_0 \bigcap \Theta^{(n)}$ denote the working null space and let $\Theta_{1,n} = \Theta^{(n)} - \Theta_{0,n}$ denote the working alternative space. The objective will be to test $\mu \in \Theta_{0,n}$ against $\mu \in \Theta_{1,n}$.

The sieve is assumed to satisfy the property $\lim_{n\to\infty} \Theta^{(n)} = \Theta$ where the convergence is in terms of set distances. Since in the present case the parameter space is metrizable, the set distances will be based on the metric on the parameter space. Thus, given a metric $d(\cdot, \cdot)$ on Θ, one could define the induced distance $D(\cdot, \cdot)$ between Θ and $\Theta^{(n)}$ as

$$D(\Theta, \Theta^{(n)}) = \inf_{\mu\in\Theta, \mu_n \in \Theta^{(n)}} d(\mu, \mu_n).$$

Thus, sieve is assume to satisfy the property:

Property C: Given any $\epsilon > 0$, there exists N_ϵ such that for every $n > N_\epsilon$ we have $D(\Theta, \Theta^{(n)}) < \epsilon$.

A common metric would be the sup distance given as

$$d(\mu_1, \mu_2) = \sup_{(x_1,x_2)\in[0,1]^2} |\mu_1(x_1, x_2) - \mu_2(x_1, x_2)|.$$

It maybe sometimes helpful to describe the set convergence in terms of d neighbor-

hood. A set Θ^ϵ will be called an ϵ-neighborhood of a set Θ if for any $\mu \in \Theta$ there exists a function $\mu^\epsilon \in \Theta^\epsilon$ such that $d(\mu(x_1, x_2), \mu^\epsilon(x_1, x_2)) < \epsilon$ As mentioned earlier, the sequence of sets $\Theta^{(n)}$ need not be nested.

To obtain such a sieve for Θ we use the Bernstein functions. As discussed in the introduction, the Bernstein polynomials form a dense set in $C[0,1]^2$ and hence we can obtain the desired sieve by approximating the class of joint and the marginal models by Bernstein polynomials. Let

$$B_{J_1,J_2}(x_1, x_2) = \sum_{j_1=1}^{J_1} \sum_{j_2=1}^{J_2} \beta_{j_1,j_2}^{(12)} B(x_1; j_1, J_1 - j_1 + 1) B(x_2; j_2, J_2 - j_2 + 1) \tag{2.5}$$

be a Bernstein polynomial of order (J_1, J_2) with vector of coefficients $\beta^{(12)} = (\beta_{11}^{(12)}, \ldots, \beta_{J_1,J_2}^{(12)})$. Similarly, let

$$B_{J_1}(x_1) = \sum_{j_1=1}^{J_1} \beta_{j_1}^{(1)} B(x_1; j_1, J_1 - j_1 + 1) \tag{2.6}$$

$$B_{J_2}(x_2) = \sum_{j_2=1}^{J_2} \beta_{j_2}^{(2)} B(x_2; j_2, J_2 - j_2 + 1) \tag{2.7}$$

be the Bernstein polynomials of order J_1 and J_2 with coefficients $\beta^{(1)} = (\beta_1^{(1)}, \ldots, \beta_{J_1}^{(1)})$ and $\beta^{(2)} = (\beta_1^{(2)}, \ldots, \beta_{J_2}^{(2)})$, respectively. We have suppressed the dependence on the coefficients in the notation of the polynomials as it should be clear from the context, but the main idea is to parameterize the class of functions in terms of their coefficients. Let $\mathcal{B}$ denote the class of all Bernstein functions on $[0,1]^2$. Then the parameter set is described in terms of the Bernstein expansions as $\Theta^B = \mathcal{B} \bigcap \Theta$, and the null set is $\Theta_0^B = \Theta^B \bigcap \mathcal{F}$. The uniform approximation property of the Bernstein polynomials imply that original parameter space can be obtained as a closure of the spaces described by the Bernstein polynomials, i.e, $\overline{\Theta^B} = \Theta$ and $\overline{\Theta_0^B} = \Theta_0$.

To be able to maneuver the parameter space in a flexible manner we need to describe the space in terms of the coefficients, thereby projecting the problem from the function space to the Euclidean space. To describe the problem in terms of the

coefficients we will use the convention that $\beta^{(12)} \in \Theta^B$ means that the Bernstein polynomial with coefficients $\beta^{(12)}$ belongs to Θ^B.

For the Bernstein functions to form an appropriate sieve they have to maintain the shape and range restrictions. Also, to satisfy the null of bliss independence, the coefficients should be such that the joint factorizes in terms of the marginals. Describing the factorization in terms of the coefficients while maintaining the shape and range restrictions is an intractable problem. A sufficient hypothesis is that the coefficients factorizes themselves and are decreasing. The sufficient hypothesis for factorization is

$$H_0^\beta : \quad \beta^{(12)} = \beta^{(1)} \otimes \beta^{(2)} \text{ for some} \beta^{(1)} \text{ and } \beta^{(2)}.$$

The condition that each member of the set Θ^B is monotonically decreasing is too complex to describe in terms of the coefficients $\beta^{(12)}$. A sufficient condition is that the coefficients of the approximating Bernstein functions are decreasing in either indices. However, this maybe too restrictive. Using a restrictive sufficient hypothesis will make the tested null a very small subset of the intended null set and hence the test will have poor power properties.

In order to have a feasible sieve we describe the shape restriction in terms of the levels of the function restricted to a pre-determined grid. Fix a grid $G(n) = \{(z_{1i}, z_{2i}), i = 1, \ldots, g\}$. In fact we will use grid that is factorizable in the sense there are marginal grids $G_1 = \{z_{1i}, i = 1, \ldots, g_1\}$ and $G_2 = \{z_{2i}, i = 1, \ldots, g_2\}$ such that $G = G_1 \oplus G_2$. Hence, $g = g_1 * g_2$.

The size of the grids g_1, g_2 will be allowed to increase with the sample size, but for notational convenience we have suppressed the dependence on the sample size. We will choose a sequence of grid that will satisfy the following property:

Property G: As $n \to \infty$, $G = G(n)$ becomes dense in $[0, 1]^2$ in the sense that given any $\epsilon > 0$, there exist N such that $n > N$ implies for any $(x_1, x_2) \in [0, 1]^2$ there exists

$(z_1, z_2) \in G(n)$ with $\max\{|x_1 - z_1|, |x_2 - z_2|\} < \epsilon$.

Considering the constrained space

$$\Theta^G = \{\mu(z_1, z_2) \in \Theta : \mu(z_1, z_2) \text{ is decreasing and restricted to [0,1] for } (z_1, z_2) \in G\}$$

i.e.,

$$\Theta^G = \{\mu(z_1, z_2) \in \Theta : \frac{\partial}{\partial x_j}\mu(x_1, x_2)\Big|_{z_{1i}, z_{2i}} \leq 0, j = 1, 2, 0 < \mu < 1 \text{ for } (z_{1i}, z_{2i}) \in G\}.$$

Because the function μ are continuous on the compact set, they are uniformly continuous and hence along with the property G, we have as $n \to \infty$, $\Theta^G \to \Theta$. To construct the sieve we thus focus on $\Theta^G \bigcap \Theta^B$. We well denote the set by Θ^G_β to emphasize the parameterization in terms of the coefficients. In terms of the approximating Bernstein sequence, because the function as well as the partial derivatives are linear in the coefficients, one could translate the constraints to linear constraints on the coefficients.

$$\begin{aligned}
\sum_{j_1=1}^{J_1}\sum_{j_2=1}^{J_2}\beta^{(12)}_{j_1,j_2}B(z_{1i}; j_1, J_1 - j_1 + 1)B(z_{2i}; j_2, J_2 - j_2 + 1) < 1, \quad i = 1, \ldots, g \\
-\sum_{j_1=1}^{J_1}\sum_{j_2=1}^{J_2}\beta^{(12)}_{j_1,j_2}B(z_{1i}; j_1, J_1 - j_1 + 1)B(z_{2i}; j_2, J_2 - j_2 + 1) < 0, \quad i = 1, \ldots, g \\
\sum_{j_1=1}^{J_1}\sum_{j_2=1}^{J_2}\beta^{(12)}_{j_1,j_2}\frac{\partial}{\partial x_1}B(x_1; j_1, J_1 - j_1 + 1)\Big|_{z_{1i}} B(z_{2i}; j_2, J_2 - j_2 + 1) < 0, \quad i = 1, \ldots, g, \\
\sum_{j_1=1}^{J_1}\sum_{j_2=1}^{J_2}\beta^{(12)}_{j_1,j_2}B(z_{1i}; j_1, J_1 - j_1 + 1)\frac{\partial}{\partial x_2}B(x_2; j_2, J_2 - j_2 + 1)\Big|_{z_{2i}} < 0, \quad i = 1, \ldots, g.
\end{aligned} \tag{2.8}$$

Thus, the sieve can be elegantly written in terms of linear constraints on $\beta^{(12)}$,

$$\Theta^G_\beta = \{\beta^{(12)} : A\beta^{(12)} \leq b\},$$

where b is the $3J_1J_2 \times 1$ vector $b = (1, \ldots, 1, 0, \ldots, 0, 0, \ldots 0)$. Note that the dimension and the size of the sieve is allowed to increase with sample size by letting the grid size g depend on n and the order of the polynomials J depend on n as well.

Next we describe the null in terms of the sieve. Given that the joint model factorizes under the null, we can describe the sieve in terms of the marginal model Bernstein coefficients $\beta^{(1)}$ and $\beta^{(2)}$. The relevant hypothesis can be written in terms of the coefficients

$$H_0 : \beta^{(12)} = \beta^{(1)} \otimes \beta^{(2)} \text{ for some } \beta^{(1)} \text{ and } \beta^{(2)} \text{ vs } H_1 : \nexists \beta^{(1)} \text{ and } \beta^{(2)}, \ni \beta^{(12)} = \beta^{(1)} \otimes \beta^{(2)}.$$

For $j = 1, 2$ consider the spaces

$$\Theta^{G_j} = \{\mu_1(x_j) : \frac{\partial}{\partial x_j}\mu_j(x_j)\Big|_{z_{ji}} \leq 0, 0 < \mu_j < 1 \text{ for } z_{ji} \in G_j\}.$$

In terms of the Bernstein approximation, the spaces Θ^{G_j} can be uniformly approximated by $\Theta^{G_j}_{\beta^{(j)}}$ where

$$\Theta^{G_j}_{\beta^{(j)}} = \{\sum_{k=1}^{J_j} \beta_k^{(j)} B(x_j; k, J_j - k + 1) : A_j\beta^{(j)} < b_j\}$$

The matrix A_1 and the vector b_1 are obtained from the constraints that the marginal model for the first agent is decreasing and range restricted on the marginal grid G_1, i.e.,

$$\begin{aligned}
\sum_{j_1=1}^{J_1} \beta_{j_1}^{(1)} B(z_{1i}; j_1, J_1 - j_1 + 1) < 1, \;\; i = 1, \ldots, g_1 \\
-\sum_{j_1=1}^{J_1} \beta_{j_1}^{(1)} B(z_{1i}; j_1, J_1 - j_1 + 1) < 0, \;\; i = 1, \ldots, g_1 \\
\sum_{j_1=1}^{J_1} \beta_{j_1}^{(1)} \frac{\partial}{\partial x_1} B(x_1; j_1, J_1 - j_1 + 1)\Big|_{z_{1i}} < 0, \;\; i = 1, \ldots, g_1
\end{aligned} \tag{2.9}$$

and the matrix A_2 and the vector b_2 are obtained similarly from the constraint that

the marginal model for the second agent is decreasing and range restricted over G_2

$$\begin{aligned}
\sum_{j_2=1}^{J_2} \beta_{j_2}^{(2)} B(z_{2i}; j_2, J_2 - j_1 + 1) < 1, \;\; i = 1, \ldots, g_2 \\
-\sum_{j_2=1}^{J_2} \beta_{j_2}^{(2)} B(z_{2i}; j_2, J_2 - j_1 + 1) < 0, \;\; i = 1, \ldots, g_2 \\
\sum_{j_2=1}^{J_2} \beta_{j_2}^{(2)} \frac{\partial}{\partial x_2} B(x_2; j_2, J_2 - j_2 + 1)\Big|_{z_{2i}} < 0, \;\; i = 1, \ldots, g_2
\end{aligned} \tag{2.10}$$

Letting $\beta = (\beta^{(1)}, \beta^{(2)})$ then the sieve for the null set can be written in terms of β as

$$\Theta_{0,\beta}^{G} = \{\beta = (\beta^{(1)}, \beta^{(2)}) : A_j \beta^{(j)} < b_j, j = 1, 2\}. \tag{2.11}$$

In terms of the joint model coefficients, it is understood that the null is

$$\{\beta^{(12)} = \beta^{(1)} \otimes \beta^{(2)} : (\beta^{(1)}, \beta^{(2)}) \in \Theta_{0,\beta}^{G}\}$$

which denotes the set of coefficient vectors that can be factorized and the corresponding marginal models satisfy the restrictions over the marginal grids. Even though the null is described via the $(J_1 + J_2)$ dimensional vector $\beta = (\beta^{(1)}, \beta^{(2)})$ with a slight abuse of notation we will equivalently express the null as $\beta = \beta^{(1)} \otimes \beta^{(2)} \in \Theta_{0,\beta}^{G}$.

Having set up the null and the alternative sets in terms of the sieve or the associated coefficient vectors, we proceed with the testing procedure. We will develop the test based on a likelihood. Since the errors in the response are mostly measurement errors, a reasonable distributional assumption is

$$\epsilon(x_1, x_2) \sim N(0, \sigma^2).$$

Of course, one has to check into the validity of the distributional assumption. The distributional assumption can be modified. In certain scenarios, heteroskedastic with variance $\sigma^2(x_1, x_2)$ depending on the dose combination, maybe appropriate. One should try to incorporate any additional information available about the errors into

the distributional assumptions. There is much scope for debating and modeling the measurement process. For the current investigation we choose the additive constant variance normal model. Apart from being a common assumption about general measurement errors, the motivating data sets did not provide evidence against such an assumption.

The distributional assumption on the errors allows us to write down a log-likelihood for the coefficients β (taken as $\beta^{(12)}$ in the unconstrained case and as $\beta^{(1)} \otimes \beta^{(2)}$ in the null constrained case) as

$$\ell(\beta|data) = c - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \mu(x_{1i}, x_{2i}))^2$$

where

$$\mu(x_{1i}, x_{2i}) = \sum_{j_1=1}^{J_1}\sum_{j_2=1}^{J_2}\beta_{j_1,j_2}^{(12)}B(x_{1i}; j_1, J_1 - j_1 + 1)B(x_{2i}; j_2, J_2 - j_2 + 1),$$

and c is a constant. The associated likelihood will be denoted by $L(\beta|data)$. The likelihood immediately provides us a with a likelihood ratio test.

$$\Lambda(data) = \frac{\max\limits_{\beta\in\Theta_{0,\beta}^{G}} L(\beta|data)}{\max\limits_{\beta\in\Theta_{\beta}^{G}} L(\beta|data)}$$

where the denominator is the maximum of the likelihood over all possible J_1J_2 dimensional coefficient vectors of the two-dimensional Bernstein polynomials that make the polynomials monotonically decreasing and range restricted over the grid G and the numerator is over all possible J_1J_2 coefficient vectors that can be factorized as $\beta^{(1)} \otimes \beta^{(2)}$ for some J_1 dimensional vector $\beta^{(1)}$ and J_2 dimensional vector $\beta^{(2)}$ such that the one dimensional Bernstein polynomials described by $\beta^{(1)}$ and $\beta^{(2)}$ are monotonically decreasing and range restricted over the grids G_1 and G_2, respectively.

