

The detection of two-component mixture alternatives. Theory and methods.

THÈSE N° 9044 (2018)

PRÉSENTÉE LE 13 DÉCEMBRE 2018
À LA FACULTÉ DES SCIENCES DE BASE
CHAIRE DE STATISTIQUE APPLIQUÉE
PROGRAMME DOCTORAL EN MATHÉMATIQUES

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Daria RUKINA

acceptée sur proposition du jury:

Prof. F. Nobile, président du jury
Prof. S. Morgenthaler, directeur de thèse
Prof. A. Gasnikov, rapporteur
Dr H. U. Burger, rapporteur
Prof. A. Davison, rapporteur



ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Suisse
2018

“Would you tell me, please, which way I ought to go from here?”
‘That depends a good deal on where you want to get to,’ said the Cat.
‘I don’t much care where -’ said Alice.
‘Then it doesn’t matter which way you go,’ said the Cat.
‘- so long as I get *somewhere*,’ Alice added as an explanation.
‘Oh, you’re sure to do that,’ said the Cat, ‘if you only walk long enough.’
— Lewis Carroll

Acknowledgements

I have walked long enough till this point, and the journey was hilly, as many routes in Switzerland. I find it a suitable metaphor to describe my way by long hikes in the mountains, which I got used to for the last five years. I would never finish this distance without the guidance of a person who knows the upcoming turns, steep paths and landscape, my supervisor Stephan Morgenthaler. He supported me all the time and never asked for the impossible. I appreciate his accuracy and acute professional intuition for research questions. I also respect his scientific convictions and diplomatic attitude towards others' points of view. I would like to thank my professor for his wise remarks, engaged discussions, and exciting problems he proposed to me. When Stephan mentioned mixture models as a possible topic, he gave me the article by Larry Wasserman called "Mixture Models: The Twilight Zone of Statistics". The last sentence of this text was "I have decided that mixtures, like tequila, are inherently evil and should be avoided at all costs." What a nice motto, - thought I. Indeed, it appeared to be a tricky problem, and for a long time I was not sure that mixtures can be tamed. Nevertheless, the beginning was intriguing, and my attention was captured. I want to admit that this specific style of encouraging was effective. Besides his professional qualities, I want to note Stephan's absolute serenity, inner confidence and benevolence, which makes him an exceptional supervisor. I join to his other students reviews about how lucky I was to have such a professor.

I want to thank my beloved one, Mikhail, for his invaluable support and eternal disputes concerning life and science. You are my most ferocious and serious critic and compass, and your high standards will stay with me as long as I think of writing a formula. Thank you for being so patient and forgiving, it was a hard time for me, and your shoulder was an essential binding element of my sometimes scattered reality. Another person that I need mention is my beautiful mother. She was my strongest emotional support and private therapist who always reminded me that my PhD is not the whole world, and that I probably should not suffer so much for the case that I am not a proper mathematician. (I hope my professor will skip this line, while reading.) Nevertheless, she urged me to accomplish my goals and be self-confident, regardless how I estimated my progress myself. Finally, I am grateful to my amazing friends, Katerina Rostova, Ksenia Astakhova, Irina Prostakova, Paulo Refinetti, and Grigory Ivanov who filled my PhD life with joy, intellectual conversations and pleasant memories.

Lausanne, 3 September 2018

D. R.

Abstract

In this thesis we deal with one of the facets of the statistical detection problem. We study a particular type of alternative, the mixture model. For the i.i.d. sample X_1, \dots, X_n we consider testing

$$\begin{aligned} \mathcal{H}_0 : X_i &\sim \mathcal{F}_0, & 1 \leq i \leq n, \\ \text{against} & & \\ \mathcal{H}_1 : X_i &\sim \mathcal{F}_1 = (1-p)\mathcal{F}_0 + p\mathcal{G}, & 1 \leq i \leq n. \end{aligned} \tag{1}$$

The null hypothesis corresponds to the absence of a signal, represented by some known distribution, e.g., Gaussian white noise, while in the alternative one assumes that among observations there might be a cluster of points carrying a signal, which is characterized by the distribution \mathcal{G} . The main research objective is to determine a *detectable* set of alternatives in the parameter space combining the parameters of \mathcal{G} and the mixture proportion p . Asymptotic theory, which dominates the articles concerned with this problem, investigates limits of testing errors when $n \rightarrow \infty$ with the parameters of \mathcal{F}_1 depending on n . On the contrary, we focus on finite sample sizes and wish to study the possibility of detecting alternatives for fixed n , given pre-specified error levels.

The first part of the thesis covers theoretical results. The main results here are developed using asymptotic considerations, which are adapted in order to yield conclusions for finite n . Specifically, we propose the approximate minimum sample size n_{\min} necessary for detection of an alternative in the case of small p . For our next result, we introduce a parametrization which relates the parameter space of \mathcal{F}_1 to the sample size, and present the regions of detectability and non-detectability. The regions of detectability are the subsets of the new parameter space (induced by the parametrization) where for a pre-specified type I error rate, the type II error rate of the likelihood ratio test (LRT) is bounded from above by some constant β_{\max} . The regions of non-detectability, conversely, are characterized by the condition that the type II error rate of any test is higher than β_{\max} . We show that for sample sizes starting from 100, there is already good agreement between the proposed theoretical bound and simulation results. Our findings are applicable not only for univariate normals, but for a broader range of distributions. To move towards the real data applications, we also check the performance of some non-parametric testing procedures proposed for this problem and some widely used distributions. We draw conclusions about tests, whose detection boundaries lie close to the boundary corresponding to the LRT. In particular, if $p > 0.1$, then mean-based tests are

Acknowledgements

generally preferred among non-parametric procedures.

In the second part of the thesis we use this argument to develop a framework for clinical trial designs aimed at detecting a *sensitive-to-therapy* subpopulation. Imagine that the hypothesis (1) is tested for patients' responses in a treatment group, and \mathcal{F}_0 represents the distribution of responses in a reference control group. The idea of modeling treatment response as a mixture of subpopulations originates from treatment effect heterogeneity. Methods studying the effects of heterogeneity in the clinical data are referred to as subgroup analyses. They are particularly widespread in oncology clinical trials, where the success of the therapy depends on cancer characteristics. However, designs accounting for possible response heterogeneity are rarely discussed, though in some cases they might help to avoid trial failure due to the lack of efficacy. For example, in neurological disorders, such as depression, pain and anxiety, the major confounding factor is a placebo effect. Its size can mask the drug-specific effect, which in turn leads to trial failure. In our work we consider two possible subgroups of patients, drug responders and drug non-responders. Given no preliminary information about patients' memberships, we propose a framework for designing randomized clinical trials that are able to detect a responders' subgroup of desired characteristics. We also propose strategies to minimize the number of enrolled patients whilst preserving the testing errors below given levels and suggest how the design along with all testing metrics can be generalized to the case of multiple centers.

The last part of the thesis is not directly related to the preceding parts. We present two supervised classification algorithms for real-data applications. One of the proposed techniques might be useful in a broad range of classification problems with not excessive dimensionality. The other algorithm is developed specifically for annotating unknown genes to known gene sets using RNASeq data. Both proposed methods are computationally inexpensive and tractable alternatives to more complicated existing models.

Keywords: adaptive clinical trial design, mixture models, multiple testing, placebo effect, statistical detection problem

Résumé

Nous travaillons, dans cette thèse, sur une des facettes du problème de détection de signaux. Nous étudions un cas particulier d'alternative, le modèle de mélange. Pour un échantillon de variables aléatoires i.i.d X_1, \dots, X_n , nous testons

$$\begin{aligned} \mathcal{H}_0 : X_i &\sim \mathcal{F}_0, & 1 \leq i \leq n \\ \text{contre} & & \\ \mathcal{H}_1 : X_i &\sim \mathcal{F}_1 = (1-p)\mathcal{F}_0 + p\mathcal{G}, & 1 \leq i \leq n. \end{aligned} \tag{1}$$

L'hypothèse nulle correspond à l'absence d'un signal représenté par une distribution connue, par exemple un bruit blanc gaussien, alors que dans l'hypothèse alternative, on peut considérer que parmi les observations, il existe un groupe de points portant un signal caractérisé par la distribution \mathcal{G} . L'objectif est de déterminer les alternatives *déTECTABLES* dans l'espace des paramètres combinant les paramètres de \mathcal{G} et la proportion du mélange p . La théorie asymptotique prédominante dans la littérature liée à ce problème étudie les limites des erreurs du test quand $n \rightarrow \infty$ avec les paramètres de \mathcal{F}_1 dépendant aussi de n . De notre côté, nous voulons contribuer à l'étude des échantillons de taille finie dans le but de trouver si la détection de l'alternative est possible pour un n fixé avec des niveaux d'erreurs prédéfinis.

La première partie de la thèse présente les résultats théoriques. Les principaux résultats sont développés dans le cadre de la théorie asymptotique adaptée pour produire des conclusions pour n fini. En particulier, nous proposons l'approximation de la taille minimale de l'échantillon nécessaire pour la détection de l'alternative dans le cas où p est petit. Pour notre deuxième résultat, nous introduisons la paramétrisation qui lie l'espace des paramètres de l'alternative à la taille de l'échantillon, et présentons les régions où la détectabilité est possible ainsi que celles où elle ne l'est pas. Les premières régions sont les sous-ensembles du nouvel espace de paramètres (induit par la paramétrisation) pour lesquels le taux d'erreur de la deuxième espèce du test du maximum de vraisemblance est bornée supérieurement par une constante β_{\max} , étant donné la taille de l'échantillon et un taux d'erreur de première espèce prédéfini. Les régions de non-détectabilité sont celles où le taux d'erreur de deuxième espèce est supérieur à β_{\max} quelque soit le test. Nous démontrons l'existence d'un accord clair entre les résultats théoriques et les simulations, quand la taille de l'échantillon dépasse 100. Contrairement aux études précédentes, nos résultats ne sont pas seulement applicables aux mélanges de distributions normales univariées, mais aussi pour une classe de distributions plus générale. Dans le but d'appliquer ces résultats à des données réelles, nous étudions la

Acknowledgements

performance de certains tests non paramétriques. Nous tirons des conclusions concernant les tests, pour lesquels les bornes de détections sont proches de celle du test du maximum de vraisemblance. En particulier, nous obtenons que, pour $p > 0.1$, les tests basés sur la moyenne de l'échantillon sont généralement préférés parmi les tests non paramétriques.

Dans la deuxième partie de la thèse, nous utilisons cet argument pour construire le cadre pour les essais cliniques visant la détection de la sous-population sensible à la thérapie. Imaginons que l'hypothèse (1) soit testée pour les réponses des patients dans un groupe de traitement, et que \mathcal{F}_0 représente la distribution des réponses dans un groupe de contrôle de référence. L'idée de modéliser la réponse des participants comme un mélange de sous-populations vient de l'hétérogénéité de l'effet du traitement. Les méthodes qui étudient les effets de l'hétérogénéité dans les données cliniques sont appelées analyses de sous-groupes. Ils sont particulièrement répandus dans les essais cliniques en oncologie, où le succès de la thérapie dépend des caractéristiques du cancer. Cependant, les essais tenant compte de l'hétérogénéité des réponses possibles sont rarement discutés, bien que dans certains cas, elles puissent aider à éviter l'échec de l'essai en raison du manque d'efficacité. Par exemple, dans les troubles neurologiques, tels que la dépression, la douleur et l'anxiété, le facteur de confusion principal est un effet placebo. Sa taille peut masquer l'effet spécifique du médicament, ce qui pourrait conduire à un échec de l'essai. Dans notre travail, nous modélisons la réponse dans le groupe de traitement comme le mélange de deux sous-populations, les répondeurs et les non-répondeurs au traitement. En l'absence d'informations préliminaires sur l'appartenance des patients, nous proposons un cadre pour la conception d'essais cliniques randomisés capables de détecter un sous-population de répondeurs portant les caractéristiques souhaitées. Nous proposons également des stratégies pour minimiser le nombre de patients inscrits tout en préservant les taux d'erreurs inférieurs à certains niveaux prédéfinis. Nous généralisons des conceptions et des erreurs au cas d'essais conduites dans plusieurs centres.

La dernière partie de la thèse n'est pas directement liée aux parties précédentes. Nous présentons deux algorithmes de classification supervisés pour les applications de données réelles. Une des techniques proposées pourrait être utile dans une large gamme de problèmes de classification pour lesquels la dimension n'est pas très grande. L'autre algorithme est développé spécifiquement pour annoter des gènes inconnus à des ensembles de gènes connus en utilisant des données ARNSeq. Les deux méthodes proposées sont des alternatives élégantes à d'autres modèles existants, tout en restant peu exigeantes sur le plan calculatoire et faciles à interpréter.

Mots clés : conception d'essais cliniques adaptatifs, effet placebo, modèles de mélange, problème de détection statistique, tests multiples

Contents

Acknowledgements	v
Abstract (English/Français)	vii
Introduction	1
1 Introduction	1
1.1 Detection problem. Formulation	2
1.1.1 The quality of testing	2
1.1.2 Asymptotic models	6
1.2 Signal detection boundary. Models and Results	7
1.2.1 Asymptotic Rare/ Weak model	7
1.3 The research motivation	10
2 Detection problems for finite sample sizes	13
2.1 Likelihood ratio test	14
2.2 Asymptotic detection	23
2.3 ϕ -divergences and testing errors	24
2.4 Testing procedures for Gaussian mixtures	31
2.4.1 Univariate Gaussian mixture	31
2.4.2 Multivariate Gaussian mixture	34
2.5 Reciprocal questions. Estimability and classification. Multivariate sparse mixture models.	37
2.6 Summary and conclusions	40
3 Clinical trials	43
3.1 Clinical trials	43
3.2 Treatment effects	46
3.3 Trial designs	47
3.3.1 Response models	48
3.4 Conventional testing for the mixture response	50
3.5 Conclusion	52

Contents

4	Randomized controlled trial (RCT) designs	53
4.1	The Model	53
4.2	The designs	54
4.2.1	Hypothesis testing	55
4.2.2	Simple RCT designs - single center RCT	56
4.2.3	Multicenter RCT designs	62
4.3	Planning the trial	65
4.3.1	Planning a single center RCT	66
4.3.2	Planning a multicenter RCT	70
4.4	Examples	75
4.5	Discussion	76
4.6	The next step. Subgroup analysis	82
4.7	Conclusion	83
5	Classification procedures for real-data applications	85
5.1	Example 1: Classification models for gene expression data in cancer patients	85
5.1.1	Background and objectives	85
5.1.2	Data and Methodology	86
5.1.3	Results and conclusions	91
5.2	Example 2: Linking non-annotated genes to existing gene sets	94
5.2.1	Background and objectives	94
5.2.2	Data and Methods	95
5.2.3	Example. Reactome respiratory electron transport gene set enrichment analysis	98
5.2.4	Conclusion	100
	Bibliography	109
	Curriculum Vitae	111

1 Introduction

How often in statistical analysis does one wish to detect that the data deviate from a given model? Probably, most of the time. This is usually referred to as goodness-of-fit testing. In practice, not all deviations are of equal interest. Depending on applications, the alternative models are “customized” to reflect the underlying mechanisms expected in the data. The testing process is then built in such a way that the evidence against the tested model is sufficient if the data come from the “customized” alternatives.

In many applications a natural way to narrow the alternatives is to attribute the potential differences from the null model to the presence of a subgroup of points following another law. The distribution of the data under the alternative is then modeled as a mixture of two distributions from the same family but with different parameters, with the restriction that one of the components has the same distribution as is assumed under the tested model. Noise, multiple dimensions with little individual differences between the mixture components, and small sample size blur the alternatives, making them overlap with the tested model.

Within this framework the goal is to determine which alternatives are detectable for the given sample size and testing error rates. And, if detection is possible, what are the *best* testing methods?

Outline of the thesis

In the Introduction, we describe the hypothesis testing problem and cover previous results on the asymptotic rates of the detection for Gaussian mixture models. In Chapter 2, we discuss the asymptotic limits of detection for a broader class of distributions and check them for finite sample sizes. A review on clinical trial designs along with the reasons to model the response in the group receiving a medication as a mixture are given in Chapter 3. In Chapter 4 we present clinical trial designs aiming at detection of a drug-sensitive subpopulation at the early stages of drug development. In Chapter 5, we present two classification procedures based on gene expression biomarker data. One is aimed at classifying patients potentially at risk for cancer,

and the second relates unannotated genes to existing gene sets.

1.1 Detection problem. Formulation

The original interest in this topic came from the field of communication systems. Let $X_i(t)$, $i = 1, \dots, n$ be signals in n channels observed in white Gaussian noise of known level $\sigma > 0$. The corresponding stochastic differential equation is

$$dX_{i\sigma}(t) = f_i(t)dt + \sigma dW(t), \quad t \in [0, 1], \quad i = 1, \dots, n.$$

Here, the signal, if exists, is of known form $f_i(t) = a_i f(t) \in L_2(0, 1)$ and $W(t)$ is the standard Wiener process. Signal detection is then formulated as testing the null hypothesis $\mathcal{H}_0: f_1 = \dots = f_n = 0$ corresponding to the absence of a signal versus the alternative $\mathcal{H}_1: f_i \neq 0$ for some $i = 1, \dots, n$, where $\{f_1, \dots, f_n\}$ is the sequence of functions in $\mathcal{F} \subset L_2(0, 1)$. It is convenient to make a transformation

$$x_i = \frac{1}{\sigma} \int_0^1 f(t) dX_{i\sigma}(t) = \theta_i + \xi_i, \quad i = 1, \dots, n;$$

where $\theta_i = \frac{a_i \|f\|^2}{\sigma}$ and $\xi_i \sim \mathcal{N}(0, 1)$. And the test is reduced to testing the mean of the multivariate vector

$$\begin{aligned} \mathcal{H}_0: \theta &= 0 \\ \mathcal{H}_1: \theta &\neq 0, \end{aligned} \tag{1.1}$$

where

$$x = \theta + \xi, \tag{1.2}$$

with $\theta = \left(\frac{a_1 \|f\|^2}{\sigma}, \dots, \frac{a_n \|f\|^2}{\sigma} \right) \in \mathbb{R}^n$ and $\xi \sim \mathcal{N}(0, I^n)$. In the following sections we will discuss which tests are usually employed to minimize testing errors, why and how the set of alternatives in (1.1) should be modified to achieve satisfactory testing errors.

To start, we first need to review some important concepts of parametric hypothesis testing theory.

1.1.1 The quality of testing

Let X be a random variable generated by the probability measure P_θ , where $\theta \in \Theta$ and realizations of X belong to the space $x \in \mathcal{X}$ with σ -algebra \mathcal{A} . Here we only consider the parametric case with $\Theta \in \mathbb{R}^n$. Given the data X from P_θ , we test $\mathcal{H}_0: \theta \in \Theta_0$, $\Theta_0 \subset \Theta$ versus $\mathcal{H}_1: \theta \notin \Theta_0$, or $\mathcal{H}_1: \theta \in \Theta_1 = \Theta \setminus \Theta_0$. The decision rule is a measurable function $\psi(x): \mathcal{X} \rightarrow [0, 1]$ which gives a probability of rejecting the \mathcal{H}_0 for some realization $x \in \mathcal{X}$. To compare tests, one considers

1.1. Detection problem. Formulation

the two possible errors. Type I error corresponds to a false rejection of the null hypothesis and type II error corresponds to a false non rejection of the null when the alternative is true. The corresponding probabilities (or rates) are expressed as

$$\begin{aligned}\alpha(\psi, \theta) &= \mathbb{E}_\theta [\psi], & \theta \in \Theta_0, \\ \beta(\psi, \theta) &= \mathbb{E}_\theta [1 - \psi], & \theta \in \Theta_1.\end{aligned}$$

The values $1 - \beta(\psi, \theta)$ and $\alpha = \sup_{\theta \in \Theta_0} (\alpha(\psi, \theta))$ are called the power and the level of the test.

The uniformly most powerful test

For one-parameter distributional families the *best* test selection is due to Neyman and Pearson who introduced the notion of the *uniformly most powerful (UMP)* test.

Definition 1.1. *The test is called UMP if it belongs to the level α tests $\Psi_\alpha = \{\psi : \alpha(\psi, \theta) \leq \alpha, \theta \in \Theta_0\}$ and has the minimal type II error rate among them:*

$$\beta(\psi^{UMP}, \theta) \leq \beta(\psi_\alpha, \theta) \quad \text{for all } \psi_\alpha \in \Psi_\alpha \quad \text{and } \theta \in \Theta_1.$$

The Neyman–Pearson lemma shows the likelihood ratio test (LRT) to be the most powerful if an alternative is simple (Θ_1 consists of one point). However, for the family of distributions with monotone likelihood ratios and for some partitions it is possible to construct UMP tests for some complex alternatives. Monotone likelihood ratio implies that for any $\theta_1, \theta_2 \in \mathbb{R} : \theta_1 > \theta_2$, $L(x) := \frac{f_{\theta_1}(x)}{f_{\theta_2}(x)}$ is a non-decreasing function.

Theorem 1.1 (Karlin–Rubin). *For the families of distributions with a monotone likelihood ratio, the UMP level α test for $\mathcal{H}_0 : \theta \leq \theta_0$ against $\mathcal{H}_1 : \theta > \theta_0$ satisfies*

$$\psi_C = \begin{cases} 1, & L(x) > C, \\ \gamma, & L(x) = C, \\ 0, & L(x) < C, \end{cases} \quad (1.3)$$

and $C : \mathbb{P}_{\theta_0}(L(x) > C) + \gamma \mathbb{P}_{\theta_0}(L(x) = C) = \alpha$. *Conversely, if the test satisfies (1.3), it is a UMP level α test.*

An example of the UMP test is the one-sided test for the univariate normal mean with a known variance. For $X \sim \mathcal{N}(\theta, \sigma^2)$ to test $\mathcal{H}_0 : \theta \leq 0$ versus $\mathcal{H}_1 : \theta > 0$, the UMP test is $\psi(X) = \mathbb{1}\{X \geq z_{1-\alpha}\}$, where $z_{1-\alpha}$ is the $(1 - \alpha)$ quantile of a standard normal distribution. However, to test for the other side of the parameter space, $\mathcal{H}_0 : \theta \geq 0$, the decision rule is $\psi(X) = \mathbb{1}\{X < z_\alpha\}$, which means that there is no UMP test for the two-sided alternative $\mathcal{H}_1 : \theta \neq 0$. This classical example demonstrates that it is not always possible to find a UMP test among all possible tests, and one should impose additional restrictions on the set of tests.

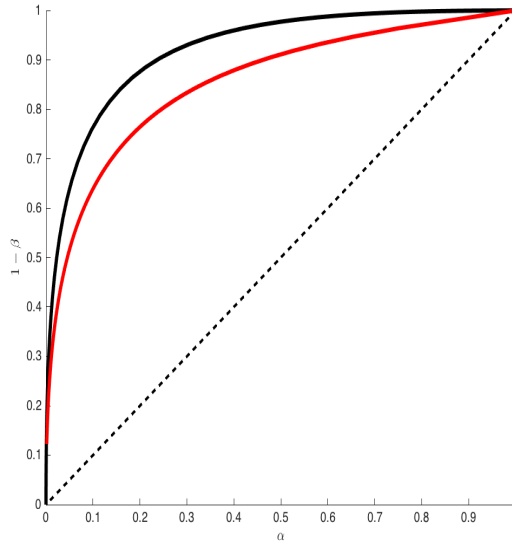


Figure 1.1 – ROC curves for three different tests of the univariate normal mean with the true parameter $\theta = 1$. The black line corresponds to the UMP test $X > C$, the red line is an unbiased test $X^2 > C$, the dashed line corresponds to random guessing.

UMP unbiased and UMP invariant tests

One of the natural requirements for the testing procedure is to preserve some type of symmetry. One would like the decision rule not to be affected by transformations of observed variables such as shifts, scaling and rotations. Two concepts that represent the desired symmetry are *unbiasedness* and *invariance*. When no UMP test exists within the class of all tests, one usually looks for a UMP test among unbiased or invariant tests.