The rejection rule based on the LRT at a nominal level α would be

$$\mathcal{R}_\alpha = \{Data : \Lambda(data) < c_\alpha\}$$

where c_α is a critical value defined as the lower α percentile of the distribution of $\Lambda(data)$. Typically, under mild assumptions on the model, for large sample size $-2\ell(\beta|data) \xrightarrow{d} \chi^2_{df}$ for some degrees of freedom df when the null hypothesis is true. Equivalently

$$\mathcal{R}_\alpha = \{-2\log(\Lambda) \geq \chi^2_{df,\alpha}\}$$

But under the increasing dimension and increasing complexity, it is not clear if the large sample distribution is a good approximation to the distribution of $-2\log(\Lambda)$. It seems that there is convergence and a χ^2 distribution is a reasonable fit, although the degrees of freedom is not obvious from the results. Theoretical investigation is needed in determining the correct value of the df for the limiting χ^2 distribution. Without any such definitive result, we propose to use a resampling method for approximating the finite sample distribution of the LRT and thereby establishing the critical region. To this end we use a residual bootstrap procedure. The main goal is to generate the distribution of the test under the null hypothesis. We use the best constrained (factorized) approximation to the data as the null model and generate replicates of the data under constrained model by adding copies of errors to the constrained model fit. To obtain the error replicates we resample (with replacement) from the residuals obtained from the unconstrained fit. Since the Bernstein model is uniformly converging to the true model regardless whether null hypothesis holds or not, the unconstrained residuals are expected to mimic the true errors. The exact steps for the residual bootstrap algorithm are as follows:

1. Fit both constrained and unconstrained models (i.e.the factorized null model and the general model but both model having shape and range restrictions) and get $\Lambda(data)$.

2. Let

$$\hat{\mu}_c(x_1, x_2) = \sum_{j_1=1}^{J_1} \sum_{j_2=1}^{J_2} \hat{\beta}_{j_1}^{(1)} \hat{\beta}_{j_2}^{(2)} B(x_1; j_1, J_1 - j_1 + 1) B(x_2; j_2, J_2 - j_2 + 1)$$

be the constrained fit where $\hat{\beta}^{(1)}, \hat{\beta}^{(2)}$ are estimated factors of $\beta^{(12)}$. Similarly let

$$\hat{\mu}_{nc}(x_1, x_2) = \sum_{j_1=1}^{J_1} \sum_{j_2=1}^{J_2} \hat{\beta}_{j_1,j_2}^{(12)} B(x_1; j_1, J_1 - j_1 + 1) B(x_2; j_2, J_2 - j_2 + 1)$$

be the fit from the maximization of the likelihood when the coefficient vector in not constrained to be factorizable as $\beta^{(12)} = \beta^{(1)} \otimes \beta^{(2)}$. Also let $e_{nc}(x_1, x_2) = y(x_1, x_2) - \hat{\mu}_{nc}(x_1, x_2)$ be the residuals from the unconstrained fit. Let $e_{nc} = (e(x_{11}, x_{21}, \ldots, e(x_{1n}, x_{2n}))$ be the residual vector from the unconstrained fit.

3. Generate bootstrap replicates by adding back unconstrained residual to constrained fit, i.e,

$$y^*(x_1, x_2) = \hat{\mu}_c(x_1, x_2) + e_{nc}^*(x_1, x_2)$$

where e_{nc}^* is a sample of size n form the n dimensional vector e_{nc} and the sampling is done with replacement. Let $y^* = (y^*(x_{11}, x_{21}, \ldots, y^*(x_{1n}, x_{2n}))$ be the entire replicate vector. We repeat the replication procedure B times for some moderately large number B to generate B bootstrap replicates $Y_1^*, \ldots, y_2^*$.

4. Generate $\Lambda_1, \ldots, \Lambda_B$ as B bootstrap replicates of Λ based on the bootstrap replicates $Y_1^*, \ldots, y_2^*$.

5. Reject if $\Lambda(data) < \Lambda_\alpha$ where Λ_α is the lower α percentile of the sample $\Lambda_1, \ldots, \Lambda_B$

2.2 Numerical illustration

We investigated the properties of the testing procedure via a limited simulation study. The number of Monte Carlo replications and the number of bootstrap replications were both 200. The maximization was carried out using the constrOptim routine in R. given the complex form of the likelihood function, Nelder-Mead search algorithm was used for optimization and the maximum iteration was fixed at 2000.

The initial vectors for the constrained optimization was fixed at $\beta^{(1)} = (0.9, 0.5, \ldots, 0.5, 0.1)$ and $\beta^{(2)} = (0.9, 0.5, \ldots, 0.5, 0.1)$, where the vectors where of appropriate lengths. When the degree is equal to 2, the initial vector is chosen as $(0.9, 0.1)$. The initial value for the unconstrained optimization was done at $\hat{\beta}^{(1)} \otimes \hat{\beta}^{(2)}$ where $\hat{\beta}^{(1)}$ and $\hat{\beta}^{(2)}$ were the estimates obtained from the constrained estimation. The error standard deviation was chosen to be a small value $\sigma = 0.01$ and a bigger value $\sigma = 0.05$.

To study the finite sample properties of the power function we generate data from parametric alternatives that are within the parameter space. To this end we consider a class of alternatives that is in the same spirit of Tukey's one degrees of freedom tests. the alternatives are of the form

$$\begin{aligned} \beta_a &= \lambda\beta^0 + (1-\lambda)\beta^* \\ \beta^* &= \omega\beta^{(*1)} + (1-\omega)\beta^{(*2)} \\ \beta^0 &= \beta^{(01)} \otimes \beta^{(02)} \\ \beta^{(*1)} &= \beta^{(11)} \otimes \beta^{(12)} \\ \beta^{(*2)} &= \beta^{(21)} \otimes \beta^{(22)} \end{aligned} \tag{2.12}$$

where the coefficient vectors $\beta^{(01)}, \beta^{(11)}, \beta^{(21)}$ are $J_1 \times 1$ vectors and the vectors $\beta^{(02)}, \beta^{(12)}, \beta^{(22)}$ are $J_2 \times 1$ vectors. In addition we also consider alternatives of the

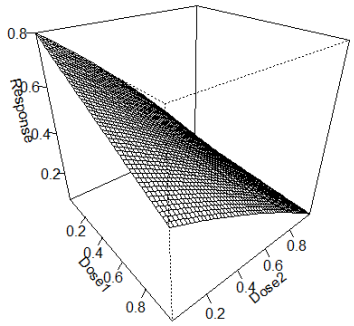
form where β^* in (2.12) is of the form

$$\beta^* = \beta^{(2)} \otimes \beta^{(1)}$$

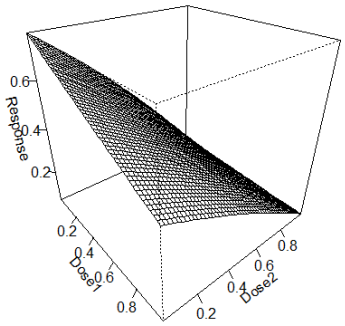
and $\beta^{(1)}$ and $\beta^{(2)}$ are two pre-determined coefficient vectors of length J_1 and J_2.

Note that the class is not one dimensional since $\beta^{(1)}$ and $\beta^{(2)}$ are unspecified nuisance parameters that are general vectors coefficients between 0 and 1. However, the single parameter $0 \leq \lambda \leq 1$ is the parameter that controls the alternative since $\lambda = 0$ results in a null value where as $\lambda > 0$ provides departure from the null toward the fixed alternative value β^*. The parameter ω in the first representation of β^* determines how close is the alternative to the null. For example for omega close to 0 or 1, β^* is again close to the null space since both ends of the segment over ω provides a factorized model. Thus, even though λ is high, the model reverts back to the null and thus the power is expected to be non-monotone when *omega* is close to zero or one. The particular choice of β^* highlights the difficulty in testing against a large alternative space and when the null itself is a high dimensional set. There are many directions from the initial null value where the path leads back toward the null set. In the second choice of β^* using the reverse factorization form $\beta^{(2)} \otimes \beta^{(1)}$, how close the alternative is to the null will be determined by the entries of the vectors $\beta^{(1)}$ and $\beta^{(2)}$ and the degree of difference in the orders J_1 and J_2. For example, regardless of the entries, when the degrees are equal, β^* belongs to the null set. Thus, the power behaves in non-monotonic fashion with respect to λ with maximum power expected when λ is far from the boundaries 0 and 1.

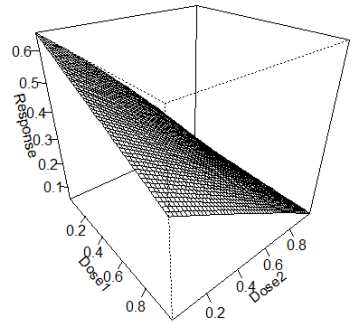
Example 1: The first example is a surface that is without any striking feature. The coefficient vectors constituting the components on the model (2.12) are $\beta_{01} = (0.9, .6, .1), \beta_{02} = (0.9, .5)$. The coefficient for the alternative are $\beta_{11} = (0.8, 0.7, .6), \beta_{12} = (0.9, 0.01)$ and $\beta_{21} = (.6, .3, .1), \beta_{22} = (.55, .45)$. Figure 2.2 shows the response surfaces for different choices of λ and ω.



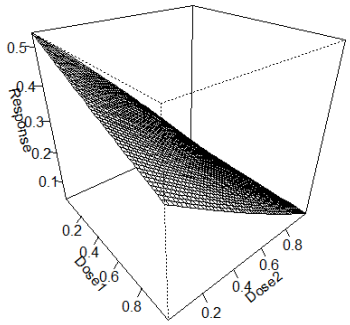
(a) $\lambda = 0$



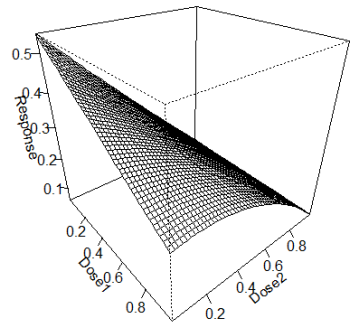
(b) $\lambda = 0.1, \omega = .5$



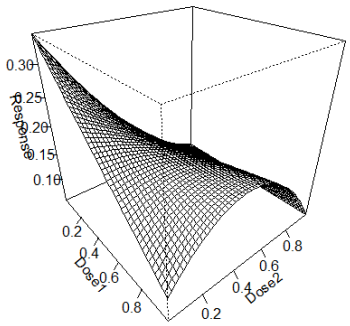
(c) $\lambda = 0.5, \omega = .5$



(d) $\lambda = 0.9, \omega = .5$



(e) $\lambda = 0.5, \omega = 0$



(f) $\lambda = 0.9, \omega = 0$

Figure 2.1: Surfaces corresponding to example 2 for different choices of λ and ω. The

Table 2.1 shows the power of the proposed test for different sample sizes and different values for λ. The value of ω is held fixed at 0.5. The power increases monotonically with λ and even for moderate sample size such as $n = 64$, the test has 100% detection at an alternative with $\lambda = 0.3$ the test maintains the nominal level for moderate to large sample sizes. The average time for a single run with sample size 100 was about 10 seconds in a standard laptop.

λ	n = 64	n = 100	n = 400
0.00	0.07	0.04	0.05
0.03	0.09	0.10	0.30
0.05	0.13	0.14	0.49
0.07	0.19	0.28	0.74
0.10	0.28	0.46	0.94
0.20	0.67	0.88	0.99
0.30	1.00	1.00	1.00
0.40	1.00	1.00	1.00
0.50	1.00	1.00	1.00
0.60	1.00	1.00	1.00
0.70	1.00	1.00	1.00
0.80	1.00	1.00	1.00
0.90	1.00	1.00	1.00
1.00	1.00	1.00	1.00

Table 2.1: Power of the proposed test for different values of λ in the alternative model described using the component vectors in Example 1 for different sample sizes and $\sigma = 0.01$ and $\omega = 0.5$.

Table 2.2 shows the power in the Example 1 set up when the error standard

deviation is slightly bigger and equal to $\sigma = 0.05$. The power is slightly affected by the increase in the error variance but still maintains the monotonic and rapid rise with increase in λ.

λ	n = 64	n = 100	n = 400
0.00	0.05	0.05	0.06
0.03	0.05	0.05	0.11
0.05	0.06	0.07	0.19
0.07	0.07	0.11	0.23
0.10	0.09	0.18	0.37
0.20	0.36	0.55	0.77
0.30	0.58	0.82	1.00
0.40	0.79	0.87	1.00
0.50	0.95	0.98	1.00
0.60	0.98	1.00	1.00
0.70	0.99	1.00	1.00
0.80	1.00	1.00	1.00
0.90	1.00	1.00	1.00
1.00	1.00	1.00	1.00

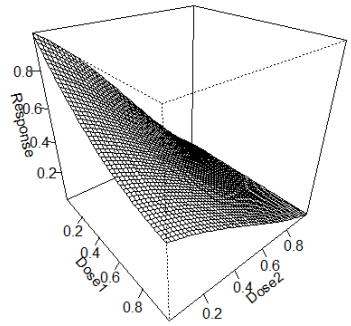
Table 2.2: Power of the proposed test for different values of λ in the alternative model described using the component vectors in Example 1 for different sample sizes and $\sigma = 0.05$ and $\omega = 0.5$.

Example 2: The second example is also a surface that is without any striking feature but does show slightly more curvature than those in Example 1. The coefficient vectors constituting the components on the model (2.12) are $\beta_{01} = (0.999, 0.8, 0.2, 0.001)$, $\beta_{02} = (0.999, 0.5, 0.45)$. The coefficient for the alternative are $\beta_{11} = (0.999, 0.9, 0.6, 0.001), \beta_{12} =$

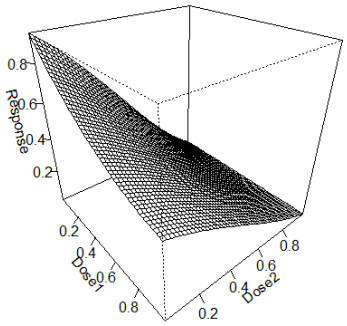
$(0.8, 0.1, 0.15)$ and $\beta_{21} = (.6, .55, .5, .1), \beta_{22} = (.5, .4, .3)$. Figure 2.2 shows the response surfaces for different choices of λ and ω.

From Figure 2.2 it seems that for a fixed ω, the departure from null is minimal even for $\lambda = 0.5$. When λ is large and close to one, then the surfaces are quite different. We know that in the case when $\lambda = 0.9$ and $\omega = .5$ the surface is not of a Bliss independence model. Table 2.3 shows the power of the proposed test for different values of λ when $\omega = 0.5$ and the error standard deviation is small and equal to $\sigma = 0.01$. The test maintains the size very well, even for smaller sample sizes. The values in Table 2.3 reaffirms the nature of departure of the parametric alternative model from the null in terms of the single parameter λ. The surfaces for $\lambda = 0$ and $\lambda = 0.9$ are visually different overall and for a small value of σ the power of the test reaches almost one at $\lambda = 0.9$ for larger sample size such as $n = 400$. However for smaller sample sizes, the observed values where the differences are significant are sparse and are unable to distinguish the surfaces based on the differences over the observed grid. For a sample of size $n = 100$ the power is 50-60% for large values of λ. The lack of power is due to the nature of the nonparametric null and alternatives. Even though the surface with $\lambda = 0.9$ and $\omega = 0.5$ is visually different from the null model with $\lambda = 0$ and also it is not a Bliss independence surface, it is not clear how close it is to "a" Bliss independence surface which need not be the one with $\lambda = 0$.

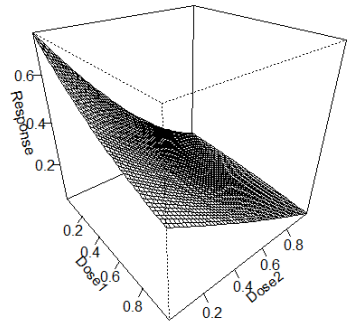
Table 2.4 shows the power for a larger error variance with $\sigma = 0.05$. The increase in the error variance has a significant impact on the power of the test. Even at $n = 400$, the power reaches only about 30% when testing an alternative with $\lambda = 0.9$. This again highlights the problem with the large alternative space and the null space. The surface at $\lambda = 0.9$, albeit visually different form the null model with $\lambda = 0$, may very well be close to another factorizable Bliss independent surface. Even though the error standard deviation is not very big in magnitude, the possible closeness of



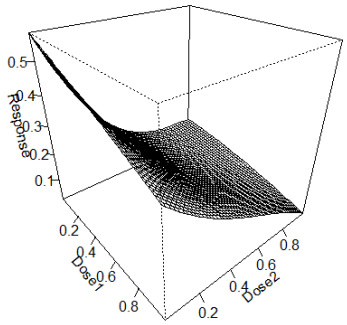
(a) $\lambda = 0$



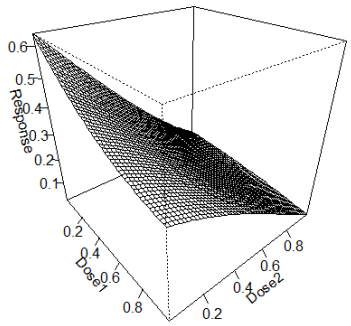
(b) $\lambda = 0.1, \omega = .5$



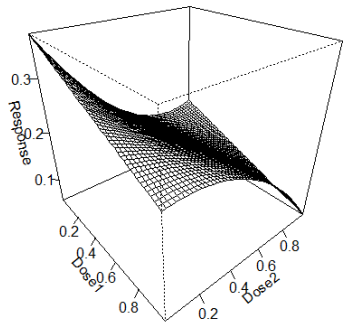
(c) $\lambda = 0.5, \omega = .5$



(d) $\lambda = 0.9, \omega = .5$



(e) $\lambda = 0.5, \omega = 0$



(f) $\lambda = 0.9, \omega = 0$

Figure 2.2: Surfaces corresponding to example 2 for different choices of λ and ω. The

λ	n = 64	n = 100	n = 400
0.00	0.04	0.05	0.05
0.03	0.05	0.06	0.09
0.05	0.07	0.08	0.10
0.07	0.08	0.10	0.14
0.10	0.10	0.13	0.15
0.20	0.13	0.17	0.24
0.30	0.18	0.25	0.23
0.40	0.22	0.29	0.32
0.50	0.24	0.31	0.42
0.60	0.28	0.32	0.47
0.70	0.34	0.37	0.51
0.80	0.37	0.43	0.66
0.90	0.40	0.52	0.81
1.00	0.45	0.65	0.99

Table 2.3: Power of the proposed test for different values of λ in the alternative model described using the component vectors in Example 2 for different sample sizes and $\sigma = 0.01$ and $\omega = 0.5$.

the alternative to a null surface compounded with the increased variability of the observed surface greatly limits the capability of the test to reject the alternative at $\lambda = 0.9$. We feel that the constant error variance assumption is too severe particularly because when the survival proportion is close to one or zero, the perturbation due to the error is quite significant and has a serious impact on estimation.

Next we consider the case when $\omega = 0$ for alternative values in Example 2. for $\omega = 0$, β^* also belongs to the null region. Thus, the power is expected to non-

λ	n = 64	n = 100	n = 400
0.00	0.02	0.05	0.05
0.03	0.03	0.06	0.05
0.05	0.03	0.06	0.07
0.07	0.05	0.08	0.08
0.10	0.05	0.07	0.08
0.20	0.06	0.07	0.10
0.30	0.08	0.08	0.12
0.40	0.08	0.09	0.14
0.50	0.07	0.10	0.17
0.60	0.08	0.11	0.20
0.70	0.08	0.11	0.22
0.80	0.09	0.12	0.25
0.90	0.09	0.13	0.29
1.00	0.08	0.15	0.32

Table 2.4: Power of the proposed test for different values of λ in the alternative model described using the component vectors in Example 2 for different sample sizes and $\sigma = 0.05$ and $\omega = 0.5$.

monotonic over the segment parametrized by λ. At either end, $\lambda = 0$ and $\lambda = 1$, the parameter belongs to the null and hence the power is expected to be around the nominal level of 5%. From Table 2.5 we see that indeed the power is around 5% for either $\lambda = 0$ or $\lambda = 1$. The power reaches maximum of 1 in the middle of the segment around $\lambda = 0.5$. Even for $n = 64$ the test has almost 100% power at $\lambda = 0.5$.

Table 2.6 shows the power properties for $\sigma = 0.05$. Again the increased variance has a drastic effect on the power. The maximum power for $n = 400$ is only about

λ	n = 64	n = 100	n = 400
0.00	0.04	0.05	0.05
0.05	0.11	0.19	0.35
0.10	0.19	0.32	0.92
0.20	0.07	0.11	0.99
0.30	0.22	0.99	1.00
0.40	0.36	1.00	1.00
0.50	0.98	1.00	1.00
0.60	0.56	1.00	1.00
0.70	0.44	0.98	1.00
0.80	0.26	0.50	1.00
0.90	0.21	0.23	0.95
0.95	0.13	0.13	0.31
1.00	0.05	0.06	0.06

Table 2.5: Power of the proposed test for different values of λ in the alternative model described using the component vectors in Example 2 for different sample sizes and $\sigma = 0.01$ and $\omega = 0$.

24% for $\lambda = 0.5$. The argument about increased variability having a dire effect on estimation particularly due to the distortion around the edges still holds. Probably a model with error variance function of the response, e.g. $\sigma^2 = \tau_0 + \tau_1\mu(1-\mu)$ for some small background τ_0 value and some τ_1 value is more appropriate. Such models are topics of future research.

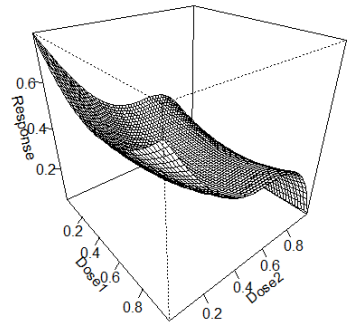
Example 3: The third example is a surface that shows distinct features such as flat regions and regions of sharp decline. The coefficient vectors constituting the components on the model (2.12) are $\beta_{01} = (0.9, 0.8, .5, .6, 0.01), \beta_{02} = (0.9, 0.4, 0.3)$. The co-

λ	n = 64	n = 100	n = 400
0.00	0.02	0.04	0.06
0.05	0.04	0.05	0.06
0.10	0.04	0.07	0.07
0.20	0.04	0.11	0.11
0.30	0.08	0.10	0.17
0.40	0.09	0.11	0.19
0.50	0.13	0.13	0.24
0.60	0.11	0.12	0.16
0.70	0.11	0.09	0.15
0.80	0.07	0.06	0.12
0.90	0.08	0.07	0.08
0.95	0.05	0.07	0.08
1.00	0.03	0.06	0.07

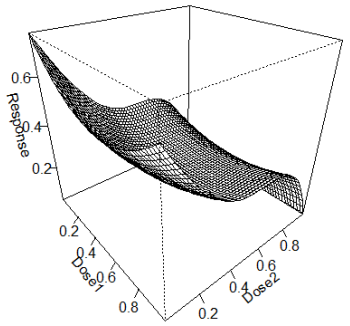
Table 2.6: Power of the proposed test for different values of λ in the alternative model described using the component vectors in Example 2 for different sample sizes and $\sigma = 0.05$ and $\omega = 0.5$.

efficient for the alternative are $\beta_{11} = (0.999, 0.9, 0.6, .002, 0.001), \beta_{12} = (0.8, 0.1, 0.15)$ and $\beta_{21} = (.6, .55, .5, .45, .1), \beta_{22} = (.5, .4, .3)$. Figure 2.2 shows the response surfaces for different choices of λ and ω.

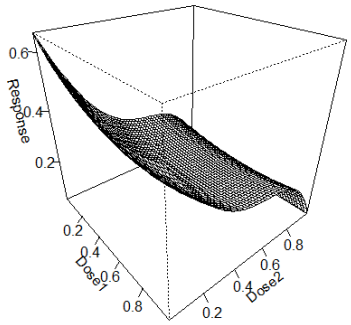
Table 2.7 shows the power for Example 3 for different values of sample sizes and λ. The value of ω is 0.5 and $\sigma = 0.01$. Given the surface has more features, the power is slightly less than that in Example 2 but still adequate even for smaller sample sizes. The size is again preserved very well across sample sizes. Table 2.8 provides the power for the same example for $\omega = 0$. The power is near the nominal level at both ends



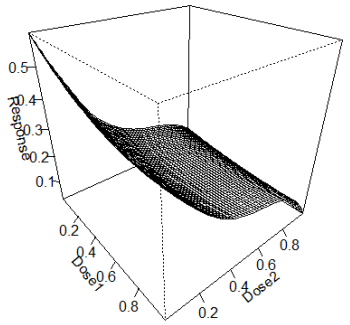
(a) $\lambda = 0, \omega = .5$



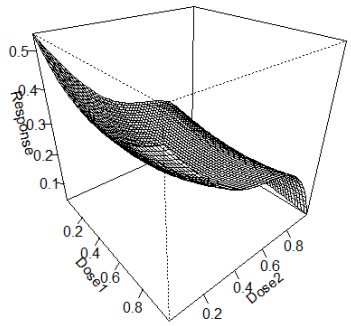
(b) $\lambda = 0, \omega = .5$



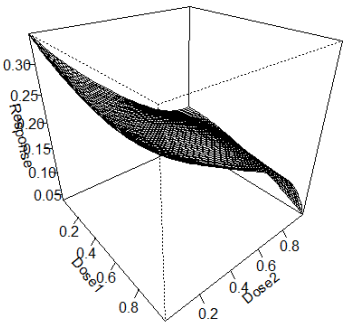
(c) $\lambda = 0, \omega = .5$



(d) $\lambda = 0, \omega = .5$



(e) $\lambda = 0, \omega = .5$



(f) $\lambda = 0, \omega = .5$

Figure 2.3: Surfaces corresponding to example 4

of the line segment parametrized by λ. However, for this example, the power hardly rises in the middle of the segment. It could be that the combination of β^0 and $\beta^{(02)}$ gives rise to vectors that are nearly factorizable.

λ	n = 64	n = 100	n = 400
0.00	0.05	0.05	0.05
0.05	0.07	0.08	0.12
0.10	0.13	0.15	0.21
0.20	0.17	0.22	0.30
0.30	0.21	0.31	0.69
0.40	0.37	0.57	0.94
0.50	0.54	0.77	0.99
0.70	0.89	0.98	1.00
1.00	1.00	1.00	1.00

Table 2.7: Power of the proposed test for different values of λ in the alternative model described using the component vectors in Example 3 for different sample sizes and $\sigma = 0.01$ and $\omega = 0.5$.

2.3 Model Selection

To model the dose response using Bernstein polynomials, we have used the uniform convergence property of the Bernstein. Thus, we have used a large enough degree such that the approximation of the true model is adequate. Of course in finite sample situation there is a severe complexity cost for using large number of basis function. Thus, to achieve reasonable approximation without sacrificing computational and statistical efficiency we propose using conventional model selection procedure to determine the degree for each sample.

λ	n = 64	n = 100	n = 400
0.00	0.05	0.05	0.05
0.05	0.07	0.07	0.07
0.10	0.08	0.09	0.11
0.20	0.09	0.12	0.14
0.30	0.10	0.15	0.17
0.40	0.12	0.16	0.19
0.50	0.12	0.18	0.22
0.70	0.11	0.14	0.18
1.00	0.06	0.06	0.05

Table 2.8: Power of the proposed test for different values of λ in the alternative model described using the component vectors in Example 3 for different sample sizes and $\sigma = 0.01$ and $\omega = 0$.

We perform limited simulation study to first check the performance of the model selection procedure in terms of how close the selected model is to the true model in an average. We create data under the null model using two dimensional Bernstein polynomials and performed the simulation experiment by varying sample size, the degrees of Bernstein polynomials, the coefficients of Bernstein polynomials for a given degree and also the relationship between the degree in the two direction of dose space, e.g, equal degrees of Bernstein polynomials in either coordinate or different degrees of Bernstein polynomials. The rest of procedure to perform the test is same as in the previous section.

We vary sample size 100, 144, 225, 324 or 400. For degree 3 Bernstein polynomial, we use coefficient vector c(.99, .66, .1) for simple cell survival curve, use coefficient vector c(.99, .66, .88) for featured cell survival curve. For degree 4 Bernstein poly-

nomial, we use coefficient vector c(.99, .66, .44, .1) for simple cell survival curve, use coefficient vector c(.99, .66, .88, .1) for featured cell survival curve. For degree 5 Bernstein polynomial, we use coefficient vector c(.99, .99, .66, .44, .1) for simple cell survival curve, use coefficient vector c(.99, .66, .88, .44, .1) for featured cell survival curve. For degree 6 Bernstein polynomial, we use coefficient vector c(.99, .99, .80, .66, .44, .1) for simple cell survival curve, use coefficient vector c(.99, .66, .88, .66, .44, .1) for featured cell survival curve. For degree 7 Bernstein polynomial, we use coefficient vector c(.99, .99, .80, .70, .55, .44, .1) for simple cell survival curve, use coefficient vector c(.99, .66, .88, .70, .55, .44, .1) for featured cell survival curve. For degree 8 Bernstein polynomial, we use coefficient vector c(.99, .99, .80, .70, .55, .44, .22, .1) for simple cell survival curve, use coefficient vector c(.99, .66, .88, .70, .55, .44, .22, .1) for featured cell survival curve.

- Equal dimension of two dimensional Bernstein data, with the true dimension (3, 3), the possible fitted dimension are $\{J_1, J_2 \in (2, \ldots, 5)\}$.
- Equal dimension of two dimensional Bernstein data, with the true dimension (4, 4), the possible fitted dimension are $\{J_1, J_2 \in (2, \ldots, 6)\}$.
- Equal dimension of two dimensional Bernstein data, with the true dimension (5, 5), the possible fitted dimension are $\{J_1, J_2 \in (2, \ldots, 7)\}$.
- Equal dimension of two dimensional Bernstein data, with the true dimension (6, 6), the possible fitted dimension are $\{J_1, J_2 \in (2, \ldots, 8)\}$.
- Equal dimension of two dimensional Bernstein data, with the true dimension (7, 7), the possible fitted dimension are $\{J_1, J_2 \in (2, \ldots, 9)\}$.
- Equal dimension of two dimensional Bernstein data, with the true dimension (8, 8), the possible fitted dimension are $\{J_1, J_2 \in (2 \ldots, 10)\}$.

- Unequal dimension of two dimensional Bernstein data, with the true dimension (3, 5), the possible fitted dimension are $\{J_1, J_2 \in (2, \ldots, 7)\}$.
- Unequal dimension of two dimensional Bernstein data, with the true dimension (4, 6), the possible fitted dimension are $\{J_1, J_2 \in (2, \ldots, 8)\}$.
- Unequal dimension of two dimensional Bernstein data, with the true dimension (5, 7), the possible fitted dimension are $\{J_1, J_2 \in (2, \ldots, 9)\}$.
- Unequal dimension of two dimensional Bernstein data, with the true dimension (6, 8), the possible fitted dimension are $\{J_1, J_2 \in (2, \ldots, 10)\}$.