Definition 1.2. A test ψ of level α is called *unbiased* if

$$1 - \beta(\psi, \theta) > \alpha, \quad \theta \in \Theta_1.$$

It is straightforward to see that any UMP test is unbiased.

In the Gaussian signal detection in one dimension, the UMP unbiased test for the two-sided alternative is $\psi = \mathbb{1}(|X| > z_{1-\alpha/2})$. It is convenient to show testing performance with a ROC curve, $1 - \beta(\psi_C, \theta_1)$ plotted against $\alpha(\psi_C, \theta_0)$ for different thresholds C . In Fig. 1.1, the ROC curves for the one-sided test of the normal mean are shown. The UMP test $\psi = \mathbb{1}(X > C)$ is in black, an unbiased test but not the UMP $\psi = \mathbb{1}(X^2 > C)$ is in red and the dashed line corresponds to random guessing (when the decision rule does not depend on the data).

As for invariant tests, we will follow Lehman's (Lehmann and Romano [2006]) introduction. Let $g : \mathcal{X} \rightarrow \mathcal{X}$ be a one-to-one transformation of the outcome space \mathcal{X} of a random variable

1.1. Detection problem. Formulation

X from a family \mathcal{P}_θ , $\theta \in \Theta$. Suppose that the distribution of the transformed random variable $\mathcal{P}_\theta(gX)$ equals $\mathcal{P}_{\bar{g}\theta}(X)$. If both Θ_0 and Θ_1 are preserved under the induced transformation of the parameter space,

$$\bar{g}\Theta_0 = \Theta_0, \quad \bar{g}\Theta_1 = \Theta_1,$$

then the transformation g is invariant for the testing problem. Usually transformations g form a group G , e.g., the group of shifts or rotations.

Definition 1.3. *The test is called invariant for the problem $\mathcal{H}_0 : \theta \in \Theta_0$ versus $\mathcal{H}_1 : \theta \in \Theta_1$ if the transformation g is invariant and $\psi(gx) = \psi(x)$ for all $x \in \mathcal{X}$.*

An invariant test is equivalently determined as the test that depends on the data X only through the so called *maximal invariant*. This is an invariant function which satisfies the condition

$$T(x_1) = T(x_2) \quad \text{implies} \quad x_2 = gx_1 \quad \text{for some } g \in G.$$

Example 1.1.1. *The UMP unbiased and invariant test for the univariate normal mean with unknown variance is*

$$\psi(X_1, \dots, X_n) = \mathbb{1} \left[t_n(X) = \frac{\sqrt{n}\bar{X}}{\sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2}} > C \right],$$

where t_n has Student's distribution with the non-centrality parameter $\sqrt{n}\theta$ and $n - 1$ degrees of freedom.

If G is the group of rotations around the origin, the maximal invariant will be the function of $\|x\|$, i.e., x can be transformed into $x' = gx$ by some $g \in G$ only if x and x' lie on the sphere with the center in the origin.

Example 1.1.2. *The group of rotations (or, more generally, orthogonal transformations) leads to the UMP invariant test for (1.1) for the multivariate normal mean. The decision rule is*

$$\psi(x) : \mathbb{1} \left[t_n(X) = \|X\|^2 = \sum_{i=1}^n X_i^2 > C \right].$$

and t_n has a non-central χ^2 distribution with the non-centrality parameter $\|\theta\|^2$ and n degrees of freedom. Here, $\|x\|^2$ denotes the squared L_2 norm in \mathbb{R}_n .

For unbiased tests the condition $1 - \beta > \alpha$ guarantees protection against random guessing over Θ_1 . Similarly, one should introduce the analogue of it for the invariant tests. This is usually achieved by the reduction of the class of alternatives, separating Θ_0 from Θ_1 . It is easier to introduce via alternative hypothesis testing procedures, the minimax tests.

The minimax approach

The *minimax* criterion considers performance under the worst alternatives. Let $\beta(\psi, \Theta_1) = \sup_{\theta \in \Theta_1} \beta(\psi, \theta)$. The *minimax* test is the test ψ that has the minimal $\beta(\psi, \Theta_1)$ over all tests of level α , Ψ_α . Since the *minimax* hypothesis testing problem focuses on the worst case scenario, it depends on Θ_1 , and in the cases when Θ_0 and Θ_1 are too close, the *minimax* type II error rate achieves its maximum, $1 - \alpha$. These testing problems are called *trivial* and are of no interest. To escape this difficulty one usually removes a neighbourhood of Θ_0 in Θ_1 . It can be shown that for Θ_1 with the removed neighborhood of Θ_0 , the UMP *invariant* test is also the *minimax* test.

1.1.2 Asymptotic models

Removing the neighborhood leads to asymptotic models. Consider the test from Example 1.1.2, but instead of $\Theta_1 = \Theta \setminus \Theta_0$, take the set of alternatives $\Theta_{1n} = \{\theta \in \mathbb{R}^n : \|\theta\| \geq \rho_n\}$, where $\lim_{n \rightarrow \infty} \rho_n = 0$. With the central limit theorem the type II error rate of the test can be approximated as

$$\beta(\psi_\alpha, \theta) = \Phi\left(z_{1-\alpha} - \frac{\rho^2}{\sqrt{2n}}\right) + o_p(1).$$

Here, the test is trivial if $\rho = o(n^{1/4})$. If hypotheses are asymptotically separated, i.e., $\frac{\rho_n}{n^{1/4}} \rightarrow \infty$, the asymptotic power of the test tends to 1 for any fixed alternative in Θ_{1n} . This property is called *consistency*.

For some classes of problems, there are asymptotically UMP tests for large sample sizes which preserve the asymptotic level $\alpha(\psi_n) \leq \alpha + o(1)$, $n \rightarrow \infty$, while maximizing the asymptotic power.

Wilks (1938) showed that under regularity conditions (Cramér, 1946) the generalized likelihood ratio test (GLRT)

$$\Lambda_n(x_1, \dots, x_n) = \left(\frac{\sup_{\theta \in \Theta_0} L(\theta|\mathbf{x})}{\sup_{\theta \in \Theta} L(\theta|\mathbf{x})} \right) < C$$

is consistent for the test $\mathcal{H}_0 : \theta \in \Theta_0$ versus $\mathcal{H}_1 : \theta \in \Theta/\Theta_0$. If the parameter set of the null consists of one point $\Theta_0 = \theta_0$ ($\dim(\Theta_0) = 1$), under the null

$$\Lambda_n = n(\hat{\theta}_{\text{MLE}} - \theta_0)^T I(\theta_0) (\hat{\theta}_{\text{MLE}} - \theta_0) + o_p(1),$$

and $-2 \log \Lambda_n \xrightarrow{d} \chi_{\dim(\Theta) - \dim(\Theta_0)}^2$, where $\hat{\theta}_{\text{MLE}} = \sup_{\theta \in \Theta} L(\theta|\mathbf{x})$ is the maximum likelihood estimate of θ . Under the alternative, $-2 \log \Lambda_n$ converges in distribution to a non-central χ^2 with $\dim(\Theta) - \dim(\Theta_0)$ degrees of freedom.

1.2 Signal detection boundary. Models and Results

Having briefly reviewed the main testing concepts, we are now ready to return to the hypothesis test (1.1). As we have seen, to have a non-trivial problem, one should remove a neighborhood of Θ_0 in Θ_1 . One way to do so is by introducing the set of alternatives Θ_{1n} as follows: each channel i carries a signal of amplitude $\mu_n > 0$ and the number of channels carrying signals is fixed to be $n^{1-\delta_n}$ with $\delta_n \in [0, 1]$, i.e., the multivariate mean vector under the alternative is of the form

$$\Theta_{1n} = \left\{ \theta = (d_1 \mu_n, \dots, d_n \mu_n), \quad d_i \in \{0, 1\}, \quad \sum_{i=1}^n d_i = n^{1-\delta_n} \right\}.$$

This model is explored from the minimax perspective in Ingster and Suslina [2002]. They discovered the conditions when the minimax error rate $\gamma_n(\Theta_{1n}) = \inf_{\psi_n} \left(\alpha(\psi_n) + \max_{\theta \in \Theta_{1n}} \beta(\psi_n, \theta) \right)$ asymptotically tends to zero or one, which correspond to asymptotic distinguishability and non-distinguishability, respectively. If there is an asymptotic distinguishability, authors proposed the minimax consistent tests of the form

$$\psi_n : \gamma_n(\psi_n, \Theta_{1n}) = \gamma_n(\Theta_{1n}) \longrightarrow 0 + o_p(1). \quad (1.4)$$

The results are summarized in the following theorem:

Theorem 1.2. *For the set of alternatives Θ_{1n} the following holds:*

- (1) $\delta_n \longrightarrow \delta \in [0, 3/4)$. If $\sqrt{n^{1-2\delta_n}(e^{\mu_n^2} - 1)} \longrightarrow 0$, then $\gamma_n \longrightarrow 1$. If $\sqrt{n^{1-2\delta_n}(e^{\mu_n^2} - 1)} \longrightarrow \infty$, then for the test $\hat{\psi}_{n, \mu_n} = \max(\psi_{n, \mu_n}, \psi_n^{thr})$, where

$$\psi_{n, \mu_n} \equiv \mathbb{1} \left(\frac{\sum_{i=1}^n \left(e^{-\frac{\mu_n^2}{2} + x_i \mu_n} - 1 \right)}{\sqrt{n(e^{\mu_n^2} - 1)}} > \frac{1}{2} \sqrt{n^{1-2\delta_n}(e^{\mu_n^2} - 1)} \right); \quad \psi_n^{thr} \equiv \mathbb{1} \left(\max_i x_i > \sqrt{2 \log n} \right),$$

$\gamma_n(\hat{\psi}_{n, \mu_n}) \longrightarrow 0$. Moreover, if $\delta_n \longrightarrow \delta \in [0, 1/2)$, the test $\tilde{\psi}_n = \mathbb{1} \left(\frac{1}{\sqrt{n}} \sum_{i=1}^n x_i > \frac{1}{2} \sqrt{n^{1-2\delta_n}(e^{\mu_n^2} - 1)} \right)$ is asymptotically consistent, $\gamma_n(\tilde{\psi}_{n, \mu_n}) \longrightarrow 0$.

- (2) $\delta_n \longrightarrow \delta \in [3/4, 1]$. If $\mu_n < \left(1 - \sqrt{1 - \delta_n} \right) \sqrt{2 \log(n)}$, then $\gamma_n \longrightarrow 1$. Otherwise, for the tests $\hat{\psi}_{n, \mu_n}$ from (1), $\gamma_n(\hat{\psi}_{n, \mu_n}) \longrightarrow 0$.

The test ψ_{n, μ_n} is a modification of the likelihood ratio test, while the tests ψ_n^{thr} and $\tilde{\psi}_n$ are non-parametric.

1.2.1 Asymptotic Rare/ Weak model

In Ingster's work (Ingster and Suslina [2002]) the number of informative channels is fixed to be $n^{1-\delta_n}$. Instead, Donoho and Jin [2004] introduced an uncertainty letting $n^{1-\delta_n}$ be the

expectation of the number of informative channels. Their alternative set is described as follows. Each channel X_i , $i = 1, \dots, n$ carries a signal of amplitude μ_n with a probability $n^{-\delta_n}$, so the data (signals) under the alternative come from a mixture of a signal and white noise. If δ_n is constant, then the probability $n^{-\delta_n}$ decreases with sample size. This assumption is common in genetic applications that aim to find associations between a phenotype and a genetic profile. When one tests for an association between single nucleotide polymorphisms (single position genetic variants, SNPs) and phenotypic variation, as the number of SNPs tested increases, the probability of any individual's SNP being associated with the phenotype decreases. The number of truly associated SNPs is believed to be small. The hypothesis test (1.1) is then transformed into testing

$$\begin{aligned} \mathcal{H}_0: X_i &\stackrel{i.i.d.}{\sim} \mathcal{N}(0, 1), & i = 1, \dots, n, \\ \text{versus} & & \\ \mathcal{H}_{1n}: X_i &\stackrel{i.i.d.}{\sim} (1 - p_n)\mathcal{N}(0, 1) + p_n\mathcal{N}(\mu_n, 1), & i = 1, \dots, n; \end{aligned} \tag{1.5}$$

where $p_n = n^{-\delta}$ with $\delta \in (1/2, 1)$ and $\mu_n = \sqrt{2r \log n}$ with $r \in (0, 1)$. As with the proposed parametrization the proportion of non-null points becomes very small $p_n \in \left(\frac{1}{\sqrt{n}}, \frac{1}{n}\right)$ and the corresponding strength of the signal μ_n is not more than the approximate value of the maximum of n standard normal random variables, $\sqrt{2 \log n}$. This model was referred to as the *Asymptotic Rare / Weak* (ARW) model.

Along with the reduction in dimensionality, the alternative in (1.2.1) is now simple, which leads to the UMP LRT test. The detectable region here is characterized by the sum of error rates of the LRT tending to zero, while in the undetectable region the sum of error rates of any test should tend to one. It was shown that the Higher Criticism (HC) test (originally introduced by Tukey 1976) is optimally adaptive, i.e., regardless of (δ, r) , $\alpha_n(\psi_n^{HC}) + \beta_n(\psi_n^{HC}, \theta) \rightarrow 0 + o_p(1)$ in the detectable region. The detection boundary is shown in Fig. 1.2 and coincides with that from Theorem 1.2:

$$r^*(\delta) = \begin{cases} \delta - \frac{1}{2}, & \frac{1}{2} < \delta \leq \frac{3}{4}, \\ \left(1 - \sqrt{1 - \delta}\right)^2, & \frac{3}{4} < \delta < 1. \end{cases} \tag{1.6}$$

The HC test is a second-level goodness-of-fit test based on the p-values of initial observations: $\text{p-value}_i = \Phi(Z > X_i)$, $i = 1, \dots, n$, where $Z \sim \mathcal{N}(0, 1)$. The test statistic is defined as

$$\text{HC}^* = \max_{1 \leq i \leq n} \frac{\sqrt{n} \left| \frac{i}{n} - \text{p-value}_i \right|}{\sqrt{\frac{i}{n} \left(1 - \frac{i}{n}\right)}},$$

which is the standardized Kolmogorov–Smirnov test of the uniform distribution of the p-values. Under the null hypothesis the p-values are distributed uniformly on $[0, 1]$, and under the alternative the HC test was shown to be asymptotically as effective as ψ_n^{thr} for $\delta \in \left[\frac{3}{4}, 1\right)$

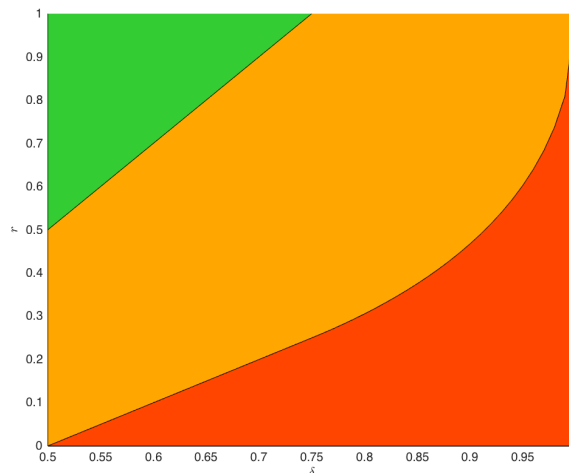


Figure 1.2 – Phase diagram for the gaussian detection problem. The red area corresponds to $r < r^*(\delta)$ the subspace where the asymptotic detection is not possible. The orange area $r^*(\delta) < r < \delta$ corresponds to the subspace where asymptotic detection and consistent estimation of p_n is possible, but identification of signal is not asymptotically possible. The green area $r > \delta$ is the subspace where both asymptotic detection and identification are possible. Misclassification error, the percentage of falsely assigned observations, was proven to tend to the random guessing error, $n\epsilon_n$, for $r < \delta$ for any testing procedure.

and more sensitive for $\delta \in (\frac{1}{2}, \frac{3}{4})$. An extensive overview of this test statistic and its various applications has is discussed in Donoho et al. [2015].

Failure of the GLRT for mixture alternative

It should be clarified why in the previous subsection the specific parametrization was introduced. Why is the alternative $\mathcal{X} \sim p\mathcal{N}(\mu, 1) + (1-p)\mathcal{N}(0, 1)$ with $\Theta_1 = \{(p, \mu) : p \in [0, 1], \mu \in \mathbb{R}\}$ not tested using the GLRT. The reason lies in the regularity conditions. Specifically, the requirement that the true value of the parameter is in the interior of the parameter's space is violated when Θ_0 contains the point $p = 0$. The consequence can be easily seen. Consider n i.i.d. observations X_1, \dots, X_n and test $\mathcal{H}_0 : X \sim \mathcal{N}(0, 1)$ against $\mathcal{H}_1 : X \sim p\mathcal{N}(\mu, 1) + (1-p)\mathcal{N}(0, 1)$, where $p \in [0, 1]$ and $\mu \in \mathbb{R}$. The likelihood ratio is then written as $\prod_{i=1}^n (1 + pZ_i)$, where $Z_i = e^{X_i\mu - \frac{\mu^2}{2}} - 1$ are i.i.d. random variables with zero mean and variance $\sigma^2 = e^{\mu^2} - 1$. Let us fix μ and explore the likelihood ratio as a function of p . Applying Jensen's inequality for the log likelihood ratio gives

$$l(p) = \sum_{i=1}^n \log(1 + pZ_i) \leq n \log\left(\frac{\sum_{i=1}^n (1 + pZ_i)}{n}\right) = n \log\left(1 + \frac{p}{n} \sum_{i=1}^n Z_i\right) \leq p \sum_{i=1}^n Z_i.$$

Therefore, if $\sum_{i=1}^n Z_i < 0$, the maximum of the log likelihood ration equals zero with $\hat{p}_{\text{MLE}} = 0$.

If $\sum_{i=1}^n Z_i \geq 0$, standard theory can be applied. The function $l(p)$ is three times differentiable

Chapter 1. Introduction

with finite $\mathbb{E}|l'(p)|$, $\mathbb{E}|l''(p)|$ and $\mathbb{E}|l'''(p)| < M$, where M does not depend on p , and the Fisher information is positive for any p . Taylor's expansion of $l(p)$ at \hat{p}_{MLE} gives

$$l(\hat{p}_{\text{MLE}}) = l(0) + l'(0)\hat{p}_{\text{MLE}} + \frac{1}{2}l''(0)\hat{p}_{\text{MLE}}^2 + o_p(1) = l'(0)\hat{p}_{\text{MLE}} + \frac{1}{2}l''(0)\hat{p}_{\text{MLE}}^2 + o_p(1).$$

Under the null, $\sqrt{n}\hat{p}_{\text{MLE}} \xrightarrow{p} -\frac{l'(0)}{l''(0)}$. Substituting this in the above expression, one obtains

$$l(\hat{p}_{\text{MLE}}) = -\frac{1}{2} \frac{(l'(0))^2}{l''(0)} + \frac{1}{2} l''(0) o_p\left(\frac{1}{n}\right) + o_p(1).$$

Note that $\frac{(l'(0))^2}{l''(0)} = \frac{(\sum_{i=1}^n Z_i)^2}{\sum_{i=1}^n Z_i^2} = \frac{\left(\sqrt{n} \frac{\sum_{i=1}^n Z_i/\sigma}{n}\right)^2}{\frac{\sum_{i=1}^n (Z_i/\sigma)^2}{n}}$, where $\left(\sqrt{n} \frac{\sum_{i=1}^n Z_i/\sigma}{n}\right)^2 \xrightarrow{d} \chi^2(1)$ and $\frac{\sum_{i=1}^n (Z_i/\sigma)^2}{n} \xrightarrow{p} 1$.

We also have $l''(0) \xrightarrow{p} nI(0)$, and it follows that $\frac{(\sum_{i=1}^n Z_i)^2}{\sum_{i=1}^n Z_i^2} \xrightarrow{d} \chi^2(1)$ and $-2l(\hat{p}_{\text{MLE}}) = \chi^2(1) + o_p(1)$, which is the standard GLRT asymptotics. From the computations above one can see that the asymptotic distribution of $-2l(\hat{p}_{\text{MLE}})$ is a mixture of zero and $\chi^2(1)$ with equal weights. Hartigan (1985) showed that $l(\mu) = \sup_p l(p, \mu)$ is unbounded in probability, which makes the GLRT asymptotically inconsistent. The general case of the failure of the GLRT on the boundary of Θ_0 is covered in Chernoff [1954]. The generalized asymptotic analysis of the maximum likelihood estimators can be found in Moran [1971], Chant [1974], Self and Liang [1987].

1.3 The research motivation

In our work we explore the detection problem in the more general setting, rather than restricting it to the Gaussian case. Given n iid observations X_1, \dots, X_n the testing problem that we will consider is

$$\begin{aligned} \mathcal{H}_0 : X_i &\sim \mathcal{F}_0, & 1 \leq i \leq n \\ \text{against} & & \\ \mathcal{H}_1 : X_i &\sim \mathcal{F}_1 = (1-p)\mathcal{F}_0 + p\mathcal{G}, & 1 \leq i \leq n, \end{aligned} \tag{1.7}$$

where p is the probability for each point to be a signal from the distribution $\mathcal{G} = \mathcal{F}_0(x - \mu)$, while the background distribution \mathcal{F}_0 could be white noise or, in the case of biomedical applications, the distribution of the observations in the control group. For some classes of alternatives we propose a parametrization for the parameter space corresponding to μ and the mixture proportion p . We also include in the parametrization admissible type I and type II error rates. With this framework, given fixed sample size n , type I error rate α and maximum type II error rate β_{max} , we want to reconstruct the regions of the parameter space of \mathcal{F}_1 with $\beta^{\text{LRT}}(\alpha, n) < \beta_{\text{max}}$ (which we refer to as the detectable set of alternatives) and with $\beta(\alpha, n) > \beta_{\text{max}}$ (non detectable set of alternatives). Though we employ asymptotic arguments, we later show the correspondence between theoretical results and empirically computed power boundaries for finite n .

Our goal is to provide a comprehensive guide for experimental design when testing (1.7). In Chapter 2 we show how the likelihood ratio statistic enables us to approach this problem for a broad range of distributions. Although the LRT gives a lower bound on the sample size required for preserving prespecified error rates, it is a parametric test. For this reason we will compare the empirical power of some non-parametric testing procedures previously proposed in the literature for the detection problem with the theoretical detection boundary of the LRT. Based on the results obtained from comparing different non-parametric procedures, in Chapter 4 we develop the framework for clinical trials design suitable for the cases when the patient population is represented as a mixture of two subgroups, drug-responders and drug non-responders.

2 Detection problems for finite sample sizes

In this chapter we present results linking the sample size for the hypothesis test (2.1) with the parameters of the distribution \mathcal{F}_1 , while keeping the fixed type I error rate $\alpha < 0.5$ and controlling type II error rate at level $\beta_{\max} < 0.5$. Suppose that we are given n iid observations X_1, \dots, X_n and wish to test

$$\begin{aligned} \mathcal{H}_0: X_i &\sim \mathcal{F}_0, & i = 1, \dots, n \\ \text{versus} & \\ \mathcal{H}_1: X_i &\sim \mathcal{F}_1 = (1-p)\mathcal{F}_0 + p\mathcal{G}, \quad p \in (0, 1], \quad i = 1, \dots, n, \end{aligned} \tag{2.1}$$

where \mathcal{F}_0 and \mathcal{G} have the same support.

Notation:

- $\varepsilon = \frac{p}{1-p}$. Note that $\lim_{p \rightarrow 0} \frac{p}{1-p} = 0$. Hereafter we consider $0 < p \leq 1/2$ or $0 < \varepsilon \leq 1$.
- $k(X) = \frac{g(X)}{f_0(X)} > 0$, where f_0 and g are the pdfs of \mathcal{F}_0 and \mathcal{G} , respectively.
- $K_m = \mathbb{E}_{f_0} [(k(X))^m]$.