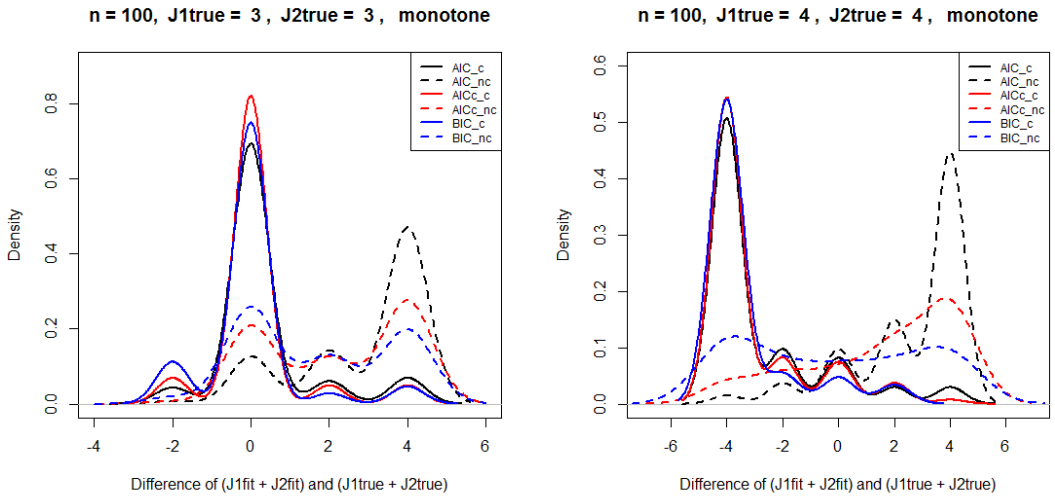
The best model is then selected based on smallest values of AIC (Akaike information criterion), BIC (Bayesian information criterion) or $AICc$ (AIC with a correction for finite sample sizes). AIC is defined as $2k - 2 * ln(L)$, here L is the maximum value of the likelihood function for the model, k is the number of estimated parameters in the model. BIC is defined as $-2lnL + +k * ln(n)$. L is the maximized value of the likelihood function of the model, k is the number of free parameters to be estimated, and n is the sample size. $AICc$ is AIC with a correction for finite sample sizes. The formula for $AICc$ is $AICc = AIC + \frac{2k*(k+1)}{n-k-1}$, n denotes the sample size and k denotes the number of parameters. For each set of simulated parameter setting, we did 100 Monte Carlo replication.While fitting the model, we can use constrained or unconstrained Bernstein polynomials to select the best degree. We study both scenarios. We study the properties of the model selection criteria by examining the differences of the fitted degrees of Bernstein polynomial and the true degrees of Bernstein polynomial. If the true degrees are (J_1, J_2) and the fitted degrees are $J_{1,fit}, J_{2,fit}$ then we look at the distribution of the qunatity $(J_{1,fit} - J_1) + (J_{2,fit} - J2)$ over the Monte Carlo replications.

Figure 2.4-2.7 shows the distribution of Δ for the different scenarios investigated. There are several points that one can make based on our experiment to find whether the model selection procedure works in the present set up. The data were generated under the null factorized model and hence for every case the fitted degree under the constrained model is closer to the true degree than those for the unconstrained model. In general, the BIC criterion does slightly better and has sharper peak around the true dimension. As expected, with larger sample sizes, e.g 400, the fitted degrees get closer to the true degrees. The model selection seems to be targeting the essential features of the data. For example, in Figure 2.4(a), each marginal curves are simple degree 3 curves and the constrained fit seems to essentially pick that up. Figure 2.4(b)- (d), because some of the coefficients are small, the fitted model has a tendency to use lower order polynomials. Whereas, in Figure 2.5, the fit for dimensions equal to 4 is better since the curves have mode features. The story is similar for the larger sample size. However, the model selection seems to be sensitive to the assumption of Bliss independence since the unconstrained model choice seemingly differs from the constrained model choice even when the null model is used. More investigation is needed to fully understand the optimal use of model selection procedure in conjunction with the testing procedure.

2.4 Discussion

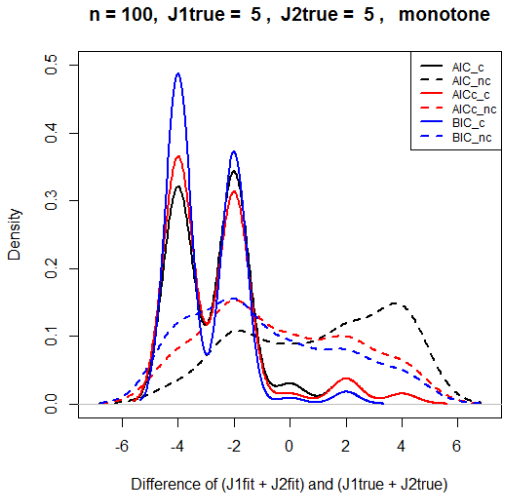
We have proposed a fully nonparametric procedure for testing Bliss independence in two agent dose response models. The procedure seems to have the expected performance of a nonparametric testing procedure. The size of the test is well maintained across a variety of scenarios. There are several factors that affect the power properties of the test and the sensitivity of the test to these factors needs to be investigated more thoroughly. Also, while the empirical evidence points to consistency of the test,

theoretical justification needs to be established to show that at any given alternative, the power of the test reaches one as sample size goes to infinity. The complexity of the procedure depends on the size and spacings of observation grid as well as the grid used to impose the constraints of monotonicity. The true relationship between these complexities and the rate of convergence of the null distribution of the test needs to be investigated and is a topic of future research.

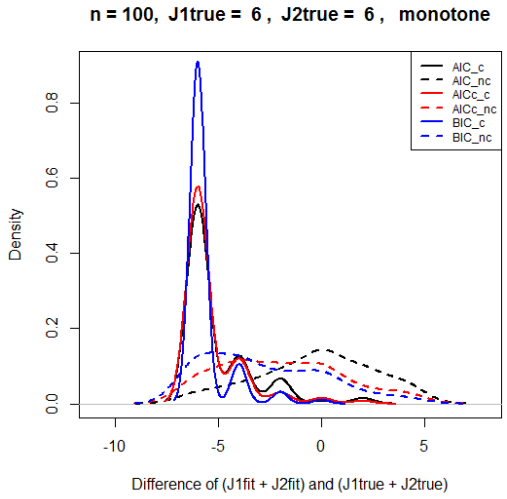


(a)

(b)

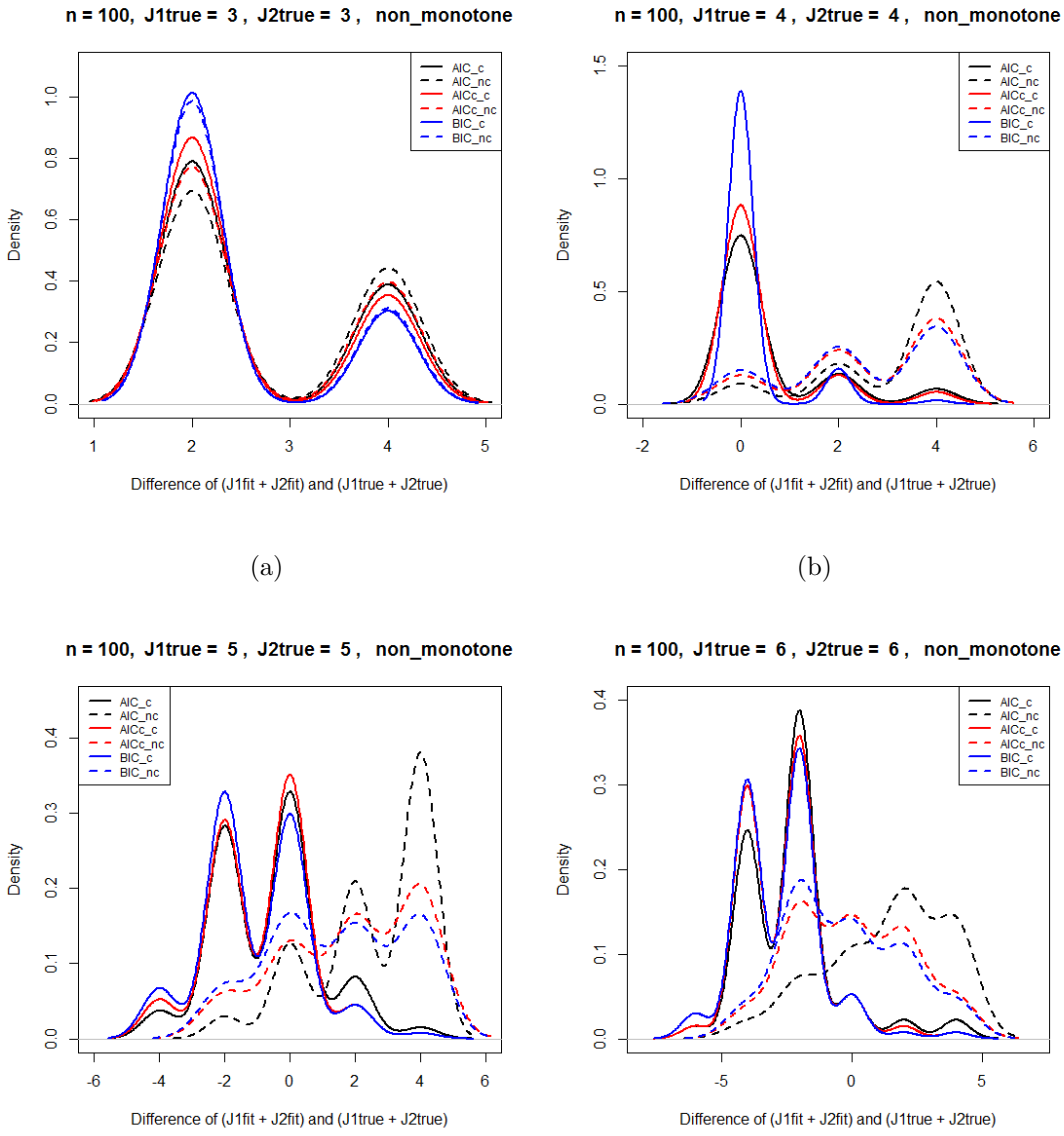


(c)



(d)

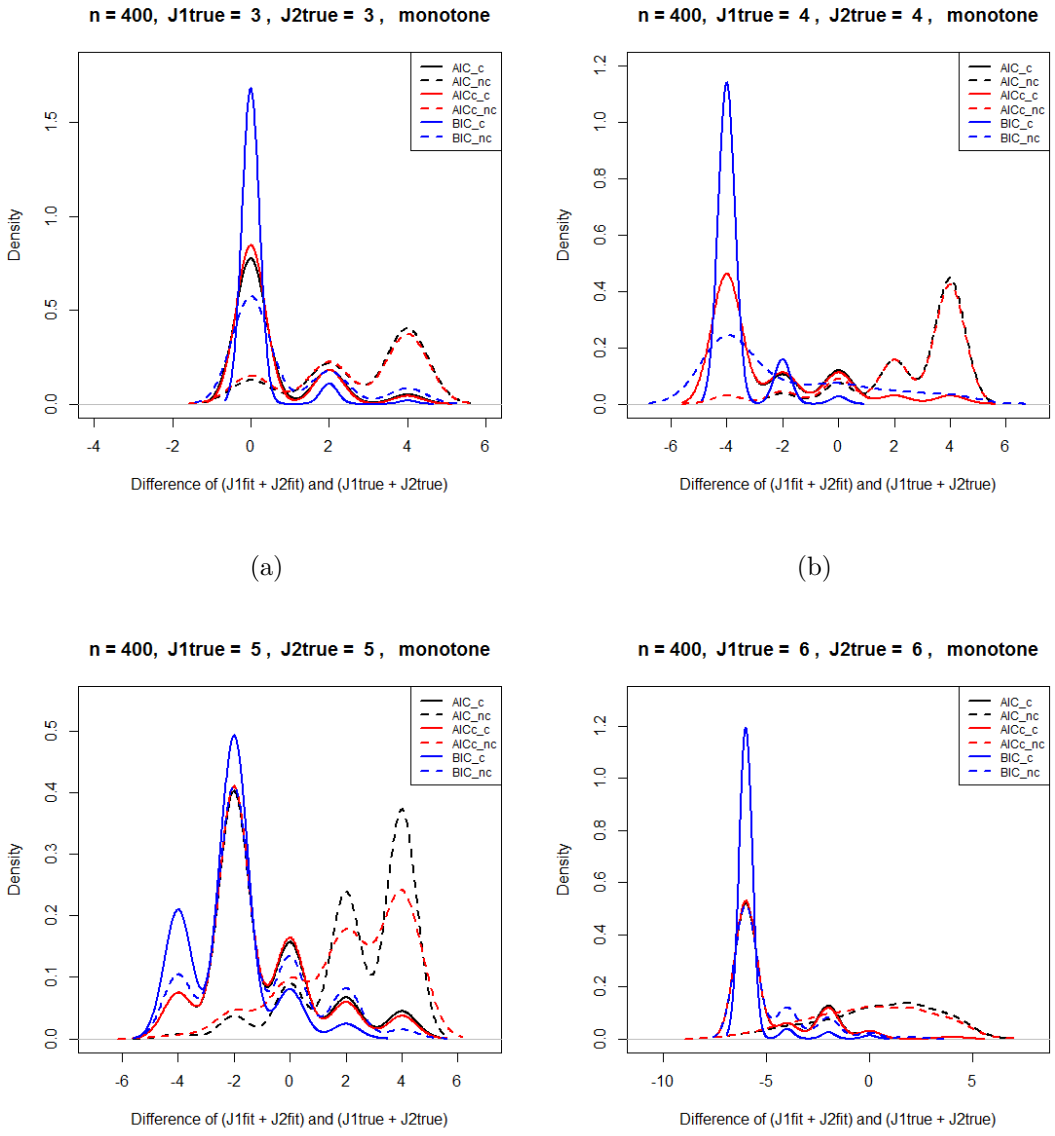
Figure 2.4: Density plot of difference of fitted dimension and true dimension of n=100, equal dimension and simple curve. Solid line represents the constrained fit and dash line represents unconstrained fit. Black line the model selection uses AIC criteria, red line the model selection uses $AICc$ criteria and blue line the model selection uses BIC criteria.



(a) (b)

(c) (d)

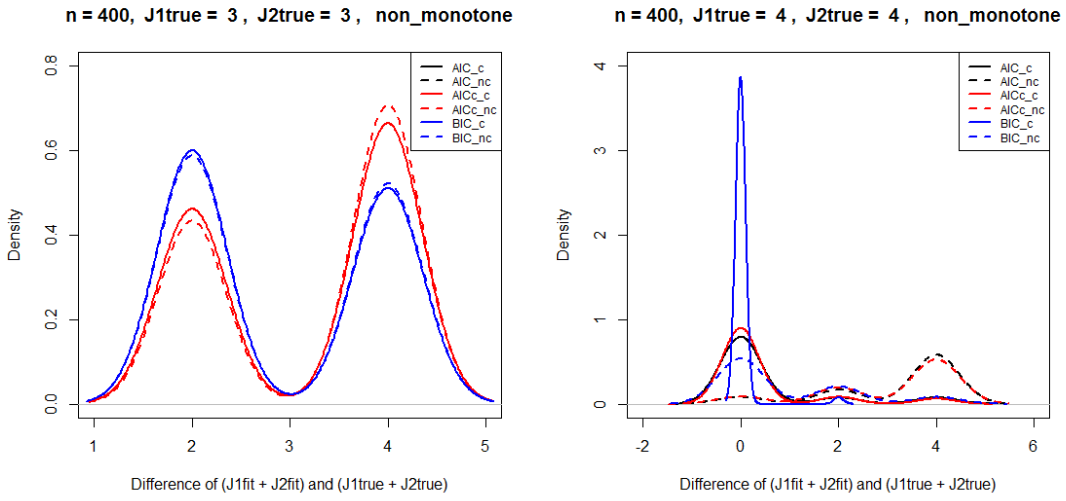
Figure 2.5: Density plot of difference of fitted dimension and true dimension of n=100, equal dimension and curve with feature. Solid line represents the constrained fit and dash line represents unconstrained fit. Black line the model selection uses AIC criteria, red line the model selection uses $AICc$ criteria and blue line the model selection uses BIC criteria.



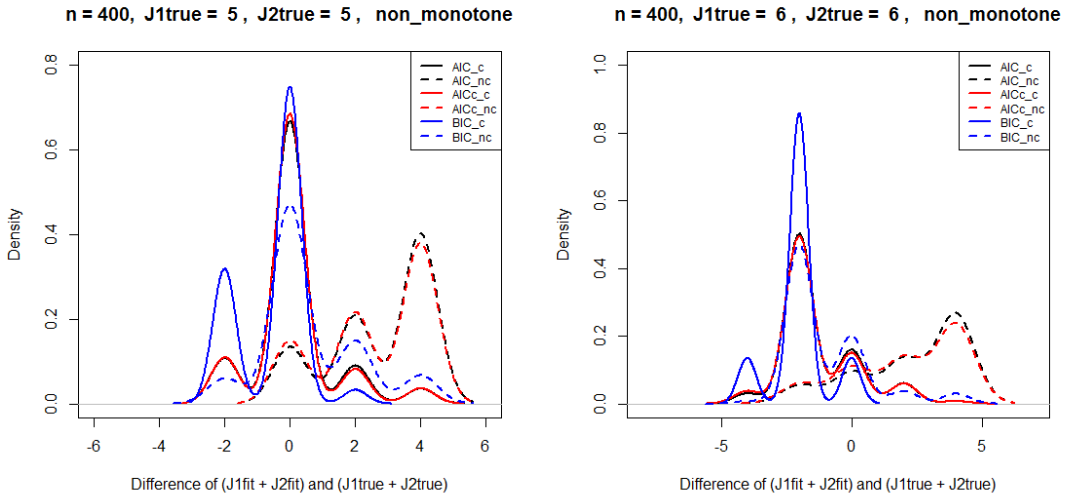
(a) (b)

(c) (d)

Figure 2.6: Density plot of difference of fitted dimension and true dimension of n=400, equal dimension and simple curve. Solid line represents the constrained fit and dash line represents unconstrained fit. Black line the model selection uses AIC criteria, red line the model selection uses $AICc$ criteria and blue line the model selection uses BIC criteria.



(a) (b)



(c) (d)

Figure 2.7: Density plot of difference of fitted dimension and true dimension of n=400, equal dimension and curve with feature. Solid line represents the constrained fit and dash line represents unconstrained fit. Black line the model selection uses AIC criteria, red line the model selection uses $AICc$ criteria and blue line the model selection uses BIC criteria.

Chapter 3

Nonparametric Test for Loewe Additivity

3.1 Introduction

As mentioned in the introduction, the two most common forms of two-agent interaction are bliss independence and Loewe additivity. In this chapter we develop nonparametric and semi-parametric tests for testing Loewe additivity. the motivation for investigating nonparametric test is similar to those given in the Bliss independence case. The response surface could have shapes that are not well approximated by conventional models and also the departure of the alternative models from a null reference model may not be adequately characterized by a few parameters. Along with Bliss independence, Loewe additive models as reference models for dose response surface has a long history. However, the models are almost always parametric. Moreover the emphasis has been on response surface modeling rather than building rigorous statistical tests for departure from Loewe additivity. This investigation attempts to fill the gap.

3.1.1 Loewe Reference Model for Zero Interaction

The study of additivity in the combined effect of two or more agents often is done via the interaction index. For a two agent situation, the interaction index (Berenbaum 1977) (at a particular effect E) can be defined as

$$I = \frac{x_1}{X_1} + \frac{x_2}{X_2} \tag{3.1}$$

where (x_1, x_2) are the doses of the two agents that produced an effect of E when the agents are combined and X_1 and X_2 are the marginal doses of the first and the second agent, respectively, that are needed to produce the same effect. The premise of te interaction index is that the left hand side should be independent of (d_1, d_2) if the agents have zero interaction or in other words they act similarly. This quantifies the ideas of additive effect studied through isobolograms, the lines of equal activity. The *isoboles* are defined as

$$isbl(E) = \{(x_1, x_2) : \mu(x_1, x_2) = E\} \tag{3.2}$$

where $\mu(x_1, x_2)$ denote the effect when the agents are combined at dose x_1 and x_2. The isoboles provide a simple graphical representation that conveys the idea of interaction without any more involved mathematical concepts. It is argued that if the agents have similar mechanism then their effects will be additive and the isobole should be a straight line. The isoboles were introduced by Fraser (1870, 1872)) to study antagonistic effect between agents. Later they were extended to the study of synergy by Loewe and Muischnek (1926) and Loewe (1953). Loewe argued that a curved isobole would indicate deviations from zero interaction. The idea was later quantified by using a combined index (CI) equal to I in (3.1). For agents whose effects interact,

the combined index (Chou and Talalay 1983) would show the nature of interaction as

$$CI \begin{cases} < 1 & \text{synergy} \\ = 1 & \text{additive} \\ > 1 & \text{antagonism} \end{cases} \tag{3.3}$$

The combined index is used is conjunction with parametric response surface methods to perform statistical analysis of two-agent interaction. One of the most common model that parametrizes the departure from zero interaction in the combined index using a single parameter is the Greco model (Greco et. al 1995). Specifically, The Greco model uses the Hill model (**REF**) as the marginal dose response model for the two agents and combines them through the interaction index. In order to describe the Greco model we describe the Hill model first. The Hill model for an effect E at a dose D is

$$E = E_0 + \frac{E_{max} - E_0}{1 + (\frac{X}{ED_{50}})^{-\gamma}} \tag{3.4}$$

where $E_0, E_{\max}$ are the background and maximum effects, respectively, and ED_{50} is the dose at which the effect is 50% of the max effect and γ is the Hill coefficient which controls the shape and the asymptote of the sigmoidal model. From the Hill model, the dose needed for reaching an effect of E is $X = ED_{50}\Big(\frac{E-E_0}{E_{max}-E}\Big)^{-1/\gamma}$. The Greco at.al 1995) model using the Hill model as the marginal response model is a solution to

$$1 = \frac{x_1}{ED_{50,1}\Big(\frac{E-E_0}{E_{max}-E}\Big)^{1/\gamma_1}} + \frac{x_2}{ED_{50,2}\Big(\frac{E-E_0}{E_{max}-E}\Big)^{1/\gamma_2}} + \frac{\tau x_1 x_2}{ED_{50,1}ED_{50,2}\Big(\frac{E-E_0}{E_{max}-E}\Big)^{1/2\gamma_1+1/2\gamma_2}}. \tag{3.5}$$

The model is a popular model but the it is still a parametric model and flexibility is limited by the specific parametric form. IN particular, the marginal model need not be adequately represented by a Hill model. Moreover, the departure from the null interaction model need not be of the specific form given by the model. Figure 3.1

shows the isobols for a typical Greco model for two different values of τ. The red line shows the isobol for $\tau = 0$ and the black line shos that for $\tau = -0.5$, a value representing antagonism.

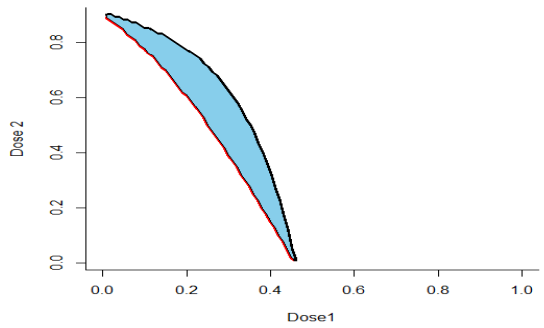


Figure 3.1: Loewe additive line and synnergistic isobol under Greco model

We will use the Greco model due to its simplicity and interpretability as our building block and generalize it to have nonparametric features.

3.1.2 A nonparametric Loewe additive reference model

The marginal models in the Greco model or any analogous parametric model are typically a Hill model or one of the variants of Hill model. The models are always low dimensional parametric model. While the models do capture many of the essential features they rely on the sigmoidal shape of the dose response. There is no scientific reason for the dose response, albeit monotonic, to have a sigmoidal shape. In particular, the dose response may show multiple regions of flat response or sharp rise. Thus, we let the marginal model of dose response to be nonparametric and model it using Bernstein polynomials. However, the linear zero interaction model is inherently a model over the positive half plane, since it is not possible for the linear equal activity lines to hold over a range of effects over a compact dose space except for the trivial linear dose response form. Therefore, it is necessary to rescale the compact dose set

using a scaling function, which could be any monotone function mapping [0, 1] to $[0, \infty)$. Let $L(\cdot)$ be such a scaling function. For our implementation we will choose the function to be $L(x) = -log(1 - x)$.

The zero interaction reference model for marginal response function $\mu_1(X_1)$ and $\mu_2(X_2)$ and joint response $\mu(x_1, x_2)$ over a dose space $\mathcal{D} = \mathbb{R}^+ \times \mathbb{R}^+$ can be written as

$$1 = \frac{x_1}{\mu_1^{-1}(E)} + \frac{x_2}{\mu_2^{-1}(E)}$$

where $\mu_1^{-1}(E)$ is the inverse function giving the dose at which the effect is E when agent one is used by itself, $\mu_2^{-1}(E)$ is the inverse function giving the dose at which the effect is E when agent two is used by itself, and (x_1, x_2) are dose on the isobole $\{(x_1, x_2) : \mu(x_1, x_2) = E\}$.

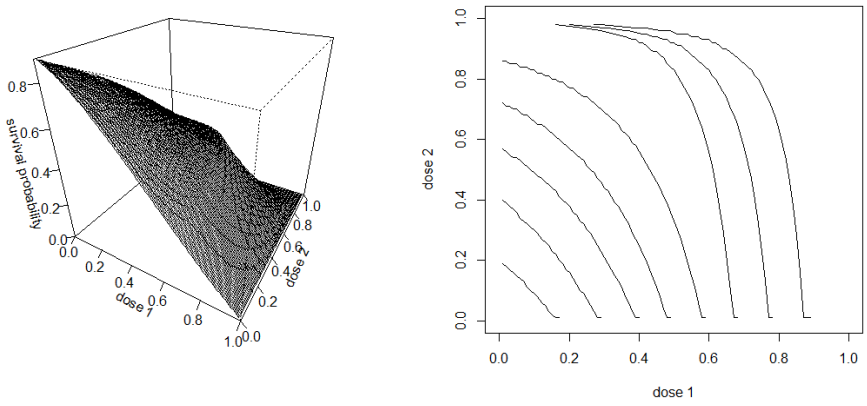
For our application, we have to generalize the idea of no interaction measured via linear isoboles over the positive half plane to a definition of zero interaction over the unit square. Let the marginal dose response models will be $\mu_j(X_j)$ and let the joint dose response function be $\mu(x_1, x_2)$. Then the zero interaction interaction model over the unit square is defined as

$$1 = \frac{L(d_1)}{L(\mu_1^{-1}(E))} + \frac{L(d_2)}{L(\mu_2^{-1}(E))} \tag{3.6}$$

where $\{(x_1, x_2) \in [0, 1]^2 : \mu(x_1, x_2) = E\}$. The fact that isoboles are straight lines in the no interaction model is not necessarily true even in the unbounded dose space. For example, Grabovsky and Tallarida (2004) report that when one of the agent is partial agonist then the isoboles can be curved. However, in the bounded dose space the idea of straight line isoboles severely limits the definition of no interaction. That is why we use a rescaling function and define the straight line Loewe additivity in the transformed space. The next two figures show the response surface for models generated using (3.6) in the following two examples:

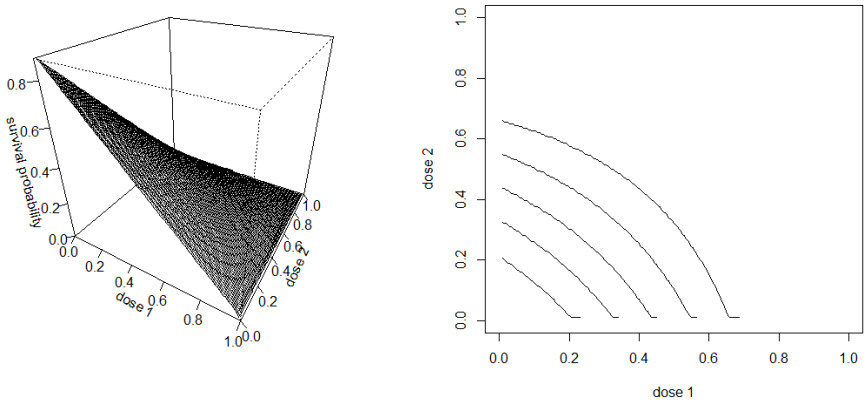
Example 4: The first example has marginal models described by Bernstein polynomials of degree 2 and 2. The coefficients are given by $\beta^{(1)} = (0.9, 0.01)$ and $\beta^{(2)} = (0.9, 0.1)$. The model is simple and the surface does not show any special features. The surface satisfy the monotonicity and range restriction. In Figure3.3 the isoboles are parallel although they are curved due to the boundedness of the dose space.

Example 5: The marginal models are given by Bernstein polynomials of order 5 and 3. The coefficients are $\beta^{(1)} = (0.9, 0.8, 0.5, 0.6, 0.01)$ and $\beta^{(2)} = (0.9, 0.4, 0.3)$. Note that the coefficients are non-monotonic, but the marginal response models satisfy monotonicity properties. For this example, the Loewe additivity lines or isoboles (in the bounded space) are not necessarily parallel. This is due to the shape of the surface which exhibits a significant shoulder like feature. Figure 3.2 shows the surface and the isoboles.



(a) surface (b) isoboles

Figure 3.2: A Bernstein based Loewe additive model surface and associated isoboles at level $\Gamma \in \{0.1, 0.2, \ldots, 0.9\}$



(a) surface (b) isoboles

Figure 3.3: A Bernstein based Loewe additive model surface and associated isoboles at level $\Gamma \in \{0.1, 0.2, \ldots, 0.9\}$

Throughout the chapter we will assume that agents are Loewe additive if they satisfy (3.6). For modeling purpose we assume that μ_j's do not belong to any specific parametric family and model them using Bernstein polynomials with monotonicity and range constraints. Thus, we assume $\mu_j(X_j)$ is well approximated by some $B_{J_j}(X_j)$ where $B_J(X)$ is the Bernstein function defined in (2.7). Let the inverse dose response functions be $B_{J_j}^{-1}(E)$. As before, we let the degree to be an increasing function of the sample size and we construct a sieve by choosing a grid over which the polynomials are constrained to be monotonic. The grid is allowed to become dense in the dose space with increasing sample size.

3.2 A semiparametric test for Loewe additivity

To test for additivity one could allow the alternative class to be fully nonparametric. In general, any class of functions that includes the linear function would work.

However, before investigating the fully nonparametric alternative, we first investigate lower dimensional departures from the zero interaction model. Specifically, we could entertain a Tukey one degree of freedom type test by extending the null interaction model using a model similar to Greco et. al (1995). One could specify and one degree interaction model based on a single interaction parameter

$$1 = \frac{L(x_1)}{L(\mu_1^{-1}(E))} + \frac{L(x_2)}{L(\mu_2^{-1}(E))} + \frac{\tau L(x_1)L(x_2)}{\sqrt{L(\mu_1^{-1}(E))L(\mu_2^{-1}(E))}}$$

However, as in the Greco model, a single degree interaction term may not be sufficient for modeling many forms of interaction surfaces, particularly when some parts of the dose space exhibits positive interaction while other parts show negative interaction. nevertheless, a single degree interaction term can used in a semi-parametric framework to test for Loewe additivity against directional alternatives. The main advantage being higher power for testing alternatives belonging to the class while having enough flexibility in modeling the marginal response functions.

To develop the test once again we need a sieve to overcome the increasing complexity of the problem. To construct the sieve we use the Bernstein approximation to the marginal models. Let the marginal models we parametrized by vectors $\beta^{(1)}$ and $\beta^{(2)}$ that restrict the model to the sieve (2.11) given in the previous chapter. To be precise the sieve for the entire parameter set $(\beta^{(1)}, \beta^{(2)}, \tau)$ is $\Theta^G_{0,\beta} \times \mathbb{R}$ where the interaction parameter τ is allowed to have any real value.