Hereafter all expectations, if not indicated, are taken with respect to \mathcal{F}_0 . Note that K_m has the following properties:

- (1) $K_1 = \mathbb{E}[k(X)] = \mathbb{E}_g[1] = 1$;
- (2) $K_m \cdot K_{m-2} \geq K_{m-1}^2$ (which follows from the Cauchy–Schwarz inequality), $m \geq 2$;
- (3) From (1) and (2) it follows that $K_m \geq K_{m-1} \geq 1$;

Theorem 2.1. *If $K_4 < +\infty$, then the minimum sample size $n_{\min}(\varepsilon)$ for the LRT in testing (2.1) satisfies*

$$\lim_{\varepsilon \rightarrow 0} \varepsilon^2 n_{\min}(\varepsilon) = \left(\frac{\Phi^{-1}(\alpha) + \Phi^{-1}(\beta_{\max})}{\sqrt{K_2 - 1}} \right)^2.$$

Chapter 2. Detection problems for finite sample sizes

For the next result we need to introduce the following conditions:

- C 1.**
1. $f_0(x)$, $x \in (-\infty, +\infty)$ is a continuous, piecewise differentiable, unimodal and symmetric univariate function;
 2. $g(x) = f_0(x - \mu)$, $\mu > 0$;
 3. $K_m < +\infty$ for all $m \geq 2$;
 4. $\lim_{\mu \rightarrow 0} \frac{K_m - 1}{K_2 - 1} < +\infty$ for all $m \geq 3$;

and the parametrization

$$\begin{aligned} \varepsilon &= \varepsilon_0 n^{-\delta}, & \varepsilon_0 &= -(\Phi^{-1}(\beta_{\max}) + \Phi^{-1}(\alpha)) > 0; \quad 0 < \delta \leq 1, \\ \mu &: K_2 = 1 + n^{2r}, & r &\in \mathbb{R}. \end{aligned} \quad (2.2)$$

Theorem 2.2. *Given parametrization (2.2) and with conditions C1 satisfied, the following results for asymptotic detectability hold:*

1. *For an alternative with pdf f_1 whose parameters lie in region*

$$R_{nd} = \left\{ (\delta, r) : r < \delta - \frac{1}{2} \right\},$$

there is a sample size $N : \beta(\alpha, f_1) > \beta_{\max}$, $\forall n \geq N$ for any testing procedure.

2. *For an alternative with pdf f_1 , whose parameters are in*

$$R_d = \left\{ (\delta, r) : \left(r < 0 \text{ and } r > \delta - \frac{1}{2} \right) \text{ or } \left(r \geq 0, r > \delta - \frac{1}{2} \text{ and } K_3 = o(n^{2r+\delta}) \right) \right\}$$

or in

$$R_d^C = \{(\delta^*, r^*) : \delta^* \leq \delta \text{ and } r^* \geq r, \text{ for } (\delta, r) \in R_d\},$$

there is a sample size $N : \beta^{LRT}(\alpha, f_1) \leq \beta_{\max}$, $\forall n \geq N$.

Further in the chapter we show how these results agree with simulations and which non-parametric testing procedures give lower values of the type II error rate.

2.1 Likelihood ratio test

In this section we prove Theorem 2.1. The likelihood ratio test (LRT) for the hypothesis testing problem (2.1), rejects \mathcal{H}_0 if $l_n \equiv \log(\Lambda_n) = \log\left(\frac{\prod_{i=1}^n f_0(x_i)}{\prod_{i=1}^n f_1(x_i)}\right) < C$. It is the most powerful test for a fixed type I error rate. The type I and type II error rates of the LRT are

$$\alpha = \mathbb{P}(l_n < C \mid \mathcal{H}_0), \quad \beta = \mathbb{P}(l_n > C \mid \mathcal{H}_1).$$

If the first two moments $\mathbb{E}[l_1|\mathcal{H}_j]$ and $\mathbb{E}[l_1^2|\mathcal{H}_j]$, $j = 0, 1$ are finite, by the central limit theorem we have

$$\sqrt{n} \left(\frac{1}{n} l_n - \mathbb{E}[l_1|\mathcal{H}_j] \right) \xrightarrow{d} \mathcal{N} \left(0, \text{Var}[l_1|\mathcal{H}_j] \right), \quad j = 0, 1, \quad n \rightarrow \infty. \quad (2.3)$$

The expectation $\mathbb{E}[l_1|\mathcal{H}_0] = \text{KL}(\mathcal{F}_0|\mathcal{F}_1)$ is the Kullback–Leibler divergence between \mathcal{F}_0 and \mathcal{F}_1 . This divergence is not symmetrical, i.e., $\text{KL}(\mathcal{F}_1|\mathcal{F}_0) = \mathbb{E}[-l_1|\mathcal{H}_1] \neq \text{KL}(\mathcal{F}_0|\mathcal{F}_1)$, and does not satisfy the triangle inequality. By Jensen's inequality one can obtain that $\text{KL}(\mathcal{F}_0|\mathcal{F}_1), \text{KL}(\mathcal{F}_1|\mathcal{F}_0) \geq 0$. Let us denote

$$\begin{aligned} \mathbb{E}[l_1|\mathcal{H}_0] &= \text{KL}(\mathcal{F}_0|\mathcal{F}_1) = \text{KL}_1, \\ \mathbb{E}[-l_1|\mathcal{H}_1] &= \text{KL}(\mathcal{F}_1|\mathcal{F}_0) = \text{KL}_2, \\ \text{Var}[l_1|\mathcal{H}_0] &= \text{Var}_1, \\ \text{Var}[l_1|\mathcal{H}_1] &= \text{Var}_2. \end{aligned}$$

If values $\rho_j = \mathbb{E}[|l_1^3|\mathcal{H}_j]$, $j = 0, 1$, are finite, one can obtain the following inequalities using the Berry–Essen theorem Berry [1941], Essen [1942],

$$\begin{cases} \mathbb{P} \left(\frac{l_n}{\sqrt{n}} - \sqrt{n} \text{KL}_1 \leq \Phi^{-1}(\alpha) \sqrt{\text{Var}_1} | \mathcal{H}_0 \right) \leq \alpha + \frac{A}{\sqrt{n}}, \\ \mathbb{P} \left(\frac{l_n}{\sqrt{n}} - \sqrt{n} \text{KL}_1 \leq \Phi^{-1}(\alpha) \sqrt{\text{Var}_1} | \mathcal{H}_0 \right) \geq \alpha - \frac{A}{\sqrt{n}}; \\ \mathbb{P} \left(\frac{l_n}{\sqrt{n}} + \sqrt{n} \text{KL}_2 \geq \Phi^{-1}(1 - \beta) \sqrt{\text{Var}_2} | \mathcal{H}_1 \right) \leq \beta + \frac{B}{\sqrt{n}}, \\ \mathbb{P} \left(\frac{l_n}{\sqrt{n}} + \sqrt{n} \text{KL}_2 \geq \Phi^{-1}(1 - \beta) \sqrt{\text{Var}_2} | \mathcal{H}_1 \right) \geq \beta - \frac{B}{\sqrt{n}}; \end{cases}$$

where $A = \frac{\tilde{C} \rho_0}{(\sqrt{\text{Var}_1})^3}$, $B = \frac{\tilde{C} \rho_1}{(\sqrt{\text{Var}_2})^3}$ and $\tilde{C} < 0.48$ is the Berry-Essen constant. In the inequalities above, we replace α with $\alpha - \frac{A}{\sqrt{n}}$ or with $\alpha + \frac{A}{\sqrt{n}}$, and β is replaced with $\beta - \frac{B}{\sqrt{n}}$ or $\beta + \frac{B}{\sqrt{n}}$. This leads to

$$\begin{cases} \mathbb{P} \left(l_n \leq n \text{KL}_1 + \Phi^{-1} \left(\alpha - \frac{A}{\sqrt{n}} \right) \sqrt{n \text{Var}_1} | \mathcal{H}_0 \right) \leq \alpha, \\ \mathbb{P} \left(l_n \leq n \text{KL}_1 + \Phi^{-1} \left(\alpha + \frac{A}{\sqrt{n}} \right) \sqrt{n \text{Var}_1} | \mathcal{H}_0 \right) \geq \alpha; \\ \mathbb{P} \left(l_n \geq -n \text{KL}_2 - \Phi^{-1} \left(\beta - \frac{B}{\sqrt{n}} \right) \sqrt{n \text{Var}_2} | \mathcal{H}_1 \right) \leq \beta, \\ \mathbb{P} \left(l_n \geq -n \text{KL}_2 - \Phi^{-1} \left(\beta + \frac{B}{\sqrt{n}} \right) \sqrt{n \text{Var}_2} | \mathcal{H}_1 \right) \geq \beta. \end{cases}$$

If n_{\min} is the minimum sample size for the LRT with type I and type II error rates $\alpha < 0.5$ and $\beta < 0.5$, respectively, then the following should be satisfied:

$$n_{\min} \text{KL}_1 + \Phi^{-1} \left(\alpha + \frac{A}{\sqrt{n_{\min}}} \right) \sqrt{n_{\min} \text{Var}_1} > -n_{\min} \text{KL}_2 - \Phi^{-1} \left(\beta + \frac{B}{\sqrt{n_{\min}}} \right) \sqrt{n_{\min} \text{Var}_1},$$

Chapter 2. Detection problems for finite sample sizes

which is equivalent to

$$\sqrt{n_{\min}} > -\frac{\Phi^{-1}\left(\alpha + \frac{A}{\sqrt{n_{\min}}}\right)\sqrt{\text{Var}_1} + \Phi^{-1}\left(\beta + \frac{B}{\sqrt{n_{\min}}}\right)\sqrt{\text{Var}_2}}{\text{KL}_1 + \text{KL}_2}. \quad (2.4)$$

Another necessary condition is that the following inequality for $n_{\min} - 1$ should be satisfied

$$(n_{\min} - 1)\text{KL}_1 + \Phi^{-1}\left(\alpha - \frac{A}{\sqrt{n_{\min} - 1}}\right)\sqrt{(n_{\min} - 1)\text{Var}_1} < -(n_{\min} - 1)\text{KL}_2 - \Phi^{-1}\left(\beta - \frac{B}{\sqrt{n_{\min} - 1}}\right)\sqrt{(n_{\min} - 1)\text{Var}_1},$$

which is equivalent to

$$\sqrt{n_{\min} - 1} < -\frac{\Phi^{-1}\left(\alpha - \frac{A}{\sqrt{n_{\min} - 1}}\right)\sqrt{\text{Var}_1} + \Phi^{-1}\left(\beta - \frac{B}{\sqrt{n_{\min} - 1}}\right)\sqrt{\text{Var}_2}}{\text{KL}_1 + \text{KL}_2}. \quad (2.5)$$

To explore the asymptotic behaviour when $\varepsilon \rightarrow 0$, we use the following inequalities. For all $l \geq 1$ and $\varepsilon > 0$:

$$\begin{aligned} \log(1 + \varepsilon) &< \varepsilon - \frac{\varepsilon^2}{2} + \dots + \frac{\varepsilon^{2l-1}}{2l-1}; \\ \log(1 + \varepsilon) &> \varepsilon - \frac{\varepsilon^2}{2} + \dots - \frac{\varepsilon^{2l}}{2l}; \\ (\log(1 + \varepsilon))^2 &> \varepsilon^2 - \varepsilon^3; \end{aligned} \quad (2.6)$$

Lemma 2.1. *For finite K_3 the following inequalities hold:*

$$\begin{aligned} \frac{\varepsilon^2}{2}(K_2 - 1) - \frac{\varepsilon^3}{3}K_3 &< \text{KL}_1 < \frac{\varepsilon^2}{2}(K_2 - 1) + \frac{\varepsilon^3}{3}K_3; \\ \varepsilon^2(K_2 - 1) - 2\varepsilon^3K_3 &< \text{Var}_1 + \text{KL}_1^2 < \varepsilon^2(K_2 - 1) + 2\varepsilon^3K_3. \end{aligned}$$

Proof. KL_1 and Var_1 can be written as

$$\text{KL}_1 = -\mathbb{E}\left[\log\left(\frac{1 + \varepsilon k(X)}{1 + \varepsilon}\right)\right] = \log(1 + \varepsilon) - \mathbb{E}[\log(1 + \varepsilon k(X))], \quad (2.7)$$

$$\text{Var}_1 + \text{KL}_1^2 = \mathbb{E}\left[\left(\log\left(\frac{1 + \varepsilon}{1 + \varepsilon k(X)}\right)\right)^2\right] = (\log(1 + \varepsilon))^2 - 2\log(1 + \varepsilon)\mathbb{E}[\log(1 + \varepsilon k(X))] + \mathbb{E}[(\log(1 + \varepsilon k(X)))^2]. \quad (2.8)$$

Using (2.6), we get:

$$\begin{aligned} \text{KL}_1 &< \varepsilon - \frac{\varepsilon^2}{2} + \frac{\varepsilon^3}{3} - \varepsilon + \frac{\varepsilon^2}{2}K_2 = \frac{\varepsilon^2}{2}(K_2 - 1) + \frac{\varepsilon^3}{3} \leq \frac{\varepsilon^2}{2}(K_2 - 1) + \frac{\varepsilon^3}{3}K_3; \\ \text{KL}_1 &> \varepsilon - \frac{\varepsilon^2}{2} - \varepsilon + \frac{\varepsilon^2}{2}K_2 - \frac{\varepsilon^3}{3}K_3 = \frac{\varepsilon^2}{2}(K_2 - 1) - \frac{\varepsilon^3}{3}K_3. \end{aligned}$$

The coefficient of the dominant term in both inequalities with KL_1 is positive. For the second moment of the log likelihood ratio we have

$$\begin{aligned}\text{Var}_1 + \text{KL}_1^2 &< \varepsilon^2 - 2 \left(\varepsilon - \frac{\varepsilon^2}{2} \right) \left(\varepsilon - \frac{\varepsilon^2}{2} K_2 \right) + \varepsilon^2 K_2 \\ &= \varepsilon^2 (K_2 - 1) + \varepsilon^3 (K_2 + 1) - \frac{\varepsilon^4}{2} K_2 \leq \varepsilon^2 (K_2 - 1) + \varepsilon^3 (K_2 + 1) \\ &\leq \varepsilon^2 (K_2 - 1) + 2\varepsilon^3 K_3.\end{aligned}$$

For the case $\varepsilon \leq 1$ the upper bound is valid, because $\varepsilon - \frac{\varepsilon^2}{2} \geq 0$ for $\varepsilon \leq 2$. To prove a lower bound, we use the inequality $(\log(1 + \varepsilon))^2 > \varepsilon^2 - \varepsilon^3$. Applying the same technique, we get

$$\text{Var}_1 + \text{KL}_1^2 > (\varepsilon^2 - \varepsilon^3) - 2\varepsilon^2 + (\varepsilon^2 K_2 - \varepsilon^3 K_3) = \varepsilon^2 (K_2 - 1) - \varepsilon^3 (K_3 + 1) \geq \varepsilon^2 (K_2 - 1) - 2\varepsilon^3 K_3.$$

This finishes the lemma. \square

With the assumptions from Lemma 2.1 it follows that the asymptotic behaviour of KL_1 and Var_1 is

$$\frac{\text{KL}_1}{\frac{\varepsilon^2}{2} (K_2 - 1)} \xrightarrow{\varepsilon \rightarrow 0} 1, \quad (2.9)$$

$$\frac{\text{Var}_1}{\varepsilon^2 (K_2 - 1)} \xrightarrow{\varepsilon \rightarrow 0} 1. \quad (2.10)$$

Lemma 2.2. *For the finite K_3 the following inequalities hold:*

$$\begin{aligned}\frac{\varepsilon^2}{2} (K_2 - 1) - \frac{4}{3} \varepsilon^3 K_3 &< \text{KL}_2 < \frac{\varepsilon^2}{2} (K_2 - 1) + \frac{4}{3} \varepsilon^3 K_3; \\ \varepsilon^2 (K_2 - 1) - 2\varepsilon^3 K_3 &< \text{Var}_2 + \text{KL}_2^2 < \varepsilon^2 (K_2 - 1) + 4\varepsilon^3 K_3.\end{aligned}$$

Proof. Notice that if $\eta(X)$ is an arbitrary random variable, $\mathbb{E}_{f_1}[\eta(X)] = \mathbb{E} \left[\eta(X) \frac{1 + \varepsilon k(X)}{1 + \varepsilon} \right]$. Therefore, KL_2 can be written as follows:

$$\text{KL}_2 = \mathbb{E} \left[\left(\frac{1 + \varepsilon k(X)}{1 + \varepsilon} \right) \log \left(\frac{1 + \varepsilon k(X)}{1 + \varepsilon} \right) \right] \quad (2.11)$$

$$= -\log(1 + \varepsilon) + \frac{1}{1 + \varepsilon} \mathbb{E} [\log(1 + \varepsilon k(X))] + \frac{\varepsilon}{1 + \varepsilon} \mathbb{E} [\log(1 + \varepsilon k(X)) k(X)]; \quad (2.12)$$

$$\text{Var}_2 + \text{KL}_2^2 = \mathbb{E} \left[\left(\log \left(\frac{1 + \varepsilon k(X)}{1 + \varepsilon} \right) \right)^2 \frac{1 + \varepsilon k(X)}{1 + \varepsilon} \right].$$

Chapter 2. Detection problems for finite sample sizes

Using (2.6), we have

$$\begin{aligned} \text{KL}_2 &< -\varepsilon + \frac{\varepsilon^2}{2} + (1 - \varepsilon + \varepsilon^2) \left(\varepsilon - \frac{\varepsilon^2}{2} K_2 + \frac{\varepsilon^3}{3} K_3 \right) + \varepsilon^2 (1 - \varepsilon + \varepsilon^2) K_2 \\ &= \frac{\varepsilon^2}{2} (K_2 - 1) + \varepsilon^3 \left(1 - \frac{1}{2} K_2 + \frac{1}{3} K_3 \right) + \varepsilon^4 \left(\frac{1}{2} K_2 - \frac{1}{3} K_3 \right) + \frac{\varepsilon^5}{3} K_3 \\ &\stackrel{(\varepsilon \leq 1)}{\leq} \frac{\varepsilon^2}{2} (K_2 - 1) + \frac{4}{3} \varepsilon^3 K_3; \end{aligned}$$

$$\begin{aligned} \text{KL}_2 &> -\varepsilon + \frac{\varepsilon^2}{2} - \frac{\varepsilon^3}{3} + (1 - \varepsilon) \left(\varepsilon - \frac{\varepsilon^2}{2} K_2 \right) + \varepsilon (1 - \varepsilon) \left(\varepsilon K_2 - \frac{\varepsilon^2}{2} K_3 \right) \\ &= \frac{\varepsilon^2}{2} (K_2 - 1) - \varepsilon^3 \left(\frac{1}{3} + \frac{1}{2} K_2 + \frac{1}{2} K_3 \right) + \frac{\varepsilon^4}{2} K_3 \\ &\geq \frac{\varepsilon^2}{2} (K_2 - 1) - \frac{4}{3} \varepsilon^3 K_3. \end{aligned}$$

For the second moment of the log likelihood ratio under the alternative one has

$$\begin{aligned} \text{Var}_2 + \text{KL}_2^2 &= \mathbb{E} \left[\left(\log \left(\frac{1 + \varepsilon k(X)}{1 + \varepsilon} \right) \right)^2 \frac{1 + \varepsilon k(X)}{1 + \varepsilon} \right] \\ &\leq \mathbb{E} \left[(1 + \varepsilon k(X)) \left(\varepsilon^2 (k^2(X) + 1) - 2 \left(\varepsilon - \frac{\varepsilon^2}{2} \right) \left(\varepsilon k(X) - \frac{\varepsilon^2 k^2(X)}{2} \right) \right) \right] \\ &= \varepsilon^2 (K_2 - 1) + \varepsilon^3 (2 - K_2 + K_3) + \varepsilon^4 \left(\frac{K_2}{2} + K_3 \right) - \frac{\varepsilon^5 K_2}{2} < \varepsilon^2 (K_2 - 1) + 4\varepsilon^3 K_3; \end{aligned}$$

and

$$\begin{aligned} \text{Var}_2 + \text{KL}_2^2 &\geq \mathbb{E} \left[(1 - \varepsilon) \left(\varepsilon^2 k^2(X) - \varepsilon^3 k^3(X) + \varepsilon^2 - \varepsilon^3 - 2\varepsilon^2 k(X) \right) \right] \\ &= \varepsilon^2 (K_2 - 1) - \varepsilon^3 (K_2 + K_3) + \varepsilon^4 (1 + K_3) \geq \varepsilon^2 (K_2 - 1) - 2\varepsilon^3 K_3. \end{aligned}$$

□

Consequently, with the assumptions of Lemma 2.2, one has the following asymptotic behaviour of the first two moments of the log likelihood ratio under the alternative

$$\frac{\text{KL}_2}{\frac{\varepsilon^2}{2} (K_2 - 1)} \xrightarrow{\varepsilon \rightarrow 0} 1, \quad (2.13)$$

$$\frac{\text{Var}_2}{\varepsilon^2 (K_2 - 1)} \xrightarrow{\varepsilon \rightarrow 0} 1. \quad (2.14)$$

Now we can return to (2.4) and (2.5). The values n_{\min} , KL_1 , KL_2 , Var_1 , Var_2 , A and B depend on ε . Obviously $n_{\min} \xrightarrow{\varepsilon \rightarrow 0} +\infty$. If ε is small enough,

$$A = \frac{\tilde{C} \rho_0}{(\sqrt{\text{Var}_1})^3} = \frac{\tilde{C} \mathbb{E} \left[\left| \log \left(\frac{1 + \varepsilon}{1 + \varepsilon k(X)} \right) \right|^3 \right]}{(\sqrt{\text{Var}_1})^3} \leq \frac{\tilde{C} \varepsilon^3 (K_3 + 1)}{\left(\varepsilon^2 (K_2 - 1) - 2\varepsilon^3 K_3 - \left(\frac{\varepsilon^2}{2} (K_2 - 1) + \frac{\varepsilon^3}{3} K_3 \right)^2 \right)^{3/2}} \xrightarrow{\varepsilon \rightarrow 0} \frac{\tilde{C} (K_3 + 1)}{(K_2 - 1)^{3/2}}. \text{ Analogously,}$$

$$B = \frac{\tilde{C} \rho_1}{(\sqrt{\text{Var}_2})^3} = \frac{\tilde{C} \mathbb{E} \left[\left| \log \left(\frac{1+\varepsilon k(X)}{1+\varepsilon} \right) \right|^3 \frac{1+\varepsilon k(X)}{1+\varepsilon} \right]}{(\sqrt{\text{Var}_2})^3} \leq \frac{\tilde{C} (\varepsilon^3 K_3 + \varepsilon^4 K_4 + \varepsilon^3 + \varepsilon^4)}{\left(\varepsilon^2 (K_2 - 1) - 2\varepsilon^3 K_3 - \left(\frac{\varepsilon^2}{2} (K_2 - 1) + \frac{4\varepsilon^3}{3} K_3 \right)^2 \right)^{3/2}} \xrightarrow{\varepsilon \rightarrow 0} \frac{\tilde{C} (K_3 + 1)}{(K_2 - 1)^{3/2}}.$$

Consequently, $\frac{A}{\sqrt{n_{\min}}} \xrightarrow{\varepsilon \rightarrow 0} 0$ and $\frac{B}{\sqrt{n_{\min}}} \xrightarrow{\varepsilon \rightarrow 0} 0$.

For sufficiently small ε , combining (2.4) and (2.5) with the results from (2.9), (2.10), (2.13), (2.14), one obtains

$$\begin{aligned} \varepsilon^2 n_{\min} &> \left(\frac{\Phi^{-1} \left(\alpha + \frac{A}{\sqrt{n_{\min}}} \right) \sqrt{\text{Var}_1} + \Phi^{-1} \left(\beta + \frac{B}{\sqrt{n_{\min}}} \right) \sqrt{\text{Var}_2}}{\text{KL}_1 + \text{KL}_2} \right)^2 \xrightarrow{\varepsilon \rightarrow 0} \left(\frac{\Phi^{-1}(\alpha) + \Phi^{-1}(\beta)}{\sqrt{K_2 - 1}} \right)^2 \\ \varepsilon^2 n_{\min} &< \left(\frac{\Phi^{-1} \left(\alpha - \frac{A}{\sqrt{n_{\min}-1}} \right) \sqrt{\text{Var}_1} + \Phi^{-1} \left(\beta - \frac{B}{\sqrt{n_{\min}-1}} \right) \sqrt{\text{Var}_2}}{\text{KL}_1 + \text{KL}_2} \right)^2 + \varepsilon^2 \xrightarrow{\varepsilon \rightarrow 0} \left(\frac{\Phi^{-1}(\alpha) + \Phi^{-1}(\beta)}{\sqrt{K_2 - 1}} \right)^2. \end{aligned}$$

This ends the proof of Theorem 2.1.

Notice that using (2.3) instead of inequalities (2.4) and (2.5), the testing error rates can be approximated as

$$\alpha \approx \Phi \left(\frac{C - n\text{KL}_1}{\sqrt{n\text{Var}_1}} \right), \quad \beta \approx \Phi \left(\frac{-C - n\text{KL}_2}{\sqrt{n\text{Var}_2}} \right), \quad (2.15)$$

and the decision boundary of the LRT approximately satisfies

$$C(\alpha) \approx \Phi^{-1}(\alpha) \sqrt{n\text{Var}_1} + n\text{KL}_1.$$

The type II error rate of the test is approximately

$$\beta \approx \Phi \left(-\Phi^{-1}(\alpha) \sqrt{\frac{\text{Var}_1}{\text{Var}_2}} - \frac{(\text{KL}_1 + \text{KL}_2) \sqrt{n}}{\sqrt{\text{Var}_2}} \right). \quad (2.16)$$

Consequently, an approximate minimum sample size is

$$n_{\min} \approx \left(\frac{\Phi^{-1}(\alpha) \sqrt{\text{Var}_1} + \Phi^{-1}(\beta_{\max}) \sqrt{\text{Var}_2}}{\text{KL}_1 + \text{KL}_2} \right)^2. \quad (2.17)$$

We use this formula for numerical calculations of n_{\min} in Fig. 2.1. From Lemmas 2.1, 2.2 and Theorem 2.1, it follows that for $\varepsilon \ll \frac{K_2-1}{K_3}$ an approximate minimum sample size is

$$n_{\min} \approx \frac{1}{\varepsilon^2} \left(\frac{\Phi^{-1}(\alpha) + \Phi^{-1}(\beta_{\max})}{\sqrt{K_2 - 1}} \right)^2. \quad (2.18)$$

To illustrate this formula, in Fig. 2.1 we plot asymptotes from (2.18) together with numerically calculated sample sizes from (2.17). In Tab.2.1 we present K_2 and K_3 for some distributions. In Fig. 2.2, we show the regions corresponding to $\varepsilon \ll \frac{K_2-1}{K_3}$ for various distributions.

Chapter 2. Detection problems for finite sample sizes

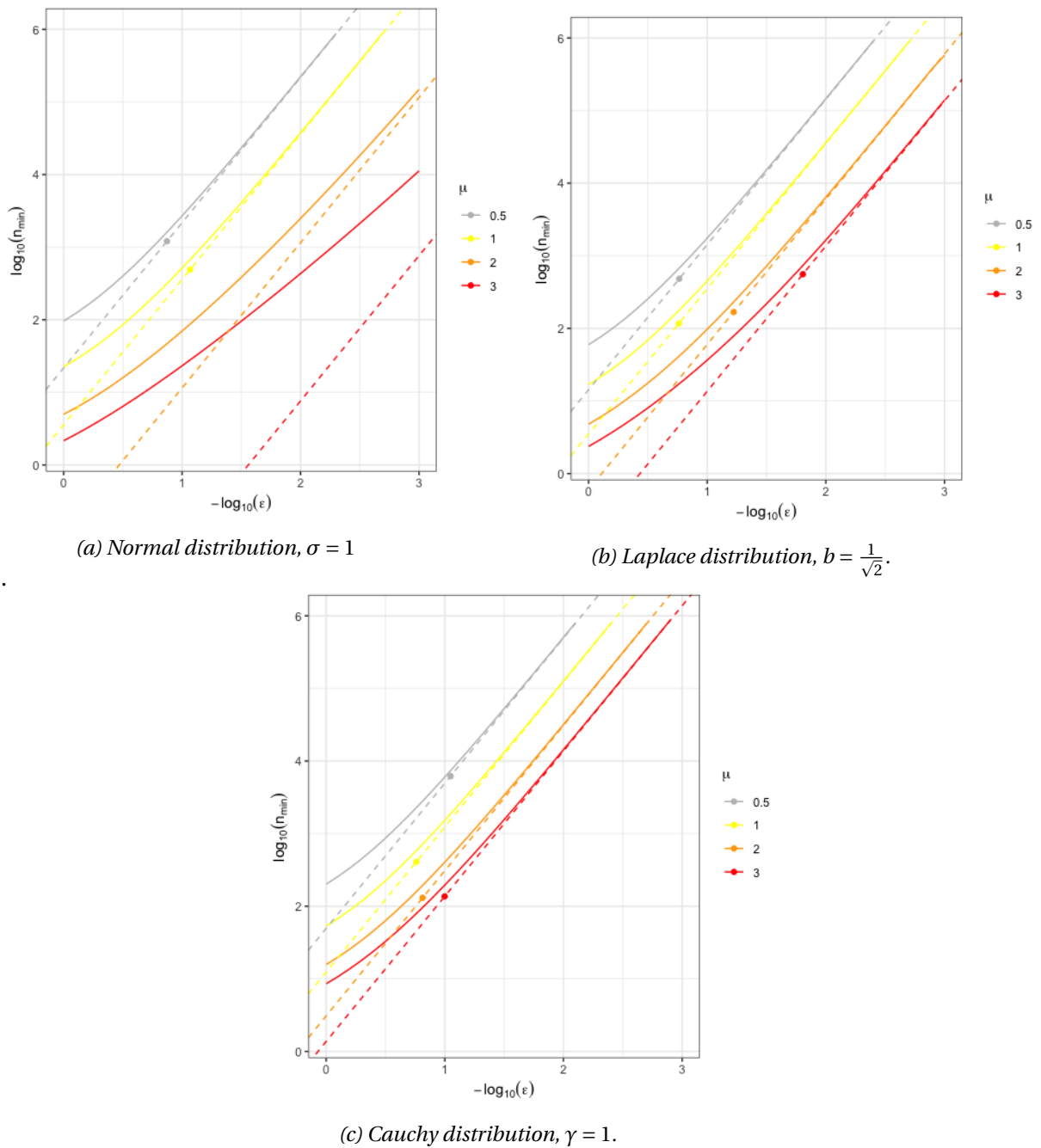
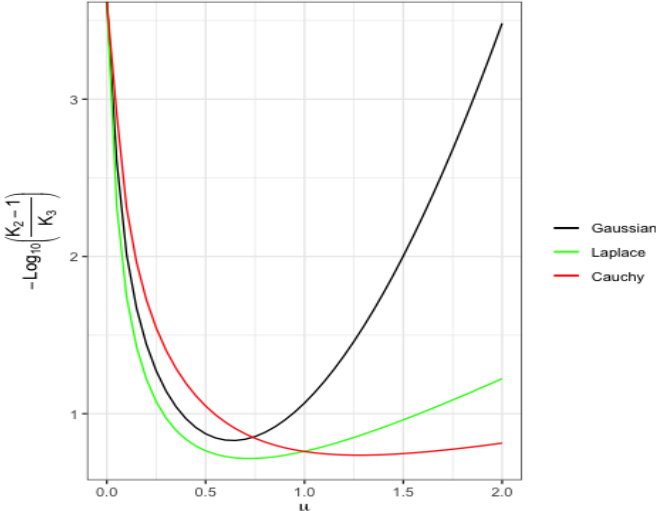


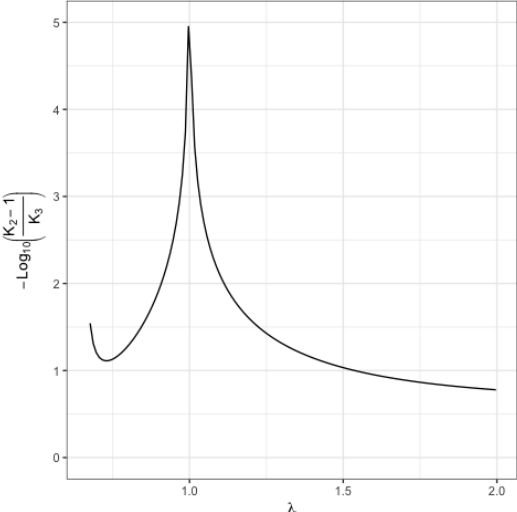
Figure 2.1 – The asymptotes for the sample size necessary to achieve a power of at least $1 - \beta_{\max}$. The parameters are for $\alpha = 0.05$, $\beta_{\max} = 0.2$. (—) - n_{\min} computed numerically, (- - -) - asymptotes from (2.18). Colored dots on the asymptotes correspond to the points where $\epsilon = \frac{K_2 - 1}{K_3}$. For the distributions, see Table 2.1.

$f_0(x)$	$g(x)$	$k(X)$	K_m	$\exists K_m$
<i>Normal distribution</i>				
$\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$	$\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	$\frac{x^2}{\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	$\frac{m(m-1)\mu^2}{\sigma^{m-1}\sqrt{m-(m-1)\sigma^2}}$	$\sigma < \sqrt{\frac{m}{m-1}}$
<i>Normal distribution, $\sigma = 1$</i>				
$\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$	$\frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2}}$	$\frac{x^2}{2} - \frac{(x-\mu)^2}{2}$	$e^{\frac{m(m-1)\mu^2}{2}}$	always
<i>Laplace distribution</i>				
$\frac{1}{\sqrt{2}} e^{-\sqrt{2} x }$	$\frac{1}{2b} e^{-\frac{ x-\mu }{b}}$	$\begin{cases} e^{-\frac{\mu-x}{b} - \sqrt{2}x}, & x < 0 \\ e^{-\frac{\mu-x}{b} + \sqrt{2}x}, & 0 \leq x < \mu \\ e^{-\frac{x-\mu}{b} + \sqrt{2}x}, & x \geq \mu \end{cases}$	$\frac{\sqrt{2}me^{\sqrt{2}(m-1)\mu + 2(m-1)be^{-\frac{m\mu}{b}}}}{(\sqrt{2})^m b^{m-1} (m^2 - 2b^2(m-1)^2)}$	$b < \frac{m}{\sqrt{2(m-1)}}$
<i>Laplace distribution, $b = \frac{1}{\sqrt{2}}$</i>				
$\frac{1}{\sqrt{2}} e^{-\sqrt{2} x }$	$\frac{1}{\sqrt{2}} e^{-\sqrt{2} x-\mu }$	$\begin{cases} e^{-\sqrt{2}\mu}, & x < 0 \\ e^{-\sqrt{2}(\mu-2x)}, & 0 \leq x < \mu \\ e^{\sqrt{2}\mu}, & x \geq \mu \end{cases}$	$\frac{me^{\sqrt{2}(m-1)\mu + (m-1)e^{-\sqrt{2}m\mu}}}{2m-1}$	always
<i>Cauchy distribution</i>				
$\frac{1}{\pi(1+x^2)}$	$\frac{1}{\pi\gamma} \left(1 + \left(\frac{x-\mu}{\gamma}\right)^2\right)^{-\frac{1}{2}}$	$\frac{1+x^2}{\gamma} \left(1 + \left(\frac{x-\mu}{\gamma}\right)^2\right)^{-\frac{3}{2}}$	$1 + \frac{m-1}{2} \left[\frac{1}{\gamma} + \frac{m}{2} - 1 \right] \mu^2 + \dots$	always
<i>Cauchy distribution, $\gamma = 1$</i>				
$\frac{1}{\pi(1+x^2)}$	$\frac{1}{\pi(1+(x-\mu)^2)}$	$\frac{1+x^2}{1+(x-\mu)^2}$	$1 + \frac{m(m-1)}{4} \mu^2 + \dots$	always
<i>Exponential distribution</i>				
e^{-x}	$\lambda e^{-\lambda x}, \lambda > 0$	$\lambda e^{-(\lambda-1)x}$	$\frac{\lambda^m}{m\lambda - (m-1)}$	$\lambda > \frac{m-1}{m}$
<i>Binomial distribution, $x = 0, \dots, N$</i>				
$\binom{N}{x} p_1^x (1-p_1)^{N-x}$	$\binom{N}{x} p_2^x (1-p_2)^{N-x}$	$\left(\frac{p_2}{p_1}\right)^x \left(\frac{1-p_2}{1-p_1}\right)^{N-x}$	$\left(\frac{p_2^m}{p_1^{m-1}} + \frac{(1-p_2)^m}{(1-p_1)^{m-1}}\right)^N$	always

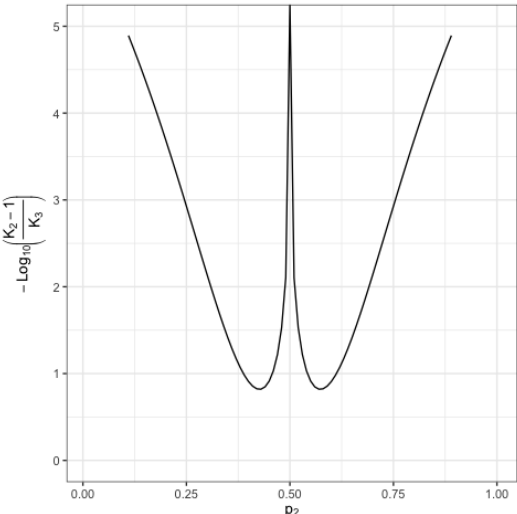
Table 2.1 – Summary information for various distributions



(a) Mixtures of Gaussian with $\sigma = 1$, Cauchy with $\gamma = 1$ and Laplace with $b = \frac{1}{\sqrt{2}}$.



(b) Exponential distribution



(c) Binomial distribution, $p_1 = 1/2, N = 20$

Figure 2.2 – The curves correspond to $\epsilon = \frac{K_2-1}{K_3}$ for various distributions.

2.2 Asymptotic detection

In the previous section we found the conditions for computing an approximate sample size given predefined testing errors. However, for some distributions the condition $\varepsilon \ll \frac{K_2-1}{K_3}$ might imply extremely small ε , which in turn implies enormous n_{\min} . However, for some practical applications, where n is fixed, one might be interested in the parameters of \mathcal{F}_1 which can be detected for this sample size. To approach this problem, we incorporate a parametrization where every alternative f_1 is a member of some sequence f_{1n} . We search for the region in the parameter space where one can asymptotically distinguish \mathcal{H}_0 and \mathcal{H}_1 with the LRT of certain power and level. In other words, the aim of an asymptotic framework is to answer whether the type II error rate for the alternative is less than β_{\max} with $n \rightarrow \infty$. The results are presented in Theorem 2.2.

Recall the conditions C1 given at the beginning of the chapter. If they are satisfied, the function f_1 is fully characterized by the pair (K_2, ε) , since K_2 is an increasing function of μ . Indeed, for the densities symmetric around zero, $K_2 = \int_{-\infty}^{+\infty} \frac{f_0^2(x-\mu)}{f_0(x)} dx$ and

$$\begin{aligned} (K_2')_{\mu} &= -2 \int_{-\infty}^{+\infty} \frac{f_0(x-\mu)f_0'(x-\mu)}{f_0(x)} dx = -2 \left(\int_{-\infty}^{\mu} \frac{f_0(x-\mu)f_0'(x-\mu)}{f_0(x)} dx + \int_{\mu}^{+\infty} \frac{f_0(x-\mu)f_0'(x-\mu)}{f_0(x)} dx \right) = \\ &= -2 \int_{\mu}^{+\infty} f_0(x-\mu)f_0'(x-\mu) \left(\frac{1}{f_0(x)} - \frac{1}{f_0(2\mu-x)} \right) dx. \end{aligned}$$

Note that $f_0'(x-\mu) \leq 0$ and $f_0(x) \leq f_0(2\mu-x)$ for $x \geq \mu$. One can prove that there exists $\bar{x} > 2\mu$, such that $f_0'(\bar{x}-\mu) < 0$ and $f_0(\bar{x}) < f_0(2\mu-\bar{x})$. Therefore, $(K_2')_{\mu} > 0$. This argument justifies the choice of the parametrization (2.2). To be consistent with the introduction, the pair (K_2, ε) represents an element of the parameter space of the alternative hypothesis Θ_{1n} .

The detectable set for the family of distributions satisfying C1 is defined as

$$S_d = \{(\delta, r) : \exists N : \beta^{LRT}(\alpha, f_1) < \beta_{\max}, \forall n \geq N\}.$$

Conversely, the non-detectable set is

$$S_{nd} = \{(\delta, r) : \exists N : \beta(\alpha, f_1) > \beta_{\max}, \forall n \geq N\}.$$

Reconstructing S_d is a challenging problem and an interesting research question. In our work we present sufficient conditions for detectability, i.e., for the experimental design with given α and β_{\max} we conclude whether f_{1n} determined by the parameters (δ, r) could be asymptotically detected. We refer to the subspace where the detection is possible as R_d , and R_{nd} stands for the subspace where detection is not possible. For some distributions for almost all pairs (δ, r) we can say whether alternatives are in R_d or in R_{nd} . For the problems with finite n , which we have mentioned in the beginning of this section, we propose the following heuristic method: find (δ, r) such that $\varepsilon = \varepsilon_0 n^{-\delta}$, $K_2 = 1 + n^{2r}$, and infer about detectability based on R_d and R_{nd} . Further in this chapter, we show that for sample sizes exceeding 100,

this heuristic works quite well.

2.3 ϕ -divergences and testing errors

In this section we prove Theorem 2.2. To start the analysis of asymptotic detectability, we refer to some general inequalities that link the testing errors and the distances between distributions under the null and under the alternative. The divergences (or distances) mentioned in this chapter are the representatives of the so-called ϕ -divergence. Let the vector of iid observations be $(X_1, \dots, X_n) \sim \tilde{\mathcal{F}}_0$ under the null and $\tilde{\mathcal{F}}_1$ under the alternative. We will consider functions $\tilde{\mathcal{F}}_0$ and $\tilde{\mathcal{F}}_1$ continuous in \mathbb{R}^n with the densities $\tilde{f}_j(\mathbf{x}) = \prod_{i=1}^n f_j(x_i)$, $j = 0, 1$. The ϕ -divergence is defined by the convex function $\phi: (0, \infty) \rightarrow \mathbb{R}$, $\phi(1) = 0$, as

$$D_\phi(\tilde{\mathcal{F}}_0, \tilde{\mathcal{F}}_1) = \int_{\mathbb{R}^n} \phi\left(\frac{\tilde{f}_1(\mathbf{x})}{\tilde{f}_0(\mathbf{x})}\right) \tilde{f}_0(\mathbf{x}) d\mathbf{x}.$$

In Section 2.1 approximate expressions for the testing errors are derived using the KLD, which is the ϕ -divergence with $\phi(t) = -\log t$. For $\phi(t) = \frac{|t-1|}{2}$, $D_\phi(\tilde{\mathcal{F}}_0, \tilde{\mathcal{F}}_1)$ is the L_1 distance. One of the results from information theory states that the sum of the testing errors is bounded from below by the L_1 distance.

Denote the rejection region for (2.1) as R . Then the lower bound on the sum of α and β of an arbitrary testing procedure depends on the L_1 distance between $\tilde{\mathcal{F}}_0$ and $\tilde{\mathcal{F}}_1$:

$$\begin{aligned} (1 - \beta) - \alpha &= \int_R (\tilde{f}_1 - \tilde{f}_0) d\mathbf{x} \leq \int_{\tilde{f}_1 > \tilde{f}_0} (\tilde{f}_1 - \tilde{f}_0) d\mathbf{x} = \int_{\tilde{f}_1 > \tilde{f}_0} |\tilde{f}_1 - \tilde{f}_0| d\mathbf{x} \\ &= \int_{\tilde{f}_0 > \tilde{f}_1} |\tilde{f}_0 - \tilde{f}_1| d\mathbf{x} = \frac{1}{2} \int_{\mathbb{R}^n} |\tilde{f}_1 - \tilde{f}_0| d\mathbf{x} \implies \alpha + \beta \geq 1 - \frac{1}{2} \int_{\mathbb{R}^n} |\tilde{f}_1 - \tilde{f}_0| d\mathbf{x}. \end{aligned} \tag{2.19}$$

In turn, the L_1 distance is bounded above by $\sqrt{2}H(\tilde{\mathcal{F}}_1, \tilde{\mathcal{F}}_0)$, where

$$H(\tilde{\mathcal{F}}_1, \tilde{\mathcal{F}}_0) = \sqrt{\frac{1}{2} \int_{\mathbb{R}^n} \left(\sqrt{\tilde{f}_1} - \sqrt{\tilde{f}_0}\right)^2 d\mathbf{x}}$$

is the Hellinger distance, the square root of ϕ -divergence with $\phi = 1 - \sqrt{t}$.

$$\begin{aligned} \int_{\mathbb{R}^n} |\tilde{f}_1 - \tilde{f}_0| d\mathbf{x} &= \int_{\mathbb{R}^n} \left| \sqrt{\tilde{f}_1} - \sqrt{\tilde{f}_0} \right| \left(\sqrt{\tilde{f}_1} + \sqrt{\tilde{f}_0} \right) d\mathbf{x} \leq \sqrt{\int_{\mathbb{R}^n} \left(\sqrt{\tilde{f}_1} - \sqrt{\tilde{f}_0}\right)^2 d\mathbf{x}} \sqrt{\int_{\mathbb{R}^n} \left(\sqrt{\tilde{f}_1} + \sqrt{\tilde{f}_0}\right)^2 d\mathbf{x}} \\ &= \sqrt{\int_{\mathbb{R}^n} \left(\sqrt{\tilde{f}_1} - \sqrt{\tilde{f}_0}\right)^2 d\mathbf{x}} \left(2 + 2 \int_{\mathbb{R}^n} \sqrt{\tilde{f}_1 \tilde{f}_0} d\mathbf{x} \right) \leq 2 \sqrt{\int_{\mathbb{R}^n} \left(\sqrt{\tilde{f}_1} - \sqrt{\tilde{f}_0}\right)^2 d\mathbf{x}} \\ \text{or } \frac{1}{2} \int_{\mathbb{R}^n} |\tilde{f}_1 - \tilde{f}_0| d\mathbf{x} &\leq \sqrt{2}H(\tilde{\mathcal{F}}_1, \tilde{\mathcal{F}}_0). \end{aligned} \tag{2.20}$$

From these inequalities it follows that

$$\alpha + \beta \geq 1 - \sqrt{2}H(\tilde{\mathcal{F}}_1, \tilde{\mathcal{F}}_0),$$

which we will use to prove asymptotic nondetectability.

Lemma 2.3. For $\alpha > 0$ and $l \geq 1$,

$$\begin{aligned} \sqrt{1+\alpha} &> 1 + \frac{\alpha}{2} - \frac{\alpha^2}{8} + \dots - \frac{(4l-3)!!}{2^{2l}(2l)!} \alpha^{2l}; \\ \sqrt{1+\alpha} &< 1 + \frac{\alpha}{2} - \frac{\alpha^2}{8} + \frac{\alpha^3}{16} \dots + \frac{(4l-1)!!}{2^{2l+1}(2l+1)!} \alpha^{2l+1}; \end{aligned}$$

Lemma 2.4. In $R_{nd} = \{(\delta, r) : r < \delta - \frac{1}{2}\}$,

$$\sqrt{2}H(\tilde{\mathcal{F}}_1, \tilde{\mathcal{F}}_0) \rightarrow 0, \quad n \rightarrow \infty.$$

Proof. For R_{nd} it is sufficient to show that the Hellinger affinity

$$A(\tilde{\mathcal{F}}_0, \tilde{\mathcal{F}}_1) = 1 - H^2(\tilde{\mathcal{F}}_0, \tilde{\mathcal{F}}_1) = \int_{\mathbb{R}^n} \sqrt{\tilde{f}_0 \tilde{f}_1} dx = \left(\int_{-\infty}^{+\infty} \sqrt{f_0(x) f_1(x)} dx \right)^n = (A(\mathcal{F}_0, \mathcal{F}_1))^n$$

tends to one when $n \rightarrow \infty$.