To test for interaction we test $H_0 : \tau = 0$ versus $H_1 : \tau \neq 0$. To test for directional departures we could use one sided alternatives. We discuss directional tests in more details in the next chapter. We use a likelihood ratio test to test the nested models, one without the restriction $\tau = 0$ and one with the restriction. The form of the likelihood ratio test is

$$\Lambda(data) = \frac{\sup_{(\beta^{(1)},\beta^{(2)}) \in \Theta^G_{0,\beta} \bigcap \{\tau = 0\}} L(\beta^{(1)}, \beta^{(2)}, \tau)}{\sup_{(\beta^{(1)},\beta^{(2)}) \in \Theta^G_{0,\beta}} L(\beta^{(1)}, \beta^{(2)}, \tau)} \tag{3.7}$$

The critical region for a test with nominal level α will be

$$\mathcal{R}_\alpha = \{Data : \Lambda(data) < c_\alpha\}$$

where c_α is the lower α percentile of the distribution of $\Lambda(data)$. While the test is based on a nested class with a single additional restriction under the null hypothesis, it is not clear if the standard asymptotics hold. This is because the marginal models are still nonparametric and the asymptotic convergence is over the constructed sieve. We follow a resampling method similar to the previous chapter to construct the critical region.

3.2.1 Numerical illustration

To investigate the power of the proposed test we performed a limited simulation experiment. The sample sizes considered were again $n = 64, 100, 400$ corresponding to equally spaced square grids with 8, 10 and 20 points. The number of Monte Carlo replications were 500 while number of bootstrap replications were 200. To implement he inverse quantile of the monotone Bernstein function we used a grid search with grid spacing equal to 0.01. The response surface were solved from the interaction index equation using R root solving routine. We consider the two examples discussed in the previous section. One of the surface is monotonic without any particular feature and the isoboles are all parallel. The other is with a pronounces shoulder and has isoboles that are not parallel.

The regular parametric models are expected do well in the first example while in the second example the fit is expected to be inadequate. The marginal Bernstein functions are chosen to have degree equal or one more than the true value. The performances are similar and only the case with fitted degree more than the true degree is reported.

Table 3.2.1 shows the power of the one degree of freedom test for Example 4 for different values of sample size and interaction parameter τ. The error standard deviation is fixed at $\sigma = 0.02$. The test is slightly oversized with size getting closer to the nominal level with increasing sample size. The power is not symmetric in terms of departure in the direction of synergistic interaction and antagonistic interaction. The synergistic interaction correspond to positive values of τ and for such values the power is slightly less than a corresponding negative value associated with antagonism. However, in either direction power functions exhibit the expected increase in terms of departure degree and sample size. The test reaches power of one even for values of τ as small as ± 0.3 for a sample of size 100. The empirical evidence points toward consistency of the test.

τ	n = 64	n = 100	n = 400
-1.00	1.00	1.00	1.00
-0.50	0.69	1:00	1.00
-0.30	0.60	0.99	1.00
-0.20	0.56	0.94	0.99
-0.10	0.41	0.76	0.80
-0.05	0.23	0.55	0.71
0.00	0.08	0.07	0.06
0.05	0.34	0.77	1.00
0.10	0.51	0.99	1.00
0.20	0.89	1.00	1.00
0.30	0.99	1.00	1.00
0.50	1.00	1.00	1.00
1.00	1.00	1.00	1.00

Table 3.2.1 shows the power of the one degree of freedom test for Example 5 for

different values of sample size and interaction parameter τ. The test maintains the size well for this particular example. The power is again not symmetric in terms of departure in the direction of synergistic interaction and antagonistic interaction. In the direction of synergistic interaction , the power rises more slowly than in the antagonistic direction. However, in either direction power functions shows the expected increase in terms of departure degree and sample size.

α	n = 64	n = 100	n = 400
-1.00	0.88	1.00	1.00
-0.50	0.71	0.95	1.00
-0.30	0.65	0.78	1.00
-0.20	0.44	0.66	1.00
-0.10	0.29	0.54	0.87
-0.05	0.16	0.28	0.65
0.00	0.05	0.04	0.05
0.05	0.29	0.51	1.00
0.10	0.44	0.98	1.00
0.20	0.79	1.00	1.00
0.30	0.97	1.00	1.00
0.50	1.00	1.00	1.00
1.00	1.00	1.00	1.00

3.2.2 Discussion

The one degree of freedom test shows the expected high power in the direction of alternative. Interestingly, the power is higher in the direction of antagonism than in the synegystic direction. The test performs well even for surfaces with multiple

features that are not necessarily captured by parametric forms, such as the Greco model.

3.3 A nonparametric test for Loewe Additivity

In order to have a fully nonparametric test, one could let the alternative model vary over all possible response surfaces. Of course, a nonparametric class that includes the Loewe additive model in the interior of the class is ideal because then a nested sequence can be formed and conventional likelihood ratio test may be carried out in a manner similar to the previous chapters.

Throughout this section we will assume that the scaling function is fully known and fixed. The analysis of course will depend on the scaling function and sensitivty of results to the choice of scaling function is a topic of future investigation. To obtain a nonparametric class one could think of the response surface as a function of the two dose reduction ratios, $\frac{L(x_1)}{L(\mu_1^{-1}(E))}$ and $\frac{L(x_2)}{L(\mu_2^{-1}(E))}$ and form a function class in two variables on the positive half plane that includes the function $(x + y)$ as a special case. However, given that the standard likelihood ratio asymptotic theory no longer necessarily holds in the infinite dimensional case, the advantage of the nested model maybe not substantial over a general alternative. Also, since we have already explored the general Bernstein formulation for the response surface over the unit square, we choose to model the alternative using general Bernstein surface only with range and monotonicity restrictions. From the asymptoic uniform approximation property of the Bernstein class, the null is included in the closure of the class. However, for a given approximation the null set of Loewe additive functions are not embedded in the Bernstein class. Thus, the hypothesis testing problem is that of nonnested alternatives.

There is a substantial literature on non-nested model testing starting with Cox

(1961,1962),Atkinson (1970), Pesaran (1974) and Pesaran and Deaton (1978). Several authors have looked at non-nested regression models in particular, which is analogous to our present problem. The literature on non-nested linear regression (linear and nonlinear) include work by by Davidson and MacKinnon (1981), Fisher and McAleer (1981), Dastoor (1983), Deaton (1982), Sawyer (1983), Gourieroux, Monfort, and Trognon (1983), and Godfrey and Pesaran (1983), Gourieroux and Monfort (1995) and Smith (1992). Using likelihood ratio test for comparing nonnested models has been looked in by Voung (1985). In particular, Voung (1989) establish the asymptotic distribution of the LRT in the non-nested case. We apply Voung's method to develop the test in the present context.

Consider again the sieves $\Theta^G_{0,\beta}$ and Θ^G_{β} for the constrained cases and the unconstrained case in the previous chpater. The only difference is that the sieve in the constrained case is being used in conjunction with the Loewe additive model. The LRT is

$$\Lambda(data) = \frac{\sup\limits_{(\beta^{(1)},\beta^{(2)})\in\Theta^G_{0,\beta}} L^*(\beta^{(1)},\beta^{(2)})}{\sup\limits_{\beta^{(12)}\in\Theta^G_{\beta}} L(\beta^{(12)})} \tag{3.8}$$

where L^* denotes the likelihood obtained using the Loewe additive model (3.6) and L is the likelihood obtained from the two dimensional Bernstein model without any restriction on the coefficients other than monotonicity and range restriction over the grid. Clearly, the likelihoods are not nested. In such situation Voung (1989) showed under mild regularity condition that

$$n^{-1/2}\Lambda(data) - E_0(\ \to N(0,\omega^{*2})$$

where the variance ω^{*2} can be consistently estimated by

$$\hat{\omega}^{2*} = n^{-1}\sum_{i=1}^{n}[\ell_i^*(\beta^{(1)},\beta^{(2)}) - \ell_i(\beta^{(12)})]^2 - [n^{-1}\sum_{i=1}^{n}(\ell_i^*(\beta^{(1)},\beta^{(2)}) - \ell_i(\beta^{(12)}))]^2$$

where ℓ_i^* and ℓ_i are the log-likelihood for the ith observation under the Loewe model and under the Bernstein model. We use a t-like test

$$T = n^{-1/2}\Lambda(data)/\sqrt{\hat{\omega}^{2*}} \tag{3.9}$$

and reject based on the asymptotic critical value. The critical region will be

$$\mathcal{R}_\alpha = \{Data : T > z_{\alpha/2}\}$$

where z_{α_2} is the upper $\alpha/2$ normal percentile.

Since our models are only asymptotically nested it not clear whether the normal limit is achieved by the likelihood ratio. Nevertheless, we designed a simulation experiment to first get an understanding about the convergence of the distribution of the LRT. The data generating model is the Greco-type model in the previous section. The specific model is the one in Example 5 with sample size is $n = 64$ and sign $= 0.02$.. The alternative models are generated by varying τ form -0.5 to 0.5. Figure 3.4 shows the distribution of the LRT (based on 200 Monte Carlo samples) in the null case ($\tau = 0$) and in two moderately far alternatives given by $\tau = -0.5$ and $\tau = 0.5$. The distribution of the LRT in the alternative case shifts significantly to the left of the null distribution and hence it is expected to have high power for such alternative values. Figure 3.5 provides the comparative study of the null distribution and the alternative distributions at $\tau = -.2$ and $\tau = 0.2$. Even though the alternatives are not far from the null, the alternative distributions have shifted to the left indicating that there will be adequate power in the testing procedure. Figure 3.6 and Figure 3.7 show that same thing but the impact of the sample size is clear. The histograms are much tighter, the null is close to having a mean of 0 and standard deviation equal to one and the alternative distributions have moved further away from the null. Thus, empirically it seems that with increasing sample size, the power of the test will go up. Table 3.2.1 provides numeric values of the power of the test (3.9). It is clear that

the power increases as the model moves away from null and for larger sample size. The test is severely oversized for $n = 64$ but the size is close to the nominal level for $n = 100$.

3.4 Discussion

We have proposed two tests for Loewe additivity. The first one is a semiparametric test and has good performance against local alternatives. The second test is fully nonparametric which uses the non-nested LRT distribution theory. One could use a resmapling based procedure in the non-nested case as well.

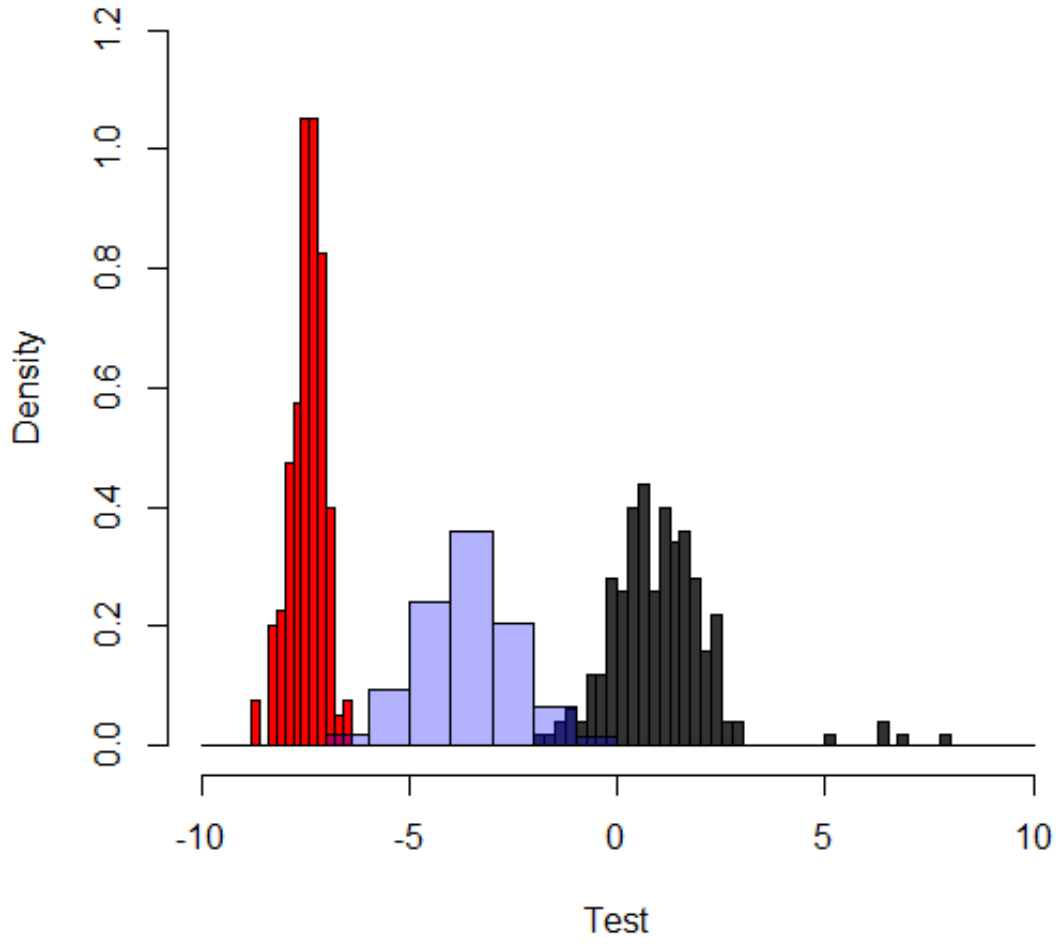


Figure 3.4: Distribution of the test statistics under different values of the interaction parameter for $n = 64$. The dark gray histogram corresponds to $\tau = 0$ or the null value, the red corresponds to $\tau = -.5$ or moderate antagonism and the light blue corresponds to $\tau - 0.5$ or moderate synergy.

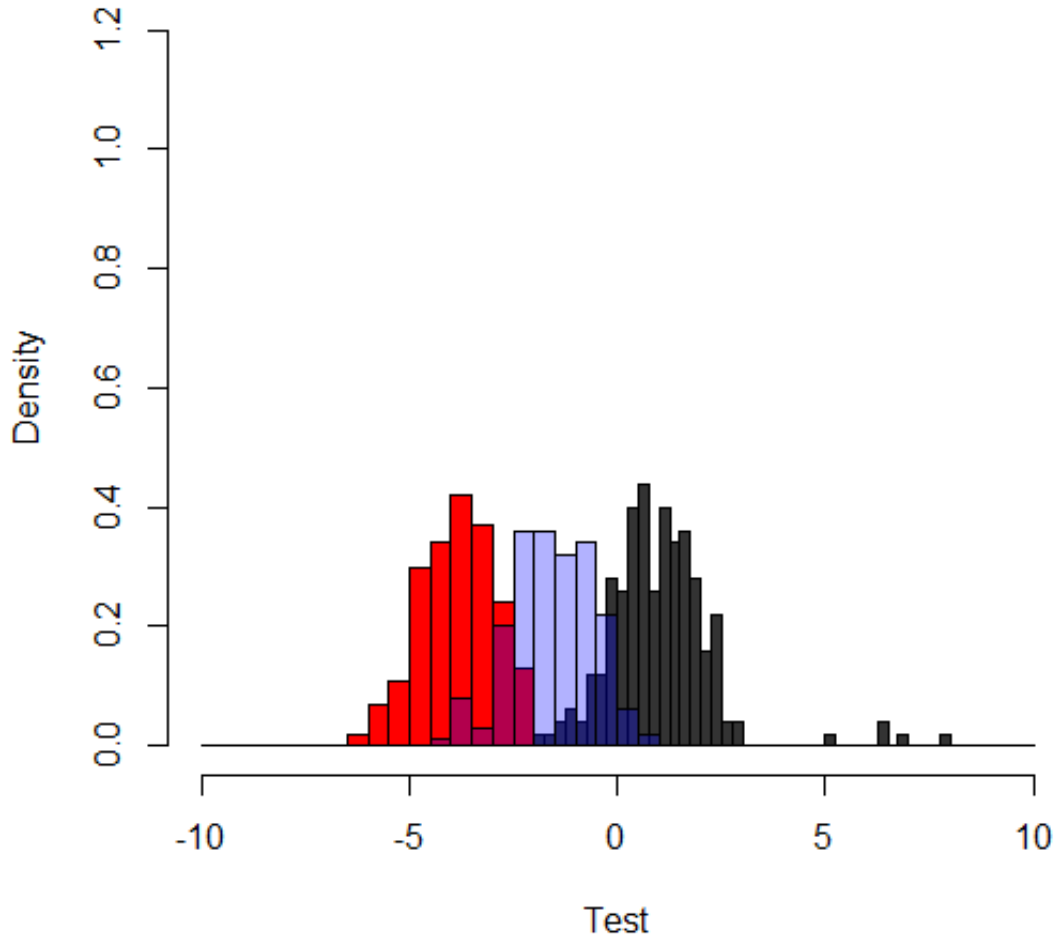


Figure 3.5: Distribution of the test statistics under different values of the interaction parameter for $n = 64$. The dark gray histogram corresponds to $\tau = 0$ or the null value, the red corresponds to $\tau = -.2$ or limited antagonism and the light blue corresponds to $\tau - 0.2$ or limited synergy.