For $r < 0$ to bound the Hellinger affinity from below we use Lemma 2.3,

$$\begin{aligned} A(\mathcal{F}_0, \mathcal{F}_1) &= \frac{1}{\sqrt{1+\varepsilon}} \mathbb{E}(\sqrt{1+\varepsilon k}) > \frac{1 + \frac{\varepsilon}{2} - \frac{\varepsilon^2}{8} K_2 + \dots - \frac{(4l-3)!!}{2^{2l}(2l)!} \varepsilon^{2l} K_{2l}}{1 + \frac{\varepsilon}{2} - \frac{\varepsilon^2}{8} + \frac{\varepsilon^3}{16} \dots - \frac{(4l-1)!!}{2^{2l+1}(2l+1)!} \varepsilon^{2l+1}} \\ &= 1 - \underbrace{\frac{\frac{\varepsilon^2}{8} (K_2 - 1) - \frac{\varepsilon^3}{16} (K_3 - 1) + \dots + \frac{(4l-3)!!}{2^{2l}(2l)!} \varepsilon^{2l} (K_{2l} - 1) + \frac{(4l-1)!!}{2^{2l+1}(2l+1)!} \varepsilon^{2l+1}}{1 + \frac{\varepsilon}{2} - \frac{\varepsilon^2}{8} + \frac{\varepsilon^3}{16} \dots - \frac{(4l-1)!!}{2^{2l+1}(2l+1)!} \varepsilon^{2l+1}}}_{z_n}. \end{aligned}$$

Choose $l : \delta > \frac{1}{2l+1}$, then $\varepsilon^{2l+1} = o(\frac{1}{n})$. If $r < \delta - \frac{1}{2}$, then $\varepsilon^2 (K_2 - 1) = o(\frac{1}{n})$. With C1 satisfied, $\varepsilon^m (K_m - 1) = o(\frac{1}{n^{1+(m-2)\delta}})$, $m \geq 2$. It follows that $n z_n \rightarrow 0$ and $A(\tilde{\mathcal{F}}_0, \tilde{\mathcal{F}}_1) \rightarrow 1$.

For $r \geq 0$ the proof follows from the inequalities

$$\begin{aligned} A(\mathcal{F}_0, \mathcal{F}_1) &> \frac{1 + \frac{\varepsilon}{2} - \frac{\varepsilon^2 K_2}{8}}{1 + \frac{\varepsilon}{2} - \frac{\varepsilon^2}{8} + \frac{\varepsilon^3}{16}} = 1 - \frac{\frac{\varepsilon^2}{8} (K_2 - 1) + \frac{\varepsilon^3}{16}}{1 + \frac{\varepsilon}{2} - \frac{\varepsilon^2}{8} + \frac{\varepsilon^3}{16}}, \\ A(\tilde{\mathcal{F}}_0, \tilde{\mathcal{F}}_1) &> \left(1 - \underbrace{\frac{\frac{\varepsilon_0^2 n^{2r-2\delta}}{8} + \frac{\varepsilon_0^3 n^{-3\delta}}{16}}{1 + \frac{\varepsilon_0 n^{-\delta}}{2} - \frac{\varepsilon_0^2 n^{-2\delta}}{8} + \frac{\varepsilon_0^3 n^{-3\delta}}{16}}}_{z_n} \right)^n. \end{aligned}$$

Chapter 2. Detection problems for finite sample sizes

Since $r < \delta - \frac{1}{2}$ and $r \geq 0$, $\delta > \frac{1}{2} > \frac{1}{3}$ and $n z_n \rightarrow 0$. Therefore, $A(\tilde{\mathcal{F}}_0, \tilde{\mathcal{F}}_1) \rightarrow 1$. \square

If α is fixed, as in our case, from the Lemma 2.4 we conclude that in R_{nd} the lower bound on $\beta(\alpha, f_1)$ tends to $1 - \alpha > 0.5 > \beta_{\max}$.

Lemma 2.5. *For the set of parameters*

$$R_d = \left\{ (\delta, r) : \left(r < 0 \text{ and } r > \delta - \frac{1}{2} \right) \text{ or } \left(r \geq 0, r > \delta - \frac{1}{2} \text{ and } K_3 = o\left(n^{2r+\delta}\right) \right) \right\},$$

the type II error rate of the LRT tends to zero when $n \rightarrow \infty$.

Proof. Here we use the result due to Chernoff [1952] on the convergence rates of the error rates for the LRT with rejection bound $C = 0$,

$$\alpha_n \leq \left(\inf_{s>0} \int_{-\infty}^{+\infty} f_1(x)^s f_0(x)^{1-s} dx \right)^n,$$

$$\beta_n \leq \left(\inf_{s>0} \int_{-\infty}^{+\infty} f_0(x)^s f_1(x)^{1-s} dx \right)^n.$$

For $s = 1/2$, the error rates of the LRT are bounded by the Hellinger affinity $A(\tilde{\mathcal{F}}_0, \tilde{\mathcal{F}}_1) = (A(\mathcal{F}_0, \mathcal{F}_1))^n$. To prove the lemma, we will show that $(A(\mathcal{F}_0, \mathcal{F}_1))^n \rightarrow 0$ for pairs (δ, r) in R_d .

If $r < 0$, we apply Lemma 2.3 to the Hellinger affinity:

$$A(\mathcal{F}_0, \mathcal{F}_1) < \frac{1 + \frac{\varepsilon}{2} - \frac{\varepsilon^2}{8} K_2 + \frac{\varepsilon^3}{16} K_3 - \dots + \frac{(4l-5)!! \varepsilon^{2l-1}}{2^{2l-1} (2l-1)!} K_{2l-1}}{1 + \frac{\varepsilon}{2} - \frac{\varepsilon^2}{8} + \dots - \frac{(4l-3)!! \varepsilon^{2l}}{2^{2l} (2l)!} \varepsilon^{2l}}$$

$$= 1 - \frac{\frac{\varepsilon^2}{8} (K_2 - 1) - \frac{\varepsilon^3}{16} (K_3 - 1) + \dots - \frac{(4l-5)!! \varepsilon^{2l-1}}{2^{2l-1} (2l-1)!} (K_{2l-1} - 1) - \frac{(4l-3)!! \varepsilon^{2l}}{2^{2l} (2l)!}}{1 + \frac{\varepsilon}{2} - \frac{\varepsilon^2}{8} + \dots - \frac{(4l-3)!! \varepsilon^{2l}}{2^{2l} (2l)!} \varepsilon^{2l}}, \quad l \geq 2,$$

$$A(\tilde{\mathcal{F}}_0, \tilde{\mathcal{F}}_1) < \left(1 - \frac{\frac{\varepsilon_0^2 n^{-2\delta}}{8} n^{2r} - \frac{\varepsilon_0^3 n^{-3\delta}}{16} (K_3 - 1) + \dots - \frac{(4l-5)!! \varepsilon_0^{2l-1} n^{-(2l-1)\delta}}{2^{2l-1} (2l-1)!} (K_{2l-1} - 1) - \frac{(4l-3)!! \varepsilon_0^{2l} n^{-2l\delta}}{2^{2l} (2l)!}}{\underbrace{1 + \frac{\varepsilon_0}{2} n^{-\delta} - \frac{\varepsilon_0^2 n^{-2\delta}}{8} + \dots - \frac{(4l-3)!! \varepsilon_0^{2l} n^{-2l\delta}}{2^{2l} (2l)!}}_{z_n}} \right)^n.$$

Let $2 \leq l : \delta > \frac{1}{2l}$, then for $r > \delta - \frac{1}{2}$ with C1 satisfied, $n z_n \rightarrow \infty$ and $A(\tilde{\mathcal{F}}_0, \tilde{\mathcal{F}}_1) \rightarrow 0$.

If $r \geq 0$, then

$$\begin{aligned}
 A(\mathcal{F}_0, \mathcal{F}_1) &< \frac{1 + \frac{\varepsilon}{2} - \frac{\varepsilon^2 K_2}{8} + \frac{\varepsilon^3 K_3}{16}}{1 + \frac{\varepsilon}{2} - \frac{\varepsilon^2}{8} + \frac{\varepsilon^3}{16} - \frac{5}{128} \varepsilon^4} = 1 - \frac{\frac{\varepsilon^2}{8} (K_2 - 1) - \frac{\varepsilon^3}{16} (K_3 - 1) - \frac{5}{128} \varepsilon^4}{1 + \frac{\varepsilon}{2} - \frac{\varepsilon^2}{8} + \frac{\varepsilon^3}{16} - \frac{5}{128} \varepsilon^4}; \\
 A(\tilde{\mathcal{F}}_0, \tilde{\mathcal{F}}_1) &< \left(1 - \frac{\frac{\varepsilon_0^2 n^{2r-2\delta}}{8} - \frac{\varepsilon_0^3 n^{-3\delta}}{16} (K_3 - 1) - \frac{5\varepsilon_0^4 n^{-4\delta}}{128}}{1 + \frac{\varepsilon_0 n^{-\delta}}{2} - \frac{\varepsilon_0^2 n^{-2\delta}}{8} + \frac{\varepsilon_0^3 n^{-3\delta}}{16} - \frac{5\varepsilon_0^4 n^{-4\delta}}{128}} \right)^n \\
 &= \left(1 - \frac{\frac{\varepsilon_0^2 n^{2r-2\delta}}{8} \left(1 - \frac{\varepsilon_0}{2} \left(\frac{K_3 - 1}{n^{2r+\delta}} \right) - \frac{5\varepsilon_0^2}{16} n^{-2(r+\delta)} \right)}{\underbrace{1 + \frac{\varepsilon_0 n^{-\delta}}{2} - \frac{\varepsilon_0^2 n^{-2\delta}}{8} + \frac{\varepsilon_0^3 n^{-3\delta}}{16} - \frac{5\varepsilon_0^4 n^{-4\delta}}{128}}_{z_n}} \right)^n.
 \end{aligned}$$

For $\{(\delta, r) : r > \delta - \frac{1}{2} \text{ and } K_3 = o(n^{2r+\delta})\}$, $n z_n \rightarrow \infty$ and, consequently, $A(\tilde{\mathcal{F}}_0, \tilde{\mathcal{F}}_1) \rightarrow 0$.

Till now we proved the detectability for the case when rejection threshold of the LRT is $C = 0$, i.e., both α and β change with n . If we fix α , then starting from some large N_0 : $\alpha_{N_0} < \alpha$ and $\beta_{N_0} > \beta(\alpha, f_1)$. Henceforth, as $\beta_n \rightarrow 0$ in R_d , $\beta(\alpha, f_1)$ also tends to zero in R_d , i.e., $\exists N : \forall n \geq N : \beta(\alpha, f_1) \leq \beta_{\max}$. \square

Notice that if $r \geq 0$ and the condition $K_3 = o(n^{2r+\delta})$ is satisfied, $r < \delta/2$. Indeed, $\text{Var}_g[k(X)] = K_3 - K_2^2 \geq 0$. Consequently, $K_3 \geq K_2^2$, and if $r \geq \delta/2$, then $4r \geq 2r + \delta$.

Example 2.3.1. From Table 2.1, we can derive the conditions for $K_3 = o(n^{2r+\delta})$ given $r \geq 0$:

1. *Gaussian mixture:* $K_2 = e^{\mu^2}$, $K_3 = e^{3\mu^2} = K_2^3$. Therefore, $K_3 = o(n^{2r+\delta})$ iff $r < \delta/4$.
2. *Laplace mixture:* $K_2 = \frac{2}{3} e^{\sqrt{2}\mu} + \frac{1}{3} e^{-2\sqrt{2}\mu}$, $K_3 = \frac{3}{5} e^{2\sqrt{2}\mu} + \frac{2}{5} e^{-3\sqrt{2}\mu}$. In order to have $K_3 = o(n^{2r+\delta})$, one needs $r < \delta/2$.
3. *Cauchy mixture.* $K_2 - 1 = \frac{\mu^2}{2}$, $K_3 = \frac{3\mu^2}{2} + \frac{3\mu^4}{8} + 1$. Consequently, $K_3 = o(n^{2r+\delta})$ iff $r < \delta/2$.

Note also that if $K_3 = o(n^{2r+\delta})$, we can approximate β as

$$\beta_n \approx \Phi\left(-\Phi^{-1}(\alpha) - \varepsilon \sqrt{n(K_2 - 1)}\right) = \Phi\left(-\Phi^{-1}(\alpha) - \varepsilon_0 n^{\frac{1}{2} + r - \delta}\right).$$

In the parameterization we put $\varepsilon_0 = -(\Phi^{-1}(\beta_{\max}) + \Phi^{-1}(\alpha))$ to ensure $\beta_n \approx \beta_{\max}$ on the line $r = \delta - \frac{1}{2}$.

Lemma 2.6. If R_d is the region of detectability from Lemma 2.5, the parameter set defined as

$$R_d^C = \{(\delta^*, r^*) : \delta^* \leq \delta \text{ and } r^* \geq r, \text{ for } (\delta, r) \in R_d\}$$

also gives the asymptotic LRT error rates $\alpha_n, \beta_n \rightarrow 0$.

Chapter 2. Detection problems for finite sample sizes

Proof. To prove the statement we will show that $A(\tilde{\mathcal{F}}_0, \tilde{\mathcal{F}}_1)$ is a monotonically non-increasing function of ε and μ .

- The monotonicity in ε follows from

$$\begin{aligned} \left(\frac{1}{\sqrt{1+\varepsilon}} \mathbb{E} \left(\sqrt{1+\varepsilon k(X)} \right) \right)'_{\varepsilon} &= \frac{1}{2(1+\varepsilon)^{3/2}} \mathbb{E} \left(\frac{k(X)-1}{\sqrt{1+\varepsilon k(X)}} \right) \\ &= \frac{1}{2(1+\varepsilon)^{3/2}} \left(\int_{k(x)<1} \frac{k(x)-1}{\sqrt{1+\varepsilon k(x)}} dx + \int_{k(x)\geq 1} \frac{k(x)-1}{\sqrt{1+\varepsilon k(x)}} dx \right) \\ &\leq \frac{1}{2(1+\varepsilon)^{3/2}} \int_{-\infty}^{+\infty} \frac{k(x)-1}{\sqrt{1+\varepsilon}} dx = 0. \end{aligned}$$

- The monotonicity of the Hellinger affinity in μ could be shown as follows

$$\begin{aligned} \left(\mathbb{E}(\sqrt{1+\varepsilon k(X)}) \right)'_{\mu} &= \mathbb{E} \left(\frac{\varepsilon k'(X)}{2\sqrt{1+\varepsilon k(X)}} \right) \stackrel{k'(x)=-\frac{f'_0(x-\mu)}{f_0(x)}}{=} \frac{\varepsilon}{2} \int_{-\infty}^{+\infty} \frac{-f'_0(x-\mu)}{\sqrt{1+\varepsilon \frac{f_0(x-\mu)}{f_0(x)}}} dx \\ &= -\frac{\varepsilon}{2} \int_{\mu}^{+\infty} f'_0(x-\mu) \left(\frac{1}{\sqrt{1+\varepsilon \frac{f_0(x-\mu)}{f_0(x)}}} - \frac{1}{\sqrt{1+\varepsilon \frac{f_0(\mu-x)}{f_0(2\mu-x)}}} \right) dx. \end{aligned}$$

Note that $f'_0(x-\mu) \leq 0$ and $f_0(x) \leq f_0(2\mu-x)$ for $x \geq \mu$. One can prove that there exists $\tilde{x} > 2\mu$, such that $f'_0(\tilde{x}-\mu) < 0$ and $f_0(\tilde{x}) < f_0(2\mu-\tilde{x})$. Because of the symmetry of $f(x)$ and the above considerations, the derivative of $A(\tilde{\mathcal{F}}_0, \tilde{\mathcal{F}}_1)$ with respect to μ is negative.

□

Three last lemmas in this section prove Theorem 2.2. In Fig. 2.3, we depict R_d and R_{nd} in the (δ, r) plane for Gaussian, Laplace and Cauchy distributions. For rather large sample sizes, the region filled with green corresponds to $\beta \leq \beta_{\max}$ and in the region filled with red $\beta \geq \beta_{\max}$.

Comparison with the ARW model

In Chapter 1 we mentioned the ARW model for sparse normal signals. To cover all alternatives, Tony Cai et al. [2011] proposed the additional parametrization for the case of dense normal signals. A *dense* signal is characterized by a higher concentration of signals and weaker signal magnitude. The combined parametrization is the following,

$$\begin{aligned} p_n &= n^{-\delta}, & 0 < \delta < 1; \\ \mu_n^{\text{sparse}} &= \sqrt{2r \log(n)}, & 0 < r < 1, \quad \frac{1}{2} < \delta < 1; \\ \mu_n^{\text{dense}} &= n^r, & -\frac{1}{2} < r < 0, \quad 0 < \delta < \frac{1}{2}. \end{aligned} \tag{2.21}$$

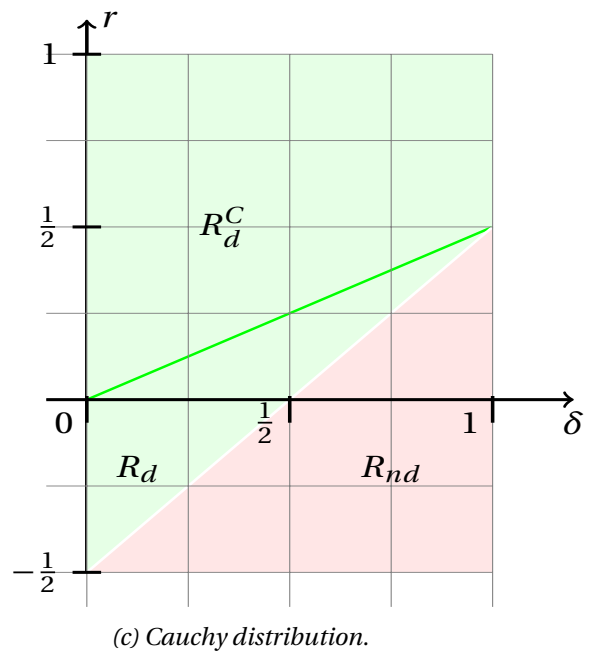
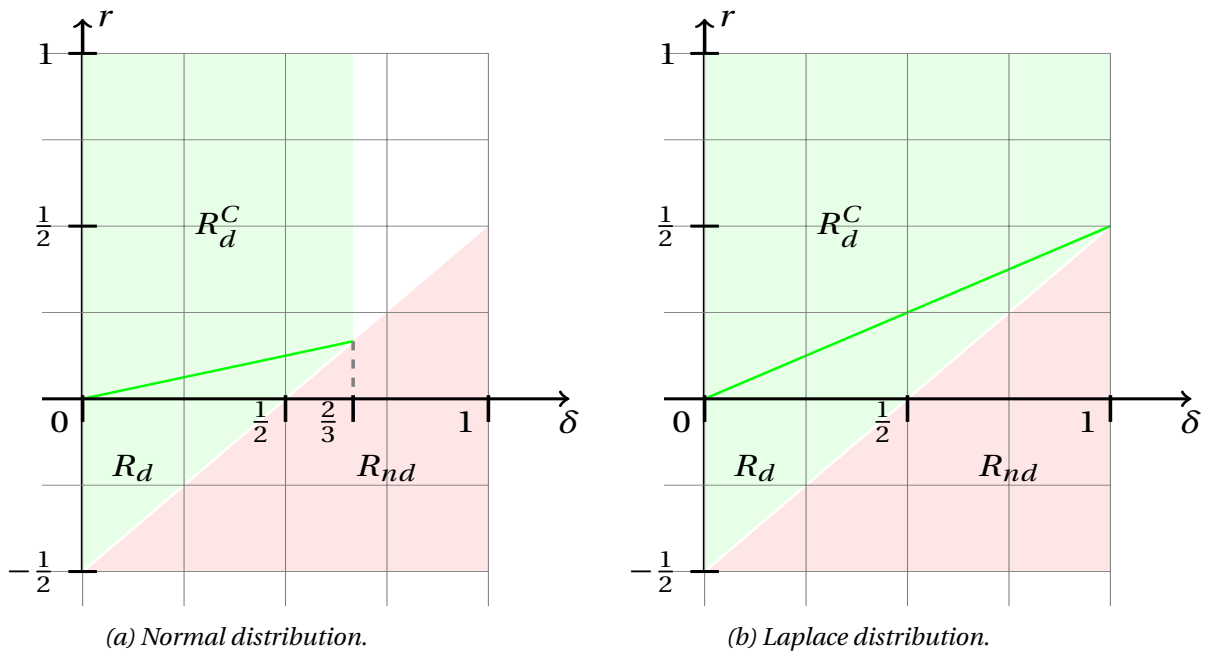


Figure 2.3 – Phase diagram of R_d , R_d^C and R_{nd} . R_d and R_d^C regions with $\beta_n < \beta_{\max}$ are filled with green, the red region corresponds to R_{nd} with $\beta_n > \beta_{\max}$. For the region filled with white the possibility of asymptotic detection has not been established.

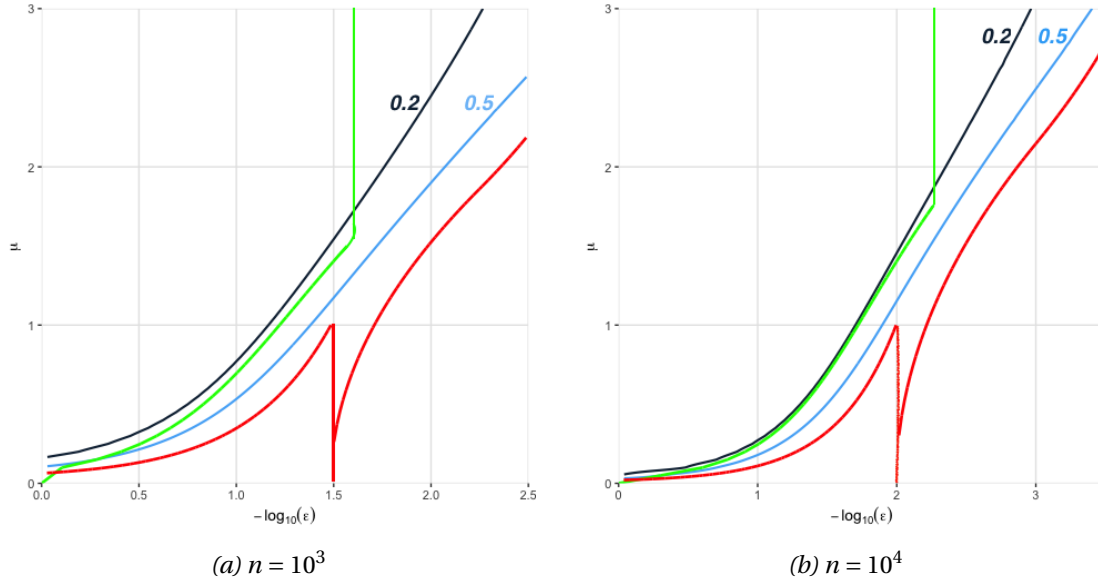


Figure 2.4 – The parameter space for the Gaussian mixture model $(1 - p)\mathcal{N}(0, 1) + p\mathcal{N}(\mu, 1)$. The blue line corresponds to the numerical contour lines for the type II error rate of the LRT; the detection boundary from (2.22) is shown in red; the region R_d is inside the area restricted by the green lines. The default parameters of the test are chosen to satisfy $\alpha = 0.05$, $\beta_{\max} = 0.2$.

The asymptotic detection boundary in the case of Gaussian mixtures was found to be

$$\rho^*(\delta) = \begin{cases} \delta - \frac{1}{2}, & 0 < \delta \leq \frac{3}{4}; \\ \left(1 - \sqrt{1 - \delta}\right)^2, & \frac{3}{4} < \delta < 1. \end{cases} \quad (2.22)$$

If $r > \rho^*(\delta)$, the sum of the type I and type II error rates of the LRT with the rejection threshold $C = 0$ tends to zero, which is referred to as the detectable region, while if $r < \rho^*(\delta)$, the sum tends to one for any testing procedure (non-detectable region). The findings were illustrated in simulations with the LRT, the Higher Criticism test, the sample mean and the maximum value testing procedures. However, even sample sizes of order 10^7 were not enough to show a clear distinction between the region of detectability and the region of non-detectability. For many applications the asymptotic results are not satisfactory, because one should often plan the experiment and estimate the errors for the chosen sample size. Our motivation was to explore the boundaries of the detection for finite sample sizes and predefined maximum error rates levels. To check how far the asymptotic boundary is from the one obtained numerically, in Fig. 2.4(a), (b) we show the power of the LRT for $n = 10^3, 10^4$. The contour lines of the type II error rate are plotted in blue, and the boundary corresponding to (2.22) is plotted in red. The detection region R_d corresponds to the area restricted by the green lines. We can see that for these sample sizes the boundary defined in (2.22) gives an overly optimistic guidance, resulting in power of less than 50%. Moreover, at $\delta = 1/2$ a discontinuity occurs, which makes the results not consistent, while our parametrization is a continuous mapping of r into μ .

For Cauchy and Laplace mixtures we do not have any references to compare, and we check

empirical power in R_d and R_{nd} based on simulations. For each point of the (ε, μ) plane, the mixture from the corresponding distribution was generated, where the proportion of points from the component with a location parameter μ is $p = \frac{\varepsilon}{1+\varepsilon}$. The results are depicted in Fig. 2.5. Apart from the LRT test, we applied the rejection procedures based on the sample mean (SM) and Higher Criticism. Rejection bounds were chosen as upper α quantiles of the corresponding test statistics under the null. One can see that there is a good correspondence of the theoretical detection boundary and the empirical one (the boundary of the light green area and the red one) starting from sample sizes of order 10^2 . It is smaller than for the Gaussian mixtures where the empirical contour line $\beta = 0.2$ is close to the theoretical boundary starting from sample size of order 10^3 (see Fig. 2.4). Interestingly, Higher Criticism does not seem to be optimal, either for Laplace or for Cauchy mixtures. Conversely, we observe that the gap between the detection boundary and the HC empirical power does not shrink when the sample size is increased. For Laplace mixtures the sample mean converges to the detection boundary and has satisfactory power for large ε .

In general, when comparing Gaussian mixtures with Laplace and Cauchy mixtures, one can see how much more difficult the detection of signal is for heavy-tailed distributions.

2.4 Testing procedures for Gaussian mixtures

The parametric LRT is not a useful tool if the distribution is unknown. In this section we compare the performance of several non-parametric tests for the case of Gaussian mixtures. All the results for the univariate case are presented in Fig. 2.6.

2.4.1 Univariate Gaussian mixture

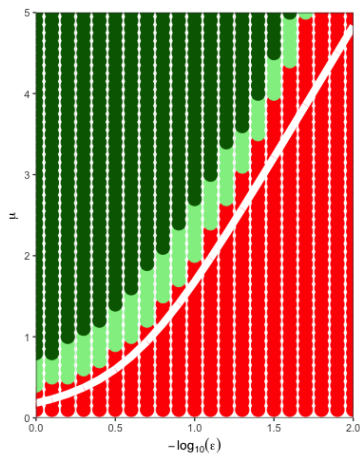
- *Higher Criticism test.* Transform each observation X_i into the p-value $p_i = \mathbb{P}(\mathcal{N}(0, 1) > X_i)$, $i = 1, \dots, n$. Then sort the p-values in an increasing order, $p_{(1)} \leq p_{(2)} \leq \dots \leq p_{(n)}$. The final statistic is the maximum standardized deviation between the empirical distribution of the p-values and their expected distribution under the null, the uniform $\mathcal{U}[0, 1]$,

$$\text{HC}(p_{(i)}) = \frac{\sqrt{n}|i/n - p_{(i)}|}{\sqrt{p_{(i)}(1 - p_{(i)})}}, \quad \text{HC}_n^* = \max_{1 \leq i \leq n} \text{HC}(p_{(i)}) = \text{HC}(p^*).$$

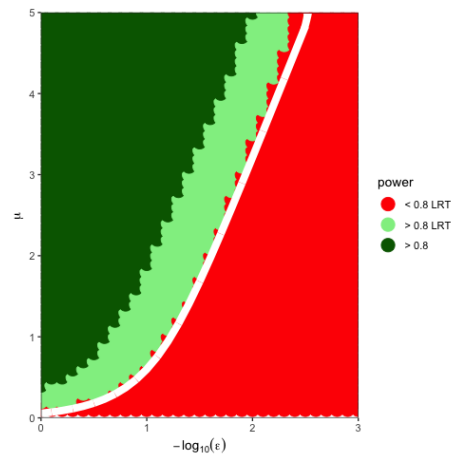
It was shown in Donoho and Jin [2004], Tony Cai et al. [2011] that HC is optimally adaptive in the detectable region $r > \rho^*(\delta)$ (see Eq.2.22).

We simulated HC performance for $n = 100$ and $n = 1000$ random variables from Gaussian mixtures with different (p, μ) . Rejection bounds for the test corresponding to $\alpha = 0.05$ were computed with 10^4 Monte-Carlo simulations: $\text{HC}_{100}^*(0.05) = 4.8$, $\text{HC}_{1000}^*(0.05) = 4.83$.

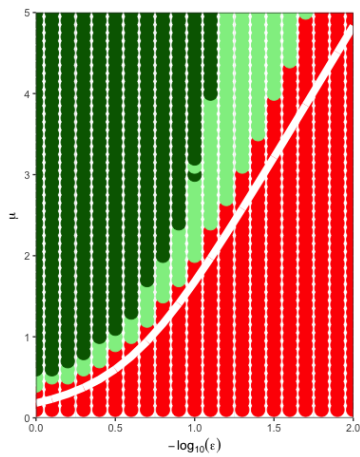
- *The maximum value.* The next non-parametric statistic we look at is the maximum of the observations. If we continue with the multiple testing analogy, the test based on



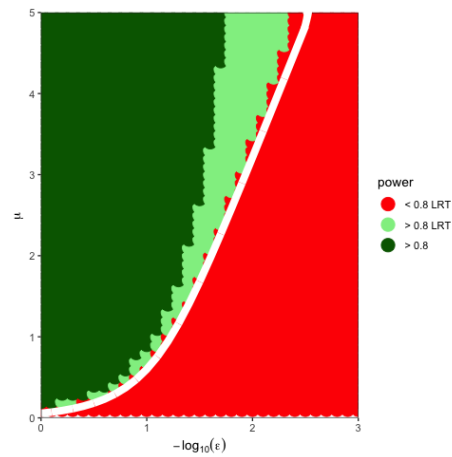
(a) Laplace mixture. The HC test $n = 10^2$



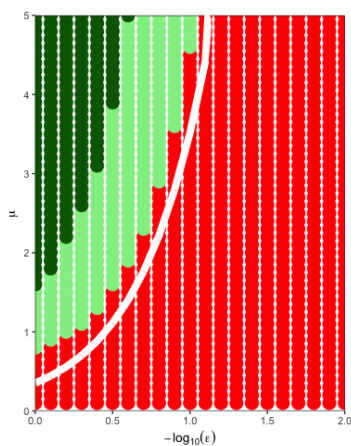
(b) Laplace mixture. The HC test, $n = 10^3$



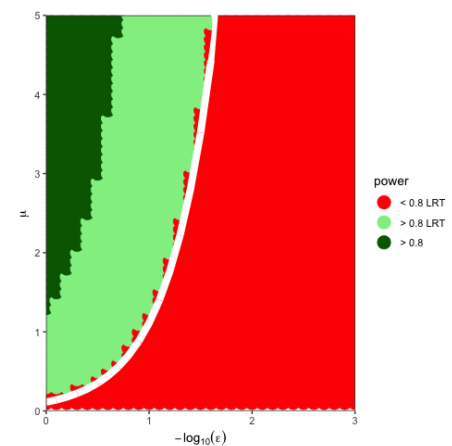
(c) Laplace mixture. The SM, $n = 10^2$



(d) Laplace mixture. The SM, $n = 10^3$



(e) Cauchy mixture. The HC test, $n = 10^2$



(f) Cauchy mixture. The HC test, $n = 10^3$

Figure 2.5 – Empirical detection power based on 10^3 simulations from the mixtures with the parameters (ϵ, μ) . The dark green area corresponds to the power of the corresponding statistic greater than 0.8; the light green area - to the power of the LRT greater or equal to 0.8. The white line corresponds to $r = \delta - \frac{1}{2}$. All tests are performed to ensure the maximum error levels, $\alpha = 0.05$, $\beta_{\max} = 0.2$.

the maximum value can be considered as the analogue of the Bonferroni correction (the minimal p-value converts to the maximal observation). Consequently, one would expect it to be more conservative than HC.

Under the null, $\frac{\max_{i=1,\dots,n} X_i}{\sqrt{2 \log(n)}} \xrightarrow{p} 1$. So, it is natural to take $\sqrt{2 \log(n)}$ as a rejection boundary to let the type I error rate tending to zero. In the case of the Gaussian mixture, the maximum value is the maximum of two components,

$$\max_{i=1,\dots,n} X_i \approx \sqrt{2 \log(n)} \cdot \max\{1, \sqrt{1-\delta} + \sqrt{r}\}.$$

Therefore, the asymptotic separation of \mathcal{H}_0 and \mathcal{H}_1 is only possible if $\sqrt{1-\delta} + \sqrt{r} > 1$. For this reason the usage of the maximum value statistic is restricted to the sparse case where the detection boundary is defined by $\rho_{\max} = (1 - \sqrt{1-\delta})^2$. It follows that the maximum value statistic is less powerful than the HC for $\delta \in (1/2, 3/4]$ and $r \geq 1$ and coincides with the HC for $\delta \in (3/4, 1)$, $r \geq 1$. In Fig. 2.6(a),(c),(b),(d) one can see for the chosen sample sizes there is a negligible distinction for *sparse* signals, whereas the distinction for the *dense* signals becomes substantial with bigger n .

For our analysis with a fixed type I error rate, the rejection bound corresponding to $\alpha = 0.05$ is determined by $x_\alpha = \Phi^{-1}(\sqrt[1-\alpha]{1-\alpha})$.

- *The sample mean.* $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ could be a reasonable choice for big ε . From the central limit theorem, the approximation of the type II error rate for big n is

$$\beta \approx \Phi\left(\frac{z_{1-\alpha} - \mu p \sqrt{n}}{\sqrt{1 + (1-p)p\mu^2}}\right), \quad (2.23)$$

where $p = \frac{\varepsilon}{1+\varepsilon}$ and $z_{1-\alpha} = \Phi^{-1}(1-\alpha)$.

Consider first the case when $r > 0$. The asymptotic limit of (2.23) is

$$\beta \approx \Phi\left(\frac{z_{1-\alpha} - \varepsilon_0 \sqrt{2r \log(n)} n^{-\delta+1/2}}{\sqrt{1 + \varepsilon_0 2r n^{-\delta} \log(n)}}\right) \xrightarrow{n \rightarrow \infty} \begin{cases} 1 - \alpha, & \delta > 1/2, \\ 0, & \delta < 1/2. \end{cases}$$

For the *dense* region with $\delta \in (0, 1/2)$ and $r < 0$, $\beta \approx \Phi(z_{1-\alpha} - \varepsilon_0 n^{\frac{1}{2}+r-\delta})$, and it follows that

$$\beta \xrightarrow{n \rightarrow \infty} \begin{cases} 1 - \alpha, & r < \delta - 1/2, \\ 0, & r > \delta - 1/2 \end{cases}.$$

On the line $r = \delta - 1/2$ one has $\beta \approx \beta_{\max}$, i.e., the sample mean is asymptotically equivalent to the LRT in the dense region.

Hence, to detect the mixture with power at least $1 - \beta_{\max}$, the smallest possible number of points containing the signal should be $\varepsilon_0 n^{-1/2}$ (the amplitude of this signal is $\sqrt{\log 2}$). For $\alpha = 0.05$, $\beta = 0.2$ and $n = 100$ and $n = 1000$ the value of $\varepsilon_0 n^{-1/2}$ is 25 and 79,

sample size, n	min # of points in the second group	s	min μ detected in R_d	$\mu\left(\varepsilon_0 n^{-\frac{2}{3}}\right)$
100	10	2	0.31	0.93
1000	24	2	0.22	1.09
100	10	3	0.26	0.76
1000	24	3	0.18	0.89
100	10	5	0.20	0.59
1000	24	5	0.14	0.69
100	10	7	0.17	0.50
1000	24	7	0.12	0.59
100	10	10	0.14	0.42
1000	24	10	0.10	0.49
100	10	15	0.11	0.34
1000	24	15	0.08	0.40

Table 2.2 – Some of the detection limiting characteristics for a multivariate gaussian mixture with sparse mean value. The preserved type I and type II error rates are 0.05 and 0.2 respectively, and $\varepsilon_0 \approx 2.49$.

respectively.

2.4.2 Multivariate Gaussian mixture

For completeness, we consider the alternative hypothesis in (2.1) to be the d -variate mixture

$$\mathbf{X}_i \sim \frac{1}{\varepsilon + 1} \mathcal{N}_d(\mathbf{0}, \Sigma) + \frac{\varepsilon}{\varepsilon + 1} \mathcal{N}_d(\boldsymbol{\mu}, \Sigma) \quad , i = 1, \dots, n; \quad \mu_j \in \{0, \mu\}, j = 1, \dots, d.$$

Denote the number of $\mu_j \neq 0$, $j = 1, \dots, d$ as s . In the multivariate normal case the moments of $k(x)$ are

$$K_2 = e^{\boldsymbol{\mu}^T \Sigma^{-1} \boldsymbol{\mu}}, \quad K_3 = e^{3\boldsymbol{\mu}^T \Sigma^{-1} \boldsymbol{\mu}}.$$

The parameterization of μ (the y-axis in Fig. 2.4a, 2.4b) in the multivariate normal case will be changed by the scaling factor $\frac{1}{\sqrt{s}}$. The more true non-null components are in $\boldsymbol{\mu}$, the easier it is to detect a mixture. The minimum theoretical ε in R_d is defined by $\varepsilon_0 n^{-\frac{2}{3}}$ with the corresponding effect of the size $\mu\left(\varepsilon_0 n^{-\frac{2}{3}}\right) = \sqrt{\log\left(1 + n^{\frac{1}{3}}\right) / s}$. The resulting values for $n = 100, 1000$ are summarized in Tab. 2.2.

To check the detection for multivariate Gaussian mixtures, we consider two mean-based tests and two variance-based tests:

- *Hotelling's test*. The multivariate analogue of the squared standardized mean-value. If

2.4. Testing procedures for Gaussian mixtures

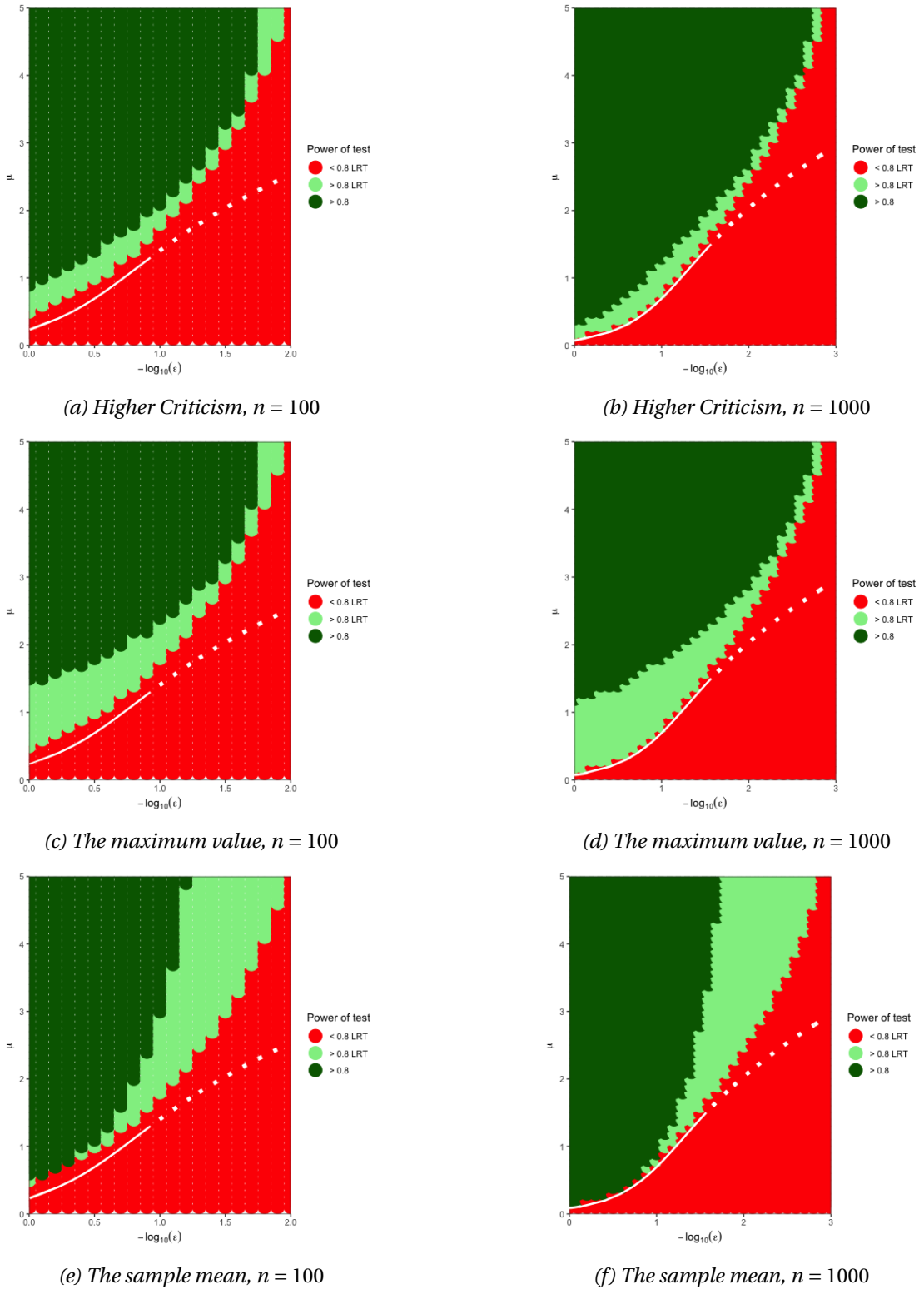


Figure 2.6 – Empirical power based on 10^3 simulations from the mixture $\frac{1}{\varepsilon+1} \mathcal{N}(0, 1) + \frac{\varepsilon}{\varepsilon+1} \mathcal{N}(\mu, 1)$. The dark green area corresponds to the powers greater than 0.8 of the explored statistic; the light green area is the region where the power of the LRT is greater or equal to 0.8. The white line corresponds to $r = \delta - \frac{1}{2}$. $\alpha = 0.05$, $\beta_{\max} = 0.2$.

the null is true and $\mathbf{x} \sim \mathcal{N}_d(\boldsymbol{\mu}, \Sigma)$, then

$$n(\bar{\mathbf{x}} - \boldsymbol{\mu})^T \Sigma^{-1}(\bar{\mathbf{x}} - \boldsymbol{\mu}) \sim \chi_d^2.$$

In our case $\boldsymbol{\mu} = \mathbf{0}$ and $\Sigma = \mathbf{I}$. The rejection region is chosen according to the $(1 - \alpha)$ quantile of $\chi^2(d)$ distribution.

- *Coordinatewise sample mean with the multiplicity correction.* The null hypothesis for the coordinatewise mean-based methods is $\mathcal{H}_0 : \mu_1 = \dots = \mu_d = 0$. Denote $\mathcal{H}_{0i} : \mu_i = 0, i = 1, \dots, d$. The method consists of two steps. First, the individual p-values for each coordinate's sample mean are computed. Then, one applies the Hochberg multiplicity correction, which places the p-values in ascending order, $p_{(1)} \leq p_{(2)} \dots \leq p_{(d)}$. The method controls the probability of at least one false positive among the set of the p-values, i.e., the family wise error rate (FWER), to be less than a certain level and proceeds as following: for the maximum rank k such that $k : p_{(k)} \leq \frac{\alpha}{d+1-k}$ the method rejects the null hypotheses corresponding to the p-values with ranks $i \leq k$. For Hochberg's method to control FWER, test statistics need to be independent or have a distribution of a certain positive dependence structure, see Sarkar [1998]. For $\Sigma = \mathbf{I}$, as in our simulations, the p-values are independent. We count how many times at least one of the null hypotheses is rejected, in this case we consider that the alternative was detected.

The next two testing methods that we consider below were proposed for the sparse detection problem in the recent work Verzelen and Arias-Castro [2017]. They explored minimax detection rates in terms of the signal-to-noise ratio $\boldsymbol{\mu}^T \Sigma^{-1} \boldsymbol{\mu}$ for the fixed ε . For different sparsity regimes they claimed that methods based on the sample variance are minimax optimal ($\gamma_n \rightarrow 0$, see Eq. 1.4) and at the same time computationally tractable. For a sparsity regime with $s \leq \sqrt{d/\log(d)}$, it is shown that if the parameters satisfy

$$p(1-p)\|\boldsymbol{\mu}\|^2 \gg \sqrt{\frac{s^2}{n} \log(d/s)}, \quad (2.24)$$

the statistic $\max_{j=1, \dots, d} \hat{\sigma}_{jj}$, where $\hat{\sigma}_{jj} = \frac{1}{n} \sum_i (x_{ij} - \bar{x}_j)^2$ is consistent. For the non-sparse regime, $s \geq \sqrt{d/\log(d)}$, if

$$p(1-p)\|\boldsymbol{\mu}\|^2 \gg \sqrt{\frac{d}{n}}, \quad (2.25)$$

the consistent test (in terms of minimax error tending to zero with the sample size tending to infinity) is based on thresholding the top eigenvalue.

- *Coordinatewise sample variance with the multiplicity correction.* Hochberg's multiplicity correction procedure was applied for each coordinate's sample variance. We found this method to be more powerful than the maximum canonical variance, due to the link between the maximum canonical variance and the overly conservative Bonferroni correction.

2.5. Reciprocal questions. Estimability and classification. Multivariate sparse mixture models.

- *The maximum eigenvalue.* Here we take the top eigenvalue of the sample's empirical covariance matrix. The critical value is obtained by 10^4 Monte-Carlo simulations under the null.

As expected, the power of the mean-based tests is higher for large values of ε and small values of μ , while the variance-based tests are more powerful for small values of ε (see Fig. 2.7, 2.8). The difference between the two mean-based tests is small for the considered sample sizes and dimensionality. Conversely, the distinction between the canonical variance and the top eigenvalue is rather big, especially between the μ is 1-sparse ($s = 1$) and d -sparse ($s = d$) cases.

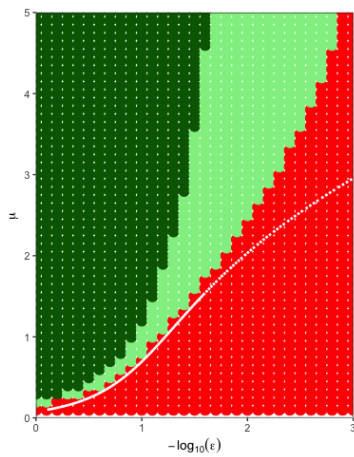
2.5 Reciprocal questions. Estimability and classification. Multivariate sparse mixture models.