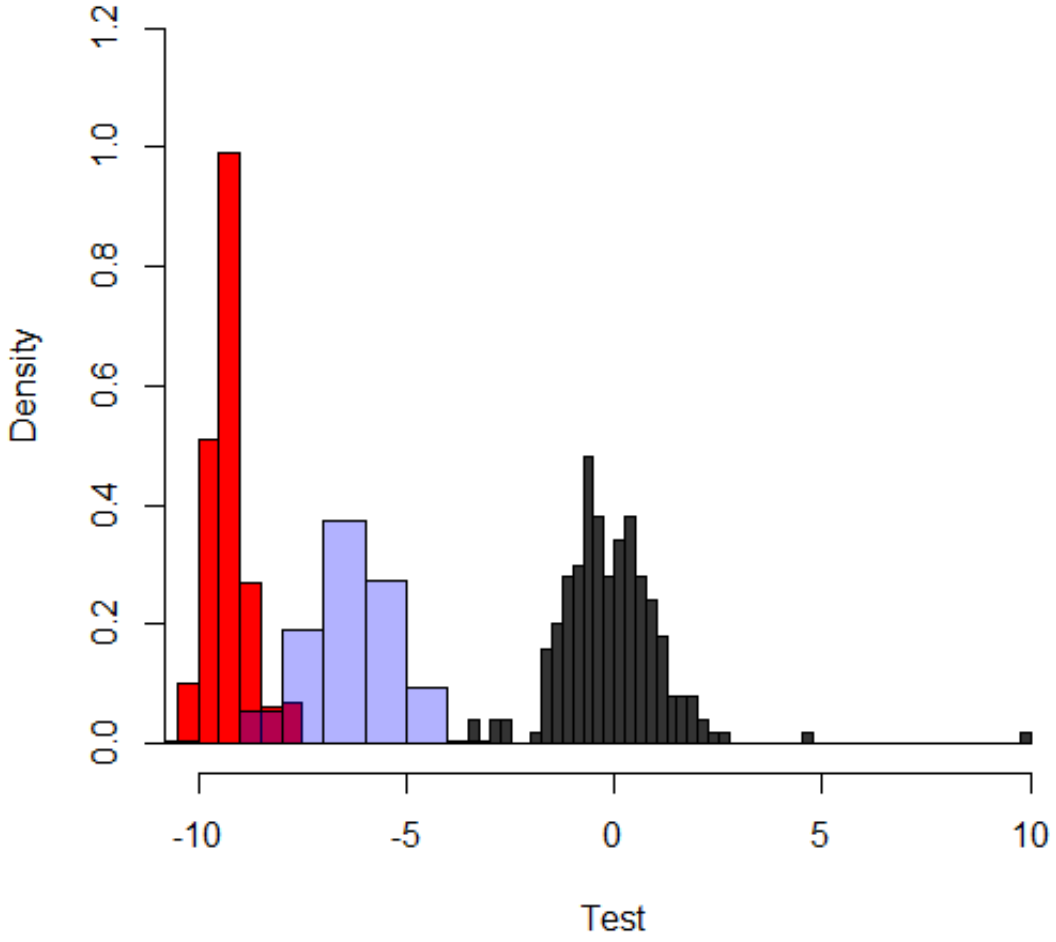


Figure 3.6: Distribution of the test statistics under different values of the interaction parameter for $n = 100$. The dark gray histogram corresponds to $\tau = 0$ or the null value, the red corresponds to $\tau = -.5$ or moderate antagonism and the light blue corresponds to $\tau - 0.5$ or moderate synergy.

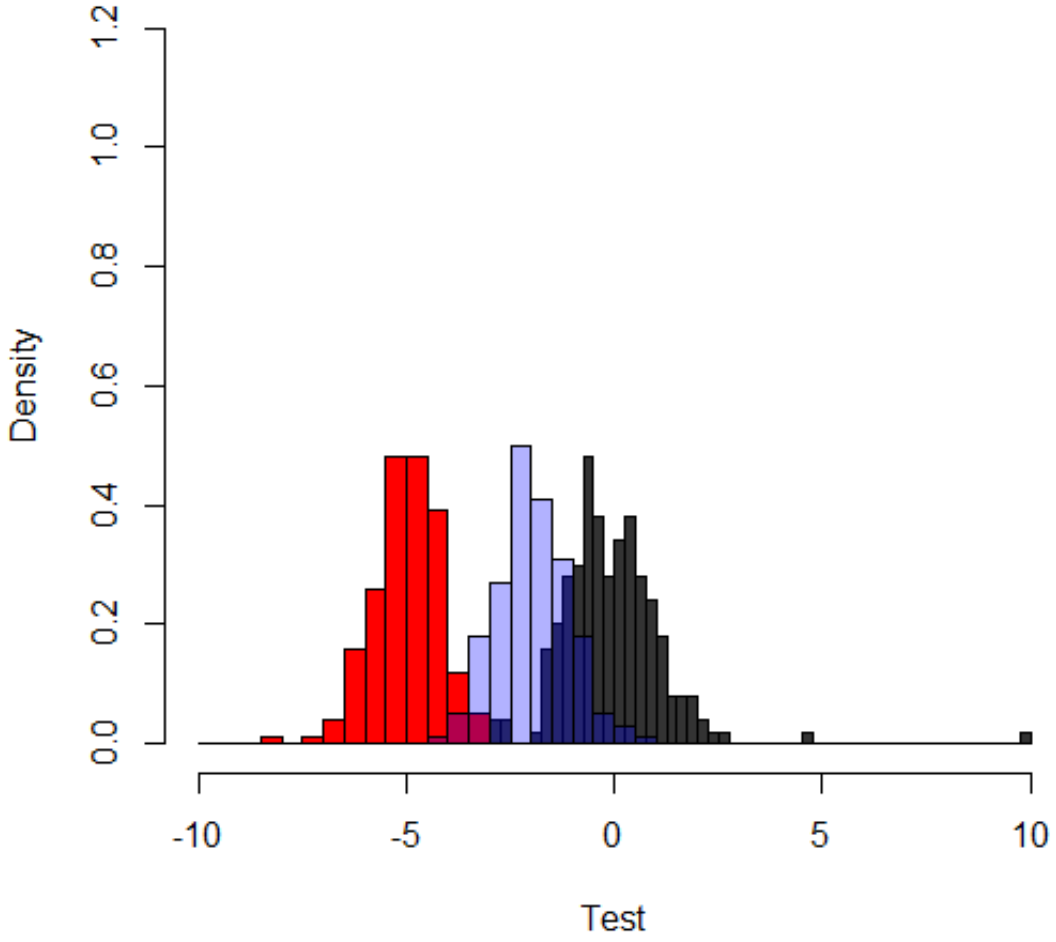


Figure 3.7: Distribution of the test statistics under different values of the interaction parameter for $n = 100$. The dark gray histogram corresponds to $\tau = 0$ or the null value, the red corresponds to $\tau = -.2$ or limited antagonism and the light blue corresponds to $\tau - 0.2$ or limited synergy.

τ	$n = 64$	$n = 100$
-0.50	1.00	1.00
-0.45	1.00	1.00
-0.40	1.00	1.00
-0.35	1.00	1.00
-0.30	0.99	1.00
-0.25	0.96	0.99
-0.20	0.87	0.93
-0.15	0.53	0.75
-0.10	0.27	0.35
-0.05	0.18	0.19
0.00	0.16	0.06
0.05	0.16	0.05
0.10	0.18	0.08
0.15	0.25	0.37
0.20	0.36	0.52
0.25	0.46	0.76
0.30	0.55	0.89
0.35	0.63	0.95
0.40	0.72	0.99
0.45	0.81	1.00
0.50	0.93	1.00

Chapter 4

Future work

There are several important directions in which the proposed methodology can be extended. In this chapter we discuss some preliminary formulation for future research projects.

4.1 Directional testing

One of the main emphasis of future investigation will be to describe tests that are consistent for one-sided alternatives. To this end, we define procedures that consider directional departure from the null under both the bliss independence model and also the Loewe additivity model. The test are defined both globally (entire dose space) and locally (subset of interest).

4.1.1 Bliss independence

In the previous chapter we demonstrated that the Bernstein model is capable of modeling the response function for a variety of shapes. We now develop a method based on the Bernstein model for testing one sided interaction, e.g. synergy or antagonism when the null hypothesis is Bliss independence. Let the marginal response at

dose levels x_1 and x_2 be $\mu_1(x_1)$ and $\mu_2(x_2)$, respectively. Also let the response at a combined dose (x_1, x_2) be $\mu(x_1, x_2)$. Then a synergystic behavior at the dose level (x_1, x_2) would mean that $\mu(x_1, x_2) < \mu_1(x_1)\mu_2(x_2)$ and similarly an antagonistic behavior at the given dose level would mean $\mu(x_1, x_2) > \mu_1(x_1)\mu_2(x_2)$. In other words, in presence of synergy one would expect a more dire effect from a dose combination that you would expect if the chemicals worked independently with no multiplicative interaction. Similarly, in presence of antagonistic behavior one would expect a one chemical to mitigate the adverse effect of the other chemical, thereby increasing the probability of cell survival under the combination dose. The response will we said to have synergy over a region $\mathcal{R}$ of dose combinations if $\mu(x_1, x_2) < \mu_1(x_1)\mu_2(x_2)$ for all $(x_1, x_2) \in \mathcal{R}$.

4.1.2 Testing at low dose combination

In many applications, the scientists are interested in testing for synergy/antagonism at relatively lower levels of the dose combination. The primary reason is that in nature, humans are generally exposed to lower doses. Also at higher doses the cell survival probabilities are usually very small with or with any interaction and hence of no practical interest. Thus, we first concentrate on the lower region of the dose space $\mathcal{R} = \{(x_1, x_2) : 0 < x_1 < U_1,\ 0 < x_2 < U_2\}$, where U_1, U_2 are appropriately chosen constants. The first important observation is that the terms in the Bernstein expansion that contribute most to the low doses are the first few terms. Thus, to capture multiplicative interaction we would consider departure in the coefficients of the initial terms from those in the null factorisable model.

Consider the null space $\Theta_0 = \{(\beta^{(12)} \in [0,1]^{J_1 J_2} : \beta^{(12)} = \beta^{(1)} \otimes \beta^{(2)}$ for some $\beta^{(1)} \in [0,1]^{J_1}, \beta^{(2)} \in [0,1]^{J_2}\}$. We will develop a low degrees of freedom test for testing departure from in this null in the low dose region. The alternatives that we consider

for antagonistic behavior are in

$$\Theta^A = \{\beta^{(12)} \in [0,1]^{J_1 J_2} : \exists\, \beta^{(1)} \in [0,1]^{J_1}, \beta^{(2)} \in [0,1]^{J_2}, 0 \le \lambda_{1,1}, \ldots, \lambda_{d_1,d_2} \le 1,$$
$$\text{such that } \beta^{(12)}_{ij} = (1-\lambda_{ij})\beta^{(1)}_1 \beta^{(2)}_j + \lambda_{ij}, 1 \le i \le d_1, 1 \le j \le d_2,$$
$$\text{and } \beta^{(12)}_{ij} = \beta^{1)}_i \beta^{(2)}_j \text{ if } i > d_1 \text{ or } j > d_2\} \tag{4.1}$$

Similarly, the alternatives for synergistic behavior are in

$$\Theta^S = \{\beta^{(12)} \in [0,1]^{J_1 J_2} : \exists\, \beta^{(1)} \in [0,1]^{J_1}, \beta^{(2)} \in [0,1]^{J_2}, 0 \le \lambda_{1,1}, \ldots, \lambda_{d_1,d_2} \le 1,$$
$$\text{such that } \beta^{(12)}_{ij} = (1-\lambda_{ij})\beta^{(1)}_1 \beta^{(2)}_j, 1 \le i \le d_1, 1 \le j \le d_2,$$
$$\text{and } \beta^{(12)}_{ij} = \beta^{1)}_i \beta^{(2)}_j \text{ if } i > d_1 \text{ or } j > d_2\} \tag{4.2}$$

Typically, the departure df ($d_1 d_2$) will be allowed to increase with the sample size n, but at a slower rate than the degree of the model $J_1 J_2$. Note that when $\lambda_{ij} = 0, 1 \le i \le d_1, 1 \le j \le d_2$ then the corresponding $\beta^{(12)}$ is factorisable and belongs to the null. Thus, if we denote the full parameter space as $\Theta^A = \Theta^A_{\beta^{(1)},\beta^{(2)},\lambda}$ to denote the parameterization through the quantities $\beta^{(1)}, \beta^{(2)}$ and $\lambda = \{\lambda_{1,1}, \ldots, \lambda_{d_1,d_2}\}$, then

$$\Theta^A_{\beta^{(1)},\beta^{(2)},0} = \Theta_0,$$

the null space. Similarly define for the synergistic alternatives.

The number of parameters in the full model is $J_1 + J_2 + d_1 d_2$ and in the null model is $J_1 + J_2$. For illustration, we describe the test and the procedure for the antagonistic alternatives. Everything can be translated to tests for synergy instead of antagonism by replacing Θ^A by Θ^S. Let Θ_M denote the sieve as in the first chapter, consisting of all functions in $C[0,1]^2$ with negative partial derivatives at the grid. Then a likelihood ratio test for antagonism at low doses, i.e.,

$$H_0 :\ \beta^{(12)} \in \Theta_0 \ \ vs \ \ H_1 :\ \beta^{(12)} \in \Theta^A - \Theta_0,$$

is given by

$$\Lambda(data) = \frac{\max\limits_{\beta^{(12)} \in \Theta_M \bigcap \Theta_0} L(\beta^{(12)}|data)}{\max\limits_{\beta^{(12)} \in \Theta_M \bigcap \Theta^A} L(\beta^{(12)}|data)}$$

Note that the proposed test is somewhat of the same flavor as Tukey's one df test in the parametric set up. Here we are considering low dimensional directional alternatives and letting the dimension of the departure from null to also increase slowly with the sample size. Computation of the LRT can done by using optimization routines that allow linear constraints. For the numerator, the maximization is the same as that of the constrained optimization in the previous chapter. For the denominator, the optimization is now over the parameters $\beta^{(1)}$, $\beta^{(2)}$ and λ, all of which satisfy linear constraints, and the linear constraints associated with the monotonicity assumption over the sieve can now be written as linear constraints involving both the β and the λ parameters.

Typically, $-2\log\Lambda(data) \to \chi^2_{d1d_2}$ which will allow us to set the critical region as

$$\mathcal{C} = \{data : -2\log\Lambda(data) > \chi^2_{d1d_2,\alpha}\}.$$

However since the dimension and complexity of the model are increasing with sample size, we resort to resampling methods to obtain better finite sample approximation to the distribution of $-2\log\Lambda(data)$. The resampling algorithm is analogus to the nondirectional case and can be implemented in R. We plan to investigate the power properties of the test.