The next step after the detection is usually estimation of the parameters. For the univariate gaussian case and parametrization (2.21), the estimation of ε_n and μ_n is discussed in Meinshausen and Rice [2006], Cai et al. [2007], Jin and Cai [2007]. Consistent estimator of p when $\delta > 1/2$ requires additional restrictions on the model, while if $\delta \in (0, 1/2)$, the estimation procedure remains general. After the estimates are obtained, the classification can be done by maximizing the posterior probabilities, see Sun and Cai [2007]. The authors also showed that the misclassification error, the percentage of falsely assigned observations, tends to the random guessing error, $n\varepsilon_n$, for any classification procedure if $r < \delta$.

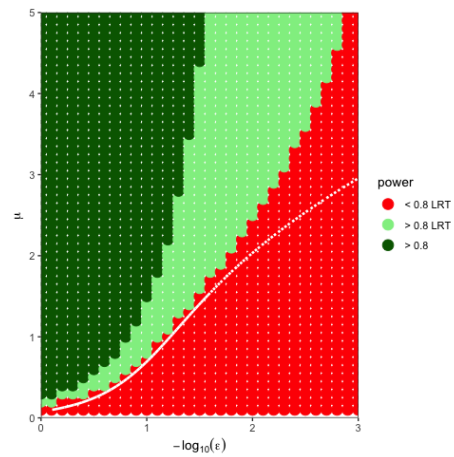
The natural transition to the multivariate case ($\text{dim} = d$) was proposed in Donoho and Jin [2008, 2009], Jin [2009]. They considered two groups of labeled multivariate observations of equal size n . Assuming the coordinatewise effect sizes to be independent, they fixed an effect size to be equal in absolute value and differ in sign for the two groups. Then the data reduced to d statistics, each of which is the mean difference between the groups in each of d dimensions. Thus, the multivariate detection problem is transformed into a univariate problem by the substitution of the sample size n with the dimension d . The proportion of signals associated with a groups' difference, ε and their strength μ are parametrized with regard to dimension similarly to (2.21). In addition, Donoho let the sample size n depend on three different growth regimes: no growth, slow and regular growth. A similar model is explored in the minimax framework in Ingster et al. [2009]. The parametrization above was motivated by the explosion of gene-wide association studies (GWAS), where the number of subjects is much smaller than the extracted genetic information represented by single nucleotide polymorphisms (SNPs). The sparsity of μ was also a natural choice in the early GWAS era. Then, researchers expected to find very few SNPs associated with a phenotype. However, today the paradigm has changed, and many studies claim that a more reasonable strategy is to search for the sets of genes that could explain the phenotype variability, rather than concentrating on single gene candidates.

The methods, such as the top eigenvalue and the maximum canonical variance, which

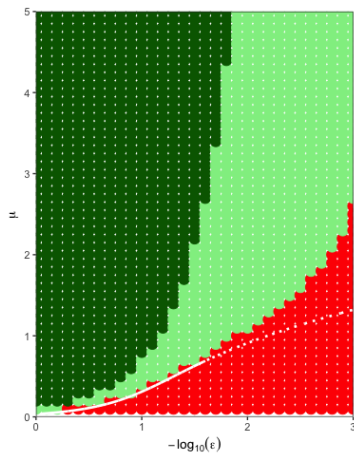
Chapter 2. Detection problems for finite sample sizes



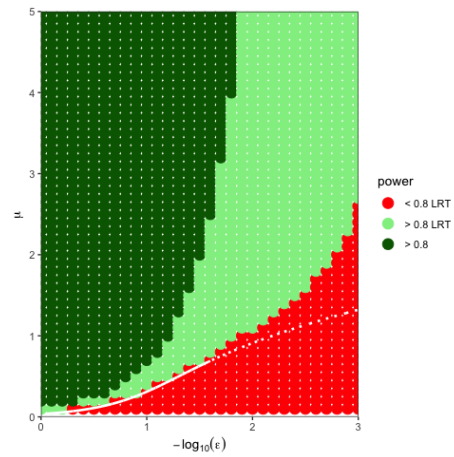
(a) The Sample mean & Hoch. correction. $s = 1$



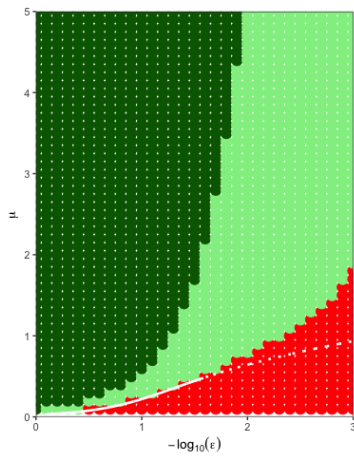
(b) Hotelling's test. $s = 1$



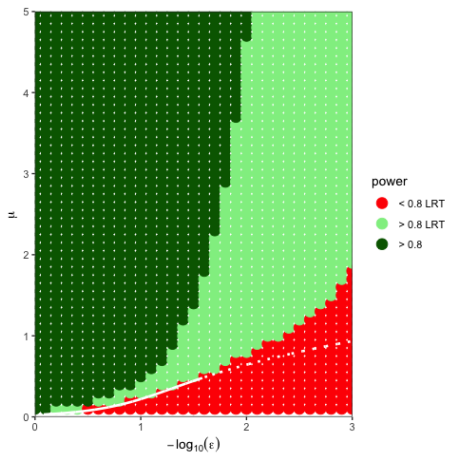
(c) The sample mean & Hoch. correction. $s = 5$



(d) Hotelling's test. $s = 5$



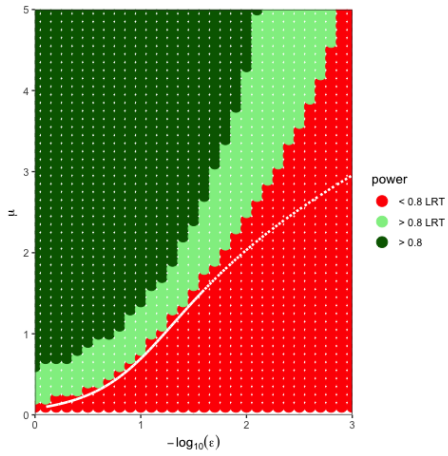
(e) The sample mean & Hoch. correction. $s = 10$



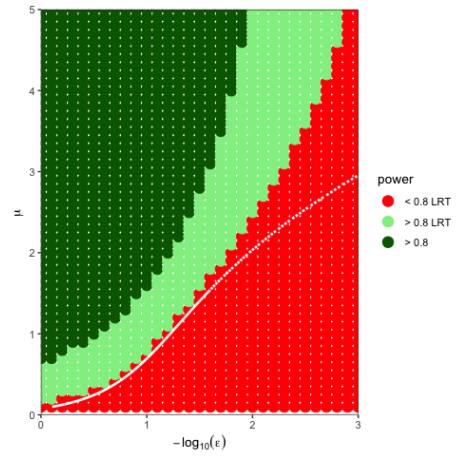
(f) Hotelling's test. $s = 10$

Figure 2.7 – Empirical power for $X \in \mathbb{R}^{10 \times 1}$ under the $\frac{1}{\varepsilon+1} \mathcal{N}(\mathbf{0}, \Sigma) + \frac{\varepsilon}{\varepsilon+1} \mathcal{N}(\boldsymbol{\mu}, \Sigma)$. The dark green area corresponds to the power of the corresponding statistic greater than 0.8; the light green area - to the power of the LRT greater or equal to 0.8. The white line corresponds to $r = \delta - \frac{1}{2}$. $\alpha = 0.05$, $\beta_{\max} = 0.2$, $n = 10^3$.

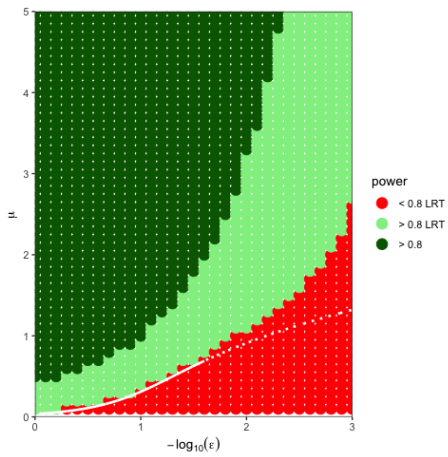
2.5. Reciprocal questions. Estimability and classification. Multivariate sparse mixture models.



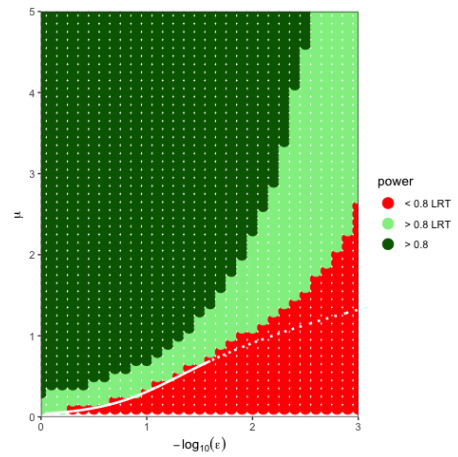
(a) The sample var. & Hoch. correction. $s = 1$



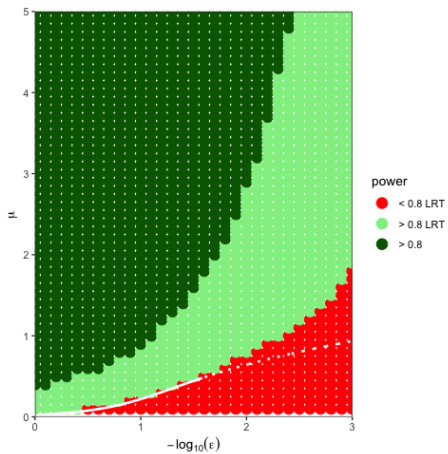
(b) The maximum eigenvalue. $s = 1$



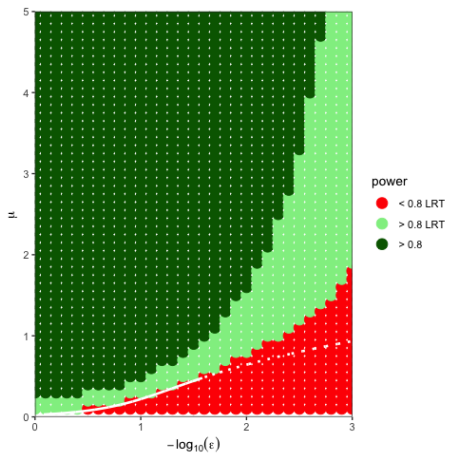
(c) The sample var. & Hoch. correction. $s = 5$



(d) The maximum eigenvalue test. $s = 5$



(e) The sample var. & Hoch. correction. $s = 10$



(f) The maximum eigenvalue test. $s = 10$

Figure 2.8 – Empirical power for $\mathbf{X} \in \mathbb{R}^{10 \times 1}$ under the $\frac{1}{\varepsilon+1} \mathcal{N}(\mathbf{0}, \Sigma) + \frac{\varepsilon}{\varepsilon+1} \mathcal{N}(\boldsymbol{\mu}, \Sigma)$. The dark green area corresponds to the power of the corresponding statistic greater than 0.8; the light green area - to the power of the LRT greater or equal to 0.8. $\alpha = 0.05$, $\beta_{\max} = 0.2$, $n = 10^3$.

were taken to illustrate the detection of sparse multivariate mixtures detection are based on Verzelen and Arias-Castro [2017]. For d -variate Gaussian mixtures they discuss the test

$$\begin{aligned} \mathcal{H}_0 : \mathbf{X} &\stackrel{iid}{\sim} \mathcal{N}(\boldsymbol{\mu}_0, \Sigma) && \text{versus} \\ \mathcal{H}_1 : \mathbf{X} &\stackrel{iid}{\sim} p \mathcal{N}(\boldsymbol{\mu}_0, \Sigma) + (1-p) \mathcal{N}(\boldsymbol{\mu}_1, \Sigma), \end{aligned}$$

where $\Delta\boldsymbol{\mu} = \boldsymbol{\mu}_1 - \boldsymbol{\mu}_0$ is s -sparse and p is fixed. The sparsity s is assumed to be known, and the parameters $s, d, \Delta\boldsymbol{\mu}$ might depend on n . The authors give an extensive instruction to what should be the minimum distance between the means of the mixture components, so that the asymptotic minimax risk tends to zero with $n \rightarrow \infty$. In addition, they suggested that coordinatewise sample moments are consistent (assuming separation between $\mathcal{N}(\boldsymbol{\mu}_0, \Sigma)$ and $\mathcal{N}(\boldsymbol{\mu}_1, \Sigma)$) estimators of $\Delta\boldsymbol{\mu}$ support.

2.6 Summary and conclusions

The detection problem for gaussian mixtures has been extensively studied. For the parametrization where the signal depends on the sample size n , the asymptotic detection rates were found for $n \rightarrow \infty$ along with the optimal procedures that achieve this boundary. However, the limits of the type I and type II error rates are not enough when one needs to plan the experiment. There should be a guidance for practitioners whether their sample size is sufficient for detection of the signal of interest with the desired accuracy. To approach this problem for the non-asymptotic setting, we used the moments of the log likelihood ratio of the densities of the given distributions under the null and under the alternative hypotheses.

Our first finding is that an approach straightforwardly leads to the asymptotes for the sample size when the weight of the non-null mixture component $p \rightarrow 0$. This result can be applied for a broad range of distributions. The second result is for the case when the mixture is parametrized in terms of the sample size n with two parameters (δ, r) . The first parameter accounts for the proportion and the second represents the effect size of the non-null signals. It is convenient to depict the power of testing procedures on the (δ, r) plane. For mixtures of two continuous symmetrical and homoscedastic distributions we found a way for splitting the (δ, r) plane into regions where the type II error rate of the LRT test is above or below some predefined β_{\max} , while keeping the level α of the test fixed. These regions have linear bounds and allow easy inference, whether or not the sample size is sufficient to detect the alternative with a desired power with the parameters within the regions. For some mixtures, e.g., Gaussian, with the theory employed, we are not able to cover the whole parameter plane. To validate the results and to fully reconstruct the regions' bounds, we provide the numeric calculation of the LRT performance using the moments of the likelihood ratio.

The LRT is a parametric test and not always applicable. For this reason, we checked the performance of the most widely used testing procedures for the univariate and multivariate Gaussian mixtures. With emphasis on planning and sample size computation, we conclude

that for dense signals (small magnitude / high frequency) the first moment statistics are preferred both for univariate and multivariate cases. For sparse signals (big magnitude / low frequency) the HC and the maximum value give almost the same detection region in the univariate case, while in the case of multiple dimensions either the methods aggregating the coordinatewise second moments or the the maximum eigenvalue should be applied depending on the sparsity of the components' mean difference.

3 Clinical trials

Because of population heterogeneity, treatment effects may vary considerably from one subject to another. In Evans et al. [2004], it has been reported that on average only 60% of treated patients react to prescription drugs, whereas the others either do not benefit or have adverse effects. If a trial showed no significant *average treatment effect*, it can be wrong to conclude that the treatment effect does not exist and that the treatment is futile. There may well be a subpopulation for whom the treatment is effective. As an example, consider the Clomethiazole Acute Stroke Study by Wahlgren et al. [1999] where the relative functional independence of patients with acute hemispheric stroke was investigated. The trial results showed no significant difference between the clomethiazole and the placebo groups (56.1% vs. 54.8%) with a p-value of 0.649. However, in the group of patients with TACS (total anterior circulation syndrome), 40.8% reached relative functional independence in contrast to 29.8% of TACS patients in the placebo group, with a p-value of 0.008.

In this chapter we discuss the models that are conventionally used to design clinical trials and the situations when they might decrease the power leading a trial to failure. We explain why the mixture models are reasonable choice in the case of partial effectiveness of the tested treatments, and how the results from Chapter 2 could be used in the design.

3.1 Clinical trials

The necessary testing procedure before a new drug or therapy is approved to enter the market, a clinical trial, consists of three phases (Fig. 3.1). According to the U.S. FDA instructions, the Phase I trial is an in vitro or a small-scale experiment conducted to assess the safety and beneficial dosage of a new agent. If the drug has passed to Phase II, safety issues and side effects are reviewed on a larger scale with up to several hundred patients from the target population. Based on these results and the preliminary evidence of the treatment's efficacy the decision whether the drug development should continue is made. Next, during a long and large Phase III trial the developer should finalize safety precautions and prove the new product's benefit for the target population.

Chapter 3. Clinical trials

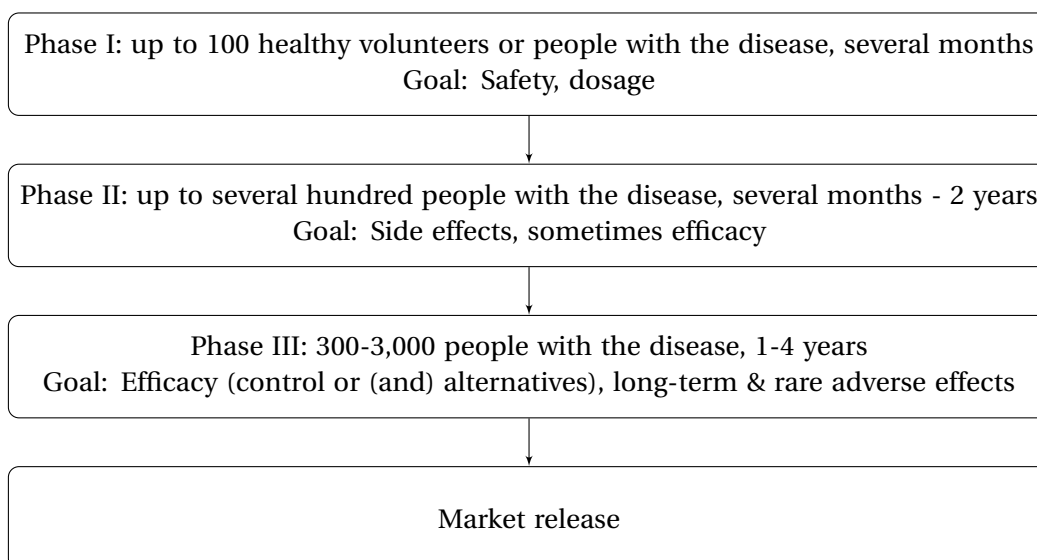


Figure 3.1 – Description of clinical trial phases according to FDA guidelines.

Despite diligent planning and validation *in vitro* and/or in animals, the majority of trials fail to prove efficacy in patients. According to Wong et al. [2017], only 13.8% of new agent's trials pass Phase III and reach the market. The authors collected data from multiple sources over the period 2000–2015 and evaluated the probability of success (POS) for phase transitions and for the clinical trials themselves across indications (see Fig. 3.2). Although success rates increase, they remain very low, causing huge financial and time losses to the pharmaceutical industry. In Fig. 3.3, the failure statistics over the period 2008–2015 collected by Harrison [2016] show that lack of efficacy is a leading reason.

The efficacy of a new drug is usually demonstrated in comparison with some control population. There are three main types of controls. In a no-treatment controlled trial, treated patients are compared to those that do not receive any medication. In placebo-controlled trials, patients in the control group receive medications without active ingredients for a treated condition (sugar pills or saline injections). When new agent is compared to some recognized effective agent, it is referred to as an active-control. The former type of controls only account for the natural course of the disease. The latter type of control may pose problems, since the absence of a difference between the average outcomes in the treatment and the control groups may be due to the ineffectiveness of both treatments or to a poorly designed study. This design is only valid if the conditions of the trial where the active-control outperformed the placebo are recreated, which is a difficult task. For these reasons, the “gold standard” is the double-blind, randomized, placebo-controlled study. There, patients are randomly allocated to control and treatment groups, and neither doctors nor patients know the true memberships.

3.2. Treatment effects

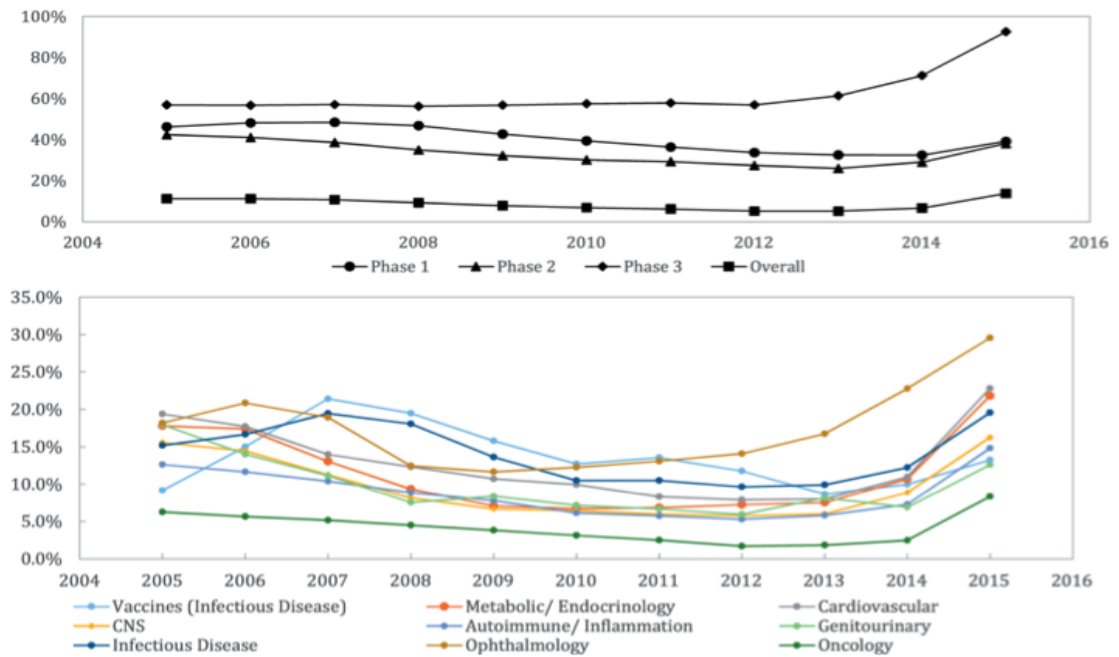


Figure 3.2 – Probability for passing the phases of clinical trial (top) and the overall success rate in the development of new compounds across indications (bottom) according to Wong et al. [2017].

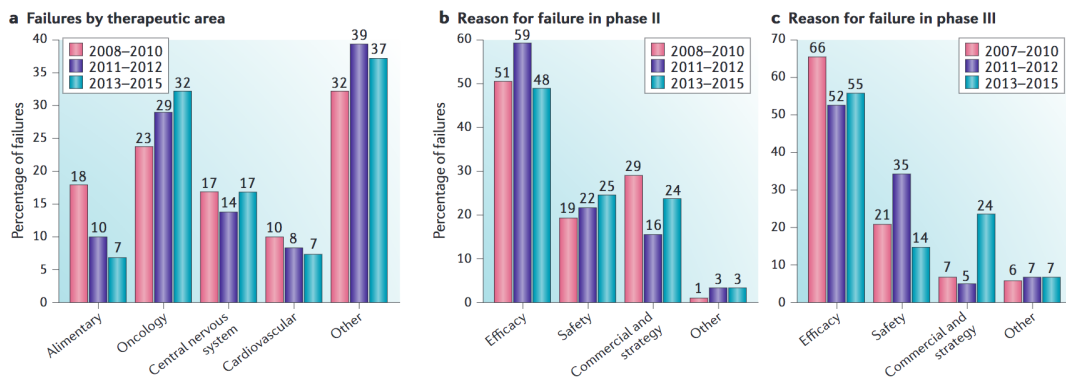


Figure 3.3 – Clinical trials' failure reasons across indications over the period 2008-2015. The data is provided by Thomson Reuters and Drugs of Today, Harrison [2016].