4.1.3 A global test for synergy/antagonism

If one is interested in testing for global effect of one-sided interaction, then following the previous section a more general alternative space can be described. First we describe the intuition behind the formulation. Consider the combined effect $\mu(x_1, x_2)$ and the assumed joint effect $\mu_1(x_1)\mu_2(x_2)$ under the assumption of no multiplicative

interaction. If there is global antagonistic effect, then the difference $\Delta(x_1, x_2) = \mu(x_1, x_2) - \mu_1(x_1)\mu_2(x_2)$ is also a smooth positive function on the domain $[0, 1]^2$. Hence we can model the difference function using a lower order Bernstein polynomial, e.g.

$$\Delta(x_1, x_2) = \sum_{k_1=1}^{K_1} \sum_{k_2=1}^{K_2} \eta^{(12)} B(x_1, k_1, K_1 - k_1 + 1) B(x_2, k_2, K_2 - k_2 + 1).$$

Thus, under the null hypothesis the response is

$$\mu(x_1, x_2) = \sum_{j_1=1}^{J_1} \sum_{j_2=1}^{J_2} \beta_{j_1}^{(1)} \beta_{j_2}^{(2)} B(x_1, j_1, J_1 - j_1 + 1) B(x_2, j_2, J_2 - j_2 + 1),$$

and under the alternative

$$\begin{aligned} \mu(x_1, x_2) = & \sum_{j_1=1}^{J_1} \sum_{j_2=1}^{J_2} \beta_{j_1}^{(1)} \beta_{j_2}^{(2)} B(x_1, j_1, J_1 - j_1 + 1) B(x_2, j_2, J_2 - j_2 + 1) \\ & + \sum_{k_1=1}^{K_1} \sum_{k_2=1}^{K_2} \eta^{(12)} B(x_1, k_1, K_1 - k_1 + 1) B(x_2, k_2, K_2 - k_2 + 1) \end{aligned}$$

However, in order to have identifiability of the parameters, one would want the factor representation to be maximal, that is the factorisable part of $\mu(x_1, x_2)$ has to be the largest factorisable function that is completely below $\mu(x_1, x_2)$ Once, that is ensured through the parameterization then a similar likelihood ratio test can be used to test the global alternative of antagonism.

$$H_0 : \eta = 0 \;\; H_1 : \eta > 0,$$

where the last expression mean at least one component of η is positive. Here note that the β parameters are again playing the role of nuisance parameters. Typically, the df for the difference function $K_1 K_2$ will be small, but will be allowed to slowly increase with the sample size.

The likelihood ratio based test is then defined in a similar manner as in the previous cases. We plan to investigate the properties of the test both numerically and theoretically.

4.2 Loewe Additivity

Under the assumption of Loewe additivity, the associated isobole will have a particular form as shown in Figure 4.1. The main intuition of developing the test for one-sided interaction based on isoboles is as follows: If there is synergy between the agents then a mixture will be more potent than the sum of individuals and hence the survival rate will drop below E for smaller dose levels than those required under no interaction. Hence, the isobole for E will be inside the isobole under Loewe additivity. Such an isobole is shown in the Figure 4.1. Under antagonistic behavior, the chemicals will mitigate each others effects and hence a the survival rate at the dose combination along the Loewe additivity line will be higher. Hence the E isobole will lie outside (away from the origin) the Loewe additivity line and is shown in Figure 4.1.

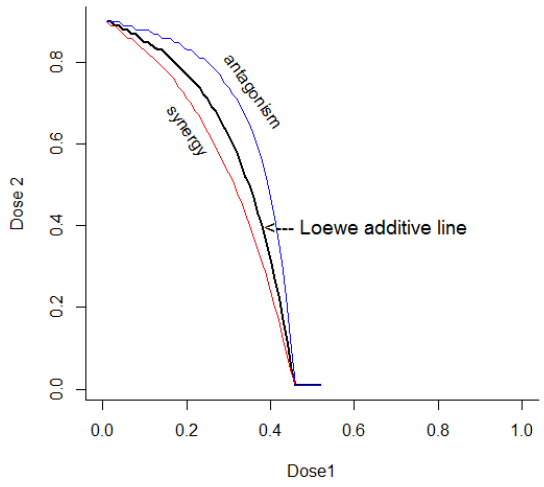


Figure 4.1: Isoboles for synergy and antagonism.

For a given situation, the interaction pattern may be more complicated than simple synergy or simple antagonism. The mixture maybe synergistic in some parts of the dose space and antagonistic in other parts of the dose space. Thus, for a given

effect E, the isoboles need not lie entirely on one side of the Loewe additive line. The pattern maybe as shown in Figure 4.2 where the red line is the true isobole and the thick line is the Loewe additive line. Thus, the true isobole for effect E intersects the Loewe additive line at effect E at a dose combination which is not on the boundary of the dose space. The intersecting isoboles generate two areas, T_S and T_A. In general, when the curves intersect multiple times, T_A will be the union of areas to the left of the Loewe additive line between the line and the true isobole and T_S will be the union of areas to the right of the Loewe additive line between the line and the true isobole. Intuitive, to test there is synergy against the null that there is no synergy one should test $T_S > 0$ versus $T_S = 0$. However, as in equivalence testing, the alternative needs to be changed to an interval hypothesis to be able to have a valiid testing procedure. Thus we propose to test the hypothesis for synergy as

$$H_0 : \ T_S > \delta_S \text{ versus } H_1 : \ T_S < \delta_S, \tag{4.3}$$

where the value of δ_S is prespecified and has to be determined by external considerations. Similarly the hypothesis for checking if there is antagonism is

$$H_0 : \ T_A > \delta_A \text{ versus } H_1 : \ T_A < \delta_A. \tag{4.4}$$

Of course the estimated isobole and the Loewe additive lines will be slightly different from the true ones. However, under a consistent model one would expect that with increasing sample size the isoboles to converge to their respective true values. The estimated areas $\hat{T}_S$ and $\hat{T}_A$ are going to be used for testing. The test for synergy would reject in favor of synergistic interaction if the value of $\hat{T}_S$ is small. That is the critical region for the test of synergy is

$$\mathcal{R}_{S,\alpha} = \{Data : \hat{T}_S < c_{S,\alpha}\}$$

where $c_{S,\alpha}$ is the lower $100(1-\alpha)\%$ percentile of the distribution of T_S under the null $T_S = \delta_S$. The value could be obtained by resampling. Similar the critical region for

the test for antagonism will be

$$\mathcal{R}_{A,\alpha} = \{Data : \hat{T}_A < c_{A,\alpha}\}$$

where $c_{A,\alpha}$ is the lower $100(1-\alpha)\%$ percentile of the distribution of T_A under the null $T_A = \delta_S$.

In the above we have suppressed the dependence of the quantities on the effect E for notational convenience but in reality $T_S = T_{S,E}$ and $T_A = T_{A,E}$. If the interest lies in a range of effects, then we could do multiple hypothesis testing. Suppose the scientist area interested in testing one-sided hypothesis for $E \in [E_L, E_U]$. The area statistics T_{S,E_k} and T_{A,E_k} can be then used to perform the multiple testing procedure or $\max\{T_{S,E_1}, \ldots, T_{S,E_K}\}$ can used depending on the nature of the problem. Since the isoboles area not necessarily equally spaced and also the area inside the Loewe additivity lines are different for different values for E, one could use a weighted test

$$T_S^w == \max\{w_1 T_{S,E_1}, \ldots, w_K T_{S,E_K}\}.$$

An intuitive choice for the weight would be the area within the Loewe additivity line (area of the triangle whose perimeters are the two axes and the Loewe additivity line). This area can be easily computed and hence the weighted test statistic can be computed. the critical region can be define again in an analogus manner for test for synergy and antagonism.

4.3 *Theory: Consistency of Test*

Given that the parameter space is infinite dimensional, it makes sense to investigate the properties of the tests in a large sample set up. An obvious question in this regard is whether the LRT is consistent, i.e. given an alternative value, does the power of the test tend to one as sample size increases to infinity. In general there are no universal

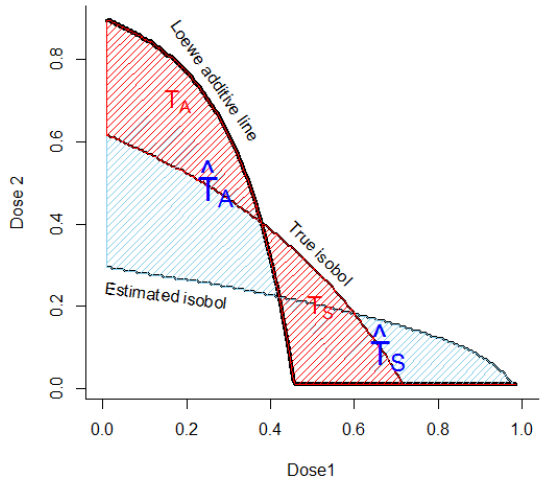


Figure 4.2: dose response when both sunnergy and antagomistic behavior are present in the dose space.

method of establishing consistency of tests in an increasing dimensional set up. There are several papers that establish the consistency of the ML estimators in regression set up with increasing number of variables. The consistency of the LRT will generally follow from that of the estimators and we intend to investigate this route.

4.4 *Semiparametric and other extensions*

Another interesting way of formulating the test without sacrificing on flexibility is to use a semi-parametric form for μ. Let S denote a class of survival functions on $[0, 1]$ and let Φ be a known link from $\mathbb{R}$ to [0,1]. Then one could model μ as

$$\mu = S(\Phi(\beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2)),$$

and one could use the Bernstein class to approximate S thereby making the Bernstein approximation the nuisance parameter and making the parameter β_{12} the interaction

parameter of interest. A likelihood based approach can be again used to perform the test.

An yet another option for building the test is to use a parameterized copula to model the the response function in terms of survival functions where the parameters of the copula will be parameterized to show departure from the null model, i.e. the product copula. Consider a class of copula parameterized by a single parameter γ such that the class includes the product copula as a special case for $\gamma = 0$,

$$\mu = C_\gamma(F_1(x_1), F_2(x_2)).$$

The one-dimensional survival functions are then treated as nuisance parameters. Then the relevant hypothesis of interest in $\gamma = 0$ and can be tested in a likelihood framework.

4.5 Real data analysis

Finally, while the project started with the motivating data on battery waste, we have analyzed the data in the present thesis because it is proprietary data. We would definitely work toward implementing the proposed methods on real data to understand the advantages and disadvantages of nonparametric schemes in real data situation.

Bibliography

[1] Atkinson, A. (1970): A Method for Discriminating Between Models (with Dis cussion)," Journal of the Royal Statistical Society, B, 32, 323– 353.

[2] Berenbaum MC. Synergy, additivism and antagonism in immunosuppression. A critical review. Clin Exp Immunol 1977; 28: 1-18.

[3] Berenbaum M.C. (1989). What is synergy?. *Pharmacol Rev* **41** 93–141.

[4] P. J. Bickel and D. A. Freedman. Bootstrapping regression models with many parameters. In Festschrift for Erich L. Lehmann, pages 2848. Wadsworth, 1983.

[5] Bliss CI. (1939). The toxicity of poisons applied jointly. *Ann Appl Biol,* **26** 585–615.

[6] Carter WH, Gennings C, Staniswalis JG, Campbell ED, and White KL. (1988). A statistical approach to the construction and analysis of isobolograms. *Int J Toxicol.* **7** 963–973.

[7] Cox, D. (1961): Tests of Separate Families of Hypothesis," Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability.

[8] Cox, D. (1962): Further Results on Tests of Separate Families of Hypotheses," Journal of Royal Statistical Society, B 24, 406– 424

[9] Dastoor, N. K. (1983): RSome Aspects of Testing Non Nested Hypotheses," Journal of Econometrics, 21, 213 228

[10] Davidson, R., and J. MacKinnon (1981): Several Tests for Model Specification in the Presence of Alternative Hypotheses," Econometrica, 49, 781 793

[11] Deaton, A. S. (1982): Model Selection Procedures, or, Does the Consumption Function Exist?," in Evaluating the Reliability of Macroeconomic Models, ed. by G. C. Show, and P. Corsi, pp. 43 65. Wiley: New York

[12] Fidler M, and Kern SE. (2006). Flexible interaction model for complex interactions of multiple anesthetics. *Anesthesiology* **105** 286–296.

[13] Finney, D. J., Ed. (1952). Probit Analysis. Cambridge, England, Cambridge University Press.

[14] Finney, D. J. and W. L. Stevens (1948). "A table for the calculation of working probits and weights in probit analysis." Biometrika 35(1-2): 191-201.

[15] Fisher, G. R., and M. McAleer (1981): Alternative Procedures and Associated Tests of Significance for Non Nested Hypotheses," Journal of Econometrics, 16, 103 119.

[16] D. A. Freedman. Bootstrapping regression models. The Annals of Statistics, 9(6):12181228, 1981.

[17] GodFrey, L. G., and M. H. Pesaran (1983): RTests of Non Nested Regression Models: Small Sample Adjustments and Monte Carlo Evidence," Journal of Econometrics, 21, 133 154

[18] Goldoni M, Johansson C. A mathematical approach to study combined effects

of toxicants in vitro: evaluation of the bliss independence criterion and the loewe additivity model. Toxicol In Vitro 2007; 21: 759-69.

[19] Gourieroux, C., and A. MonFort (1994): RTesting Non Nested Hypotheses," in Handbook of Econometrics, Volume IV, ed. by R. F. Engle, and D. L. McFadden. Elsevier, Oxford

[20] Gourieroux, C., A. MonFort, and A. Trognon (1983): RTesting Nested or Non Nested Hypotheses," Journal of Econometrics, 21, 83 115

[21] Greco W.R., Bravo G. and Parsons J.C. (1995). The search for synergy: a critical review from a response surface perspective. *Pharmacol Rev.* **47** 331–85.

[22] Hewlett PS. (1969). Measurement of the potencies of drug mixtures. *Biometrics* **25** 477–487.

[23] Kong M, and Lee JJ. (2006). A generalized response surface model with varying relative potency for assessing drug interaction. *Biometrics* **62** 986-995.

[24] Loewe S. (1953). The problem of synergism and antagonism of combined drugs. *Arzneimittelforschung* **3** 285–290.

[25] Greco W.R., Park H.S. and Rustum Y.M. (1990). Application of a new approach for the quantitation of drug synergism to the combination of cisdiamminedichloroplatinum and 1-beta-D-arabinofuranosylcytosine. *Cancer Res.* **50** 5318–5327.

[26] Machado S.G. and Robinson G.A. A direct, general approach based on isobolograms for assessing the joint action of drugs in pre-clinical experiments. *Stat Med.* **13** 2289-2309.

[27] Minto CF, Schnider TW, Short TG, Gregg KM, Gentilini A, and Shafer SL. (2000). Response surface model for anesthetic drug interactions. *Anesthesiology,* **92** 1603-1616.

[28] Murphy, S. and Van der Vaart, A. (1997). Semiparametric likelihod ratio inference. Ann.Statist. 25, 1471-1509.

[29] Pesaran, M. H. (1974): On the General Problem of Model Selection," Review of Economic Studies, **41** 153–171.

[30] Pesaran, M. H., and S. Deaton (1978): Testing Non Nested Non Linear Regression Models. *Econometrica*, **46**, 677–694

[31] Plummer JL. and Short T.G. (1990). Statistical modeling of the effects of drug combinations. *J Pharmacol Methods.* **23** 297-309.

[32] Sawyer, K. R. (1983): RTesting Separate Families of Hypotheses: An Information Criterion. *Journal of the Royal Statistical Society B*,**45**, 89– 99.

[33] Smith, R. J. (1992): Non Nested Tests for Competing Models Estimated by Generalized Method of Moments. *Econometrica,* **60,** 973–980

[34] Tukey, J. (1949). One degree of freedom for non-additivity. *Biometrics,* **5** 232242

[35] Vuong, Q. H. (1989): RLikelihood Ratio Tests for Model Selection and Non Nested Hypothesis," Econometrica, 57(2), 307 333