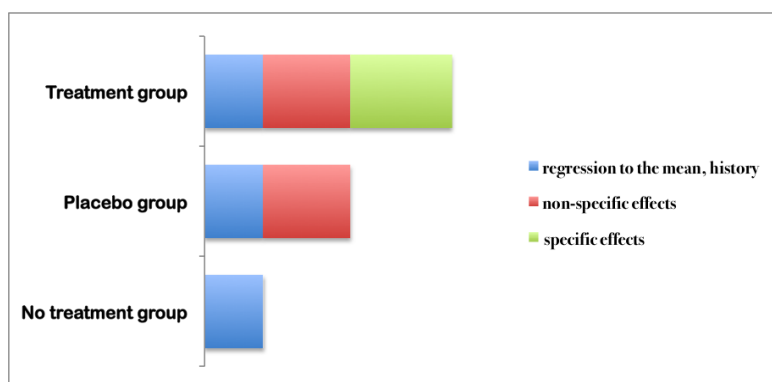


Figure 3.4 – Additivity of the effects in a RCT

3.2 Treatment effects

The improvement in a treatment group is considered to be due to a combination of three types of factors, see Fig. 3.4,: natural factors, which include the natural course of the disease, the Hawthorn effect and regression to the mean; non-specific (or physiologically driven) factors; and specific (attributed to the drug/therapy itself) factors. Control and treatment groups are compared with respect to the average outcome within groups. This means that the additivity and the independence of all three factors contributing to the outcome are fundamental assumptions in concluding the efficacy of new agent. This nowadays conventional model was proposed more than 60 years ago by Beecher [1955]. Although this model remains popular, various critiques appeared in, e.g., Grünbaum [1986], Kleijnen et al. [1994], Kirsch [2000], Walach [2001], Moerman and Jonas [2002], Boussageon et al. [2008], Walach and Loef [2015], and Boehm et al. [2017]. Many of them were concerned about the statistical interactions between the pharmacological and the placebo effects. A correlation of 78% between the placebo and the treatments response rates in 141 long-term trials for different disease categories was reported in Walach et al. [2005]. An even higher correlation was reported in antidepressant trials (see Kirsch and Sapirstein [1998]). There, the specific treatment effect was estimated to be 25% of an overall therapeutic effect.

Strong placebo effect in psychiatric disorders (Khan et al. [2004], Khan et al. [2010]) is another hot spot. Papakostas and Fava [2009] explored whether the probability of receiving a placebo affects the treatment effect, with other design features fixed. In a meta-analysis of 182 RCTs, authors revealed three significant relationships. First, the greater the probability of receiving a placebo (which depends on the number of treatment arms and a randomization scheme), the greater the difference in the response rate between the treatment and the placebo arms. Second, the higher the depression severity, the more likely a patient will respond to a medication rather than to a placebo. And the last relationship, which has been widely discussed, the later the release date of the publication, the smaller the relative efficacy for antidepressants to a placebo.

		Information	
		P	T
Treatment	P	n_{11}	n_{12}
	T	n_{21}	n_{22}

The balanced placebo design (BPD). Usually $n_{11} = n_{12}$ and $n_{21} = n_{22}$.

		Patients							
		1	2	3	4	5	6	7	...
Treatment order	first	P	T	P	P	T	T	T	...
	second	P	T	T	T	T	P	P	...

The balanced cross-over design (BCD).

Table 3.1 – The modifications of designs for the placebo effect estimation.

To summarize, the treatment effect in an RCT is probably not as simple as it is generally considered to be, see Fig. 3.4, at least for some treated conditions. The reported confounding variables and non-linear relationships affecting the treatment response should be accounted for in trial designs.

3.3 Trial designs

As mentioned above, recent findings point out that the beliefs (or expectations) of the enrolled patients affect the difference in the placebo-treatment response rates. One way to evaluate a “pure” treatment effect and to check an additivity assumption is to manipulate the information provided to patients about the treatment. As an example, take the balanced placebo design (BPD), after participants have been randomly allocated to treatment and control groups, in each group some patients are correctly informed about their treatment, while the others are misinformed. The placebo effect is estimated based on the outcome in the misinformed group of patients who received a placebo. Conversely, the treatment effect is measured among the misinformed group receiving the active treatment. Another example is balanced cross-over designs, which are used to test several treatments on the same patients. There, participants are divided into four groups, two of which follow a standard protocol, changing placebo and the active treatment, while the remaining two groups receive either placebo or drug in both treatments. Based on the observations from the two latter groups, the expectation components in placebo and treatment effects are estimated.

Although the BPD and the BCD are easily implemented, they have many unwanted side-effects, distorting the desired estimates. To avoid possible problems, randomization-to-randomization (R2R) designs where patients are randomly given a random probability of group assignment are employed. The estimation of the effects due to the active treatment and placebo are then computed conditionally on the individual’s probability.

Another approach is to use designs where the placebo effect is deliberately decreased and the treatment effect is magnified. These designs (sometimes referred to as enrichment designs) usually involve some type of inclusion/exclusion criterion. It can be pre-testing with only a placebo arm or sequential schemes with an elimination of the placebo responders at each stage of the trial and subsequent re-randomizing the selected participants. Because of many

design parameters and high dependence on the available data, these methods are highly controversial and pose questions about robustness and reproducibility. Moreover, a common issue is that regulatory authorities are less likely to approve complex alternative designs. Consequently, either the clinical data should be analyzed in a different way or an intuitive extension (depending on the response models) to conventional designs should be proposed.

3.3.1 Response models

Starting from early 70s when D. Rubin proposed his causal model (RCM) (Rubin [1974]), in many research articles, e.g., Zhang et al. [2013], Jamshidian et al. [2014], the response $Y(t, s)$ is modeled as a function of treatment, $t \in \{0, 1\}$ (placebo or active treatment), and the participant's belief, $s \in \{0, 1\}$, - strong or weak belief in the efficacy of the treatment. In Zhang et al. [2013], for example, the average treatment effect is defined as

$$\delta_{.s} = \mathbb{E}[Y(1, s)] - \mathbb{E}[Y(0, s)],$$

and the average placebo effect is

$$\delta_{t.} = \mathbb{E}[Y(t, 1)] - \mathbb{E}[Y(t, 0)].$$

The interaction between t and s is

$$\tau = \delta_{.1} - \delta_{.0} = \delta_{1.} - \delta_{0.}.$$

With independence conditions of the type $S \perp Y(t, s) | \mathbf{X} = t$; $t, s = 0, 1$, where \mathbf{X} is a vector of explanatory variables, one can estimate the effects $\delta_{.s}$, $\delta_{t.}$. The causal models raise many questions, both about the dichotomized beliefs and the treatment effect estimates. For the latter, the model estimates an 'intention-to-treat' effect, which is a causal effect of the treatment assignment, rather than a causal effect of the tested compound.

More realistic assumptions are based on the heterogeneity of the patients' responses to the treatment. The most widely used probabilistic models accounting for different subpopulations in the observed data are finite mixture models. Muthén et al. [2002] were the first to use the mixture approach for a longitudinal trial with two arms: a placebo and active treatment. The patients' outcomes were modeled using a random-effects model (REM), whose coefficients depend on the latent subgroup membership and on the treatment arm. Consider k subpopulations, in each of which the response to the drug and the response to a placebo might be different. Given T observed time points, the model for the individual outcome at time t is given by

$$y_{it} = \eta_{0i} + \eta_{1i}a_t + \eta_{2i}a_t^2 + \varepsilon_{it},$$

where $a_1 = 0$, $a_t > 0$, $t = 1, \dots, T$ are the time variables, $\varepsilon_i \sim \mathcal{N}(\mathbf{0}, \Theta)$, $\Theta \in \mathbb{R}^{T \times T}$. The random

effects are

$$\begin{aligned}\eta_{0i} &= \alpha_{0k} + \xi_{0i}, \\ \eta_{1i} &= \alpha_{1k} + \gamma_{1k}I_i + \xi_{1i}, \\ \eta_{2i} &= \alpha_{2k} + \gamma_{2k}I_i + \xi_{2i},\end{aligned}$$

where $\xi_i \sim \mathcal{N}(\mathbf{0}, \Psi)$, $\Psi \in \mathbb{R}^{3 \times 3}$ and $I = \{0, 1\}$ for the control and the treatment group, respectively. Thus, the mean trajectory for the k -th subpopulation in the control group is fully determined by $\alpha_{0k}, \alpha_{1k}, \alpha_{2k}$. Assuming independence of ε_i and ξ_i , the conditional density of \mathbf{y}_i , given the class membership c_i and the treatment \mathbf{x}_i (the vector consisting of either all 0's or all 1's), is

$$f(\mathbf{y}_i | c_i = k, \mathbf{x}_i) \sim \mathcal{N}(\boldsymbol{\mu}_i, \Sigma_i) \quad \text{with} \quad \boldsymbol{\mu}_i = \Lambda_k(\boldsymbol{\alpha}_k + \boldsymbol{\gamma}_k \mathbf{x}_i), \quad \Sigma_i = \Lambda_k \Psi \Lambda_k^T + \Theta.$$

$$\boldsymbol{\alpha}_k = \begin{pmatrix} \alpha_{0k} \\ \alpha_{1k} \\ \alpha_{2k} \end{pmatrix}, \quad \boldsymbol{\gamma}_k = \begin{pmatrix} 0 \\ \gamma_{1k} \\ \gamma_{2k} \end{pmatrix}, \quad \Lambda_k = \begin{pmatrix} 1 & 0 & 0 \\ 1 & a_2 & a_2^2 \\ \vdots & \vdots & \vdots \\ 1 & a_T & a_T^2 \end{pmatrix}.$$

Specifying the number of subgroups k , the problem of parameter estimation becomes a problem with incomplete data which can be solved by the EM-algorithm. The conditional likelihood is

$$l(\mathbf{y}_i | c_i, \mathbf{x}_i, \boldsymbol{\eta}_i) = \sum_{i=1}^n \log(\mathbf{y}_i | c_i, \mathbf{x}_i, \boldsymbol{\eta}_i) + \log(\boldsymbol{\eta}_i | c_i, \mathbf{x}_i).$$

To select k , the authors compare the BIC scores. The proposed growth mixture model with random effects appeared to support an idea of the existence of the subclasses and improved accuracy for treatment effect estimation, see Wong et al. [2017], Petkova et al. [2009], Stull et al. [2011], He and Entsuaah [2014].

Muthén and Brown [2009] studied a specified and simplified model for the subpopulations. Following the Angrist, Imbens and Rubin (AIR) model (Angrist et al. [1996]), Muthén and Brown investigated the mean differences between active treatment and placebo groups across the *principal strata*, the homogeneous subgroup of patients, of four latent classes *Never Responders*, *Drug Only Responders*, *Placebo Only Responders*, and *Always Responders*, see Table 3.2. The class membership is assumed to be independent of the treatment arm and determined by some other covariates. The marginal placebo and drug treatment effects are then

$$\begin{aligned}\mu_0 &= \pi_n \mu_{n0} + \pi_d \mu_{d0} + \pi_p \mu_{p0} + \pi_a \mu_{a0}, \\ \mu_1 &= \pi_n \mu_{n1} + \pi_d \mu_{d1} + \pi_p \mu_{p1} + \pi_a \mu_{a1}.\end{aligned}$$

This model can be reduced by putting $\pi_p = 0$ (the existence of the *Placebo Only Responders* stratum seems doubtful), $\mu_{n0} = \mu_{n1}$ and $\mu_{a0} = \mu_{a1}$ (the *Always* and *Never Responders* should

		Treatment group	
		Non-responder	Responder
Placebo group	Non-responder	Never responder: $\pi_n, \mu_{n0}, \mu_{n1}$	Drug only responder: $\pi_d, \mu_{d0}, \mu_{d1}$
	Responder	Placebo only responder: $\pi_p, \mu_{p0}, \mu_{p1}$	Always responder: $\pi_a, \mu_{a0}, \mu_{a1}$

Table 3.2 – The principal strata parameters in the placebo (0) and treatment arms (1).

have no difference between the arms). The difference between the experimental arms is then $\mu_1 - \mu_0 = \pi_d(\mu_{d1} - \mu_{d0})$. On the other hand, $\hat{\pi}_d = (1 - \hat{\pi}_a)$, since the proportion of Always Responders is directly estimated as the proportion of the responders in the placebo group.

One could argue that the assumption of equality of the strata proportions between the arms does not hold in real clinical trials and that the model with so many parameters to estimate is still redundant. Though estimation is possible, for the design one should employ parsimonious and easily tractable models.

3.4 Conventional testing for the mixture response

Let the treatment group be a mixture of two subpopulations, placebo only responders $Z \sim \mathcal{N}(\mu^C, 1)$ and drug responders $Z \sim \mathcal{N}(\mu^T, 1)$ (we assume that these people may also experience an improvement due to the placebo). Under these assumptions, responses in the treatment group follow $X \sim (1 - p)\mathcal{N}(\mu^C, 1) + p\mathcal{N}(\mu^T, 1)$, where p is the prevalence of drug responders in the population. What would be the power of the conventional mean difference testing with the statistic $\bar{X} = \bar{X}_{\text{Treatment arm}} - \bar{X}_{\text{Control arm}}$? The test is

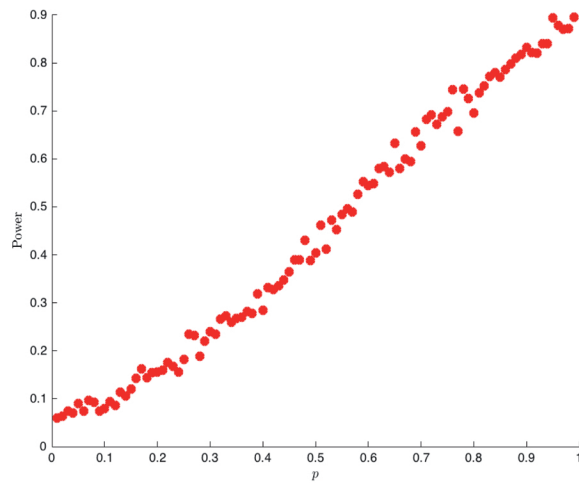
$$\begin{aligned} \mathcal{H}_0 : \gamma &= 0 \quad \text{versus} \\ \mathcal{H}_1 : \gamma &> 0, \end{aligned} \tag{3.1}$$

where $\gamma = p(\mu^T - \mu^C)$. For Gaussian responses the distribution of the groups' mean difference under the alternative is $\bar{X} \sim \mathcal{N}(\gamma, \frac{2}{n})$. For the minimum effect γ_{\min} and for the the maximum type II error rate β_{\max} , the level α test rejects the null if $\bar{X} > z_{1-\alpha}\sqrt{\frac{2}{n}}$, and the number of participants is

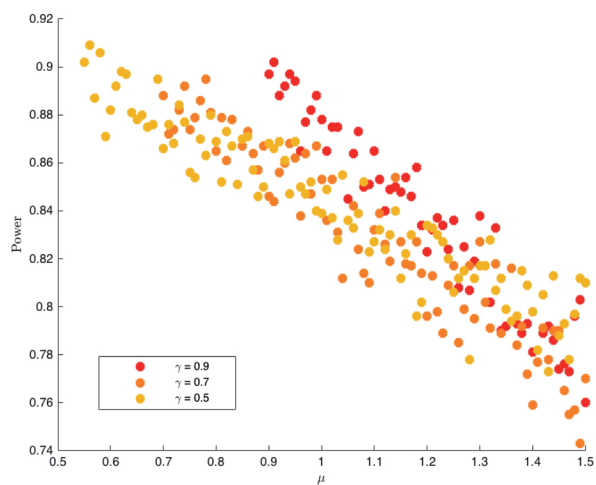
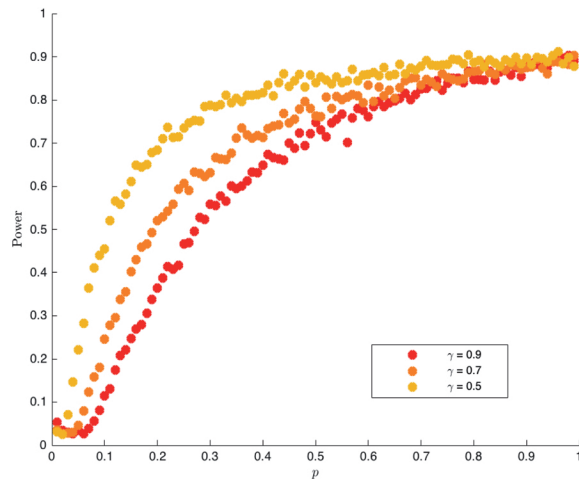
$$n = \left(\sqrt{2} \frac{z_{1-\alpha} + z_{1-\beta_{\max}}}{\gamma_{\min}} \right)^2. \tag{3.2}$$

- Assume an equal treatment-specific effect for all participants in the treatment group, $\gamma_{\min} = \mu_{\min}$. If the true alternative is a mixture, the power of the test substantially decreases when p decreases, see Fig. 3.5(a).
- One could also consider the effect as the set $\{(\mu^T - \mu^C, p) : (\mu^T - \mu^C)p = \gamma_{\min}\}$. In Fig. 3.5(b,c), we show simulation results for the mean-based test (3.1) when the data

3.4. Conventional testing for the mixture response



(a) The sample size from (3.2) is computed assuming $\mu = 0.7$, $p = 1$. The data are generated from $(1 - p)\mathcal{N}(0, 1) + p\mathcal{N}(0.7, 1)$.



The sample size from (3.2) is computed assuming $\gamma = \gamma_{\min}$. The data are generated from the mixture with $\{(\mu, p) : \mu p = \gamma\}$.

Figure 3.5 – The power of the test (3.1) when the data follow $(1 - p)\mathcal{N}(0, 1) + p\mathcal{N}(\mu p, 1)$. The design is built to guarantee $\alpha = 0.05$, $\beta \leq 0.1$.

is generated from the mixture $(1 - p)\mathcal{N}(0, 1) + p\mathcal{N}(\mu, p)$ with $\mu p = \gamma$ and the sample size is from (3.2). One can see that p influences the power more than μ . Effects with the bigger magnitude but rare frequency will be detected less often than those with smaller μ and larger prevalence.

3.5 Conclusion

The examples above illustrate that misspecifying an alternative when building a trial design can lead to a lack of power. Moreover, if the test is explicitly built assuming the mixture alternatives, the region of (μ, p) corresponding to the potentially interesting effect sizes should be accurately defined to ensure the power in this region is above a certain level.

Recall that in Chapter 2 we saw that if an alternative belongs to the region of dense signals (small magnitude / high frequency), statistics based on the first moment of the observed data are optimal in the sense that the detection boundary for these tests lies close to the detection boundary for the most powerful likelihood ratio test. The dense region there corresponds to alternatives with $\delta < 0.5$. As far as we know, except for the cases with rare diseases which are characterized by the known single genetic mutations, the percentage of population that is assumed to react to the tested drug is rarely less than 10%. If the disease is not rare and there are no known biomarkers for identifying the subgroup, clinical trial size of around few hundred patients allows us to consider the dense regime and use the sample mean for the inference.

In the next chapter we consider the designs where the response in the treatment group Z^T is modeled as a two-component mixture density $(1 - p)\mathcal{N}(\mu^C, \sigma^2) + p\mathcal{N}(\mu^T, \sigma^2)$ representing the treatment responses of *placebo responders* and *drug responders*. The treatment-specific effect is $\mu = \frac{\mu^T - \mu^C}{\sigma}$ and p is the prevalence of the drug responders in the population. Other patients in the treatment group react as if they had received a placebo. We develop a specific framework for one- and two-stage RCT designs that are able to detect a sensitive subgroup based solely on the responses. We also extend designs to multicenter RCTs using multiple testing procedures.

4 Randomized controlled trial (RCT) designs

The idea that there is a proportion of non-responders in the treatment group is often highlighted in clinical papers, e.g., Rubin [1974], Frangakis and Rubin [2002], Muthén and Brown [2009], Leiby et al. [2009], He and Entsuah [2014]. If one has specific biomarkers which point to the subjects who are likely to respond to an intervention, then a subgroup analysis is usually conducted in addition to the standard procedure. Most of the time, however, there is no such information about the sensitive subgroup, and standard designs may fail to prove the drug's efficacy.

Moreover, even if the proportion of the sensitive patients is big, there is another hurdle which can mask a drug-specific effect, namely the placebo. When its effect is comparable with a drug, the detection of a subgroup is complicated substantially. This problem may occur in treatments for neurological disorders, where many trials fail to show efficacy, and the reason is supposed to be a large placebo effect, see Agid et al. [2013], Tuttle et al. [2015], Holmes et al. [2016]. For our purpose, the placebo effect is a combination of natural course of the disease, the Hawthorne effect and other non-specific effects influencing the treatment outcome.

Given the assumptions above, we elaborate a framework in which we are able to detect the sensitive subgroup based on the following model.

4.1 The Model

Let the response to the drug, Z , be a normally distributed random variable. We assume that there are two groups of patients: placebo-only responders with $Z \sim \mathcal{N}(\mu^C, \sigma^2)$ and placebo-and-drug responders with $Z \sim \mathcal{N}(\mu^T, \sigma^2)$, where $\mu^T > \mu^C$. The drug-specific effect, $\mu^T - \mu^C$, and the placebo effect are additive. The response in the treatment group is then modelled as a mixture (see Fig. 4.1)

$$Z^T \sim (1 - p)\mathcal{N}(\mu^C, \sigma^2) + p\mathcal{N}(\mu^T, \sigma^2),$$

where $p \geq 0$ is the proportion of placebo-and-drug responders (the sensitive subpopulation). The design is constructed to determine whether $p > 0$. Denote the standardized drug-specific

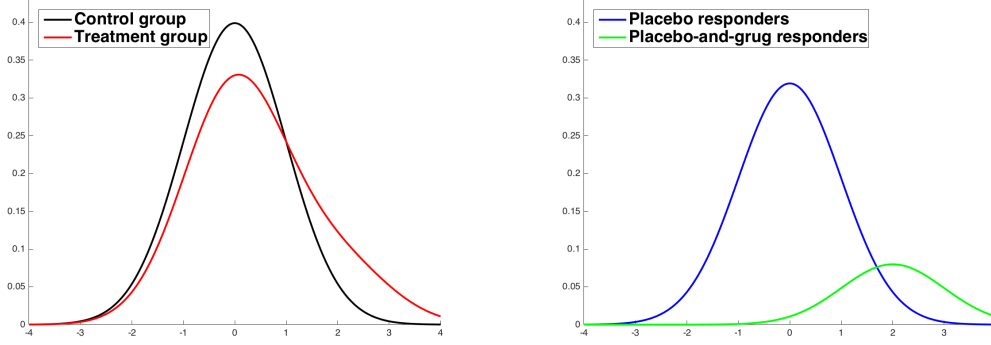


Figure 4.1 – Modelling drug response as a mixture of two continuous distributions has been proposed to be more reasonable in RCT for some types of indications. On the left, the control group is centered at zero and the mixture created in the treatment group is asymmetric with an enlarged upper tail. If we could separate the responders and non-responders in the treatment group, we would observe the graphs as in the right hand side.

effect as $\mu = \frac{\mu^T - \mu^C}{\sigma}$. Given the standardized estimates of responses in the treatment group,

$$X_i = \frac{Z_i^T - \overline{Z^C}}{\sigma}, \quad i = 1, \dots, n,$$

the average value \overline{X} has expectation μp and is distributed as $\sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} \mathcal{N}\left(\frac{k}{n}\mu, \frac{2}{n}\right)$. The test based on the average of observations can be used in deciding between $\mathcal{H}_0 : p = 0$ against $\mathcal{H}_1 : p > 0$. As saw in Chapter 2, the performance of the mean-based test is close to that of the LRT test for $p > 0.1$. Along with its simplicity the mean value statistic can also be used with non-normal responses that may be encountered in clinical trials.

To control the power of detection, we consider the minimal power over a region of strong effect, the set of (μ, p) that are deemed to be of interest. We propose one- and two-stage designs for the single and multicenter trials. This framework is mainly aimed at the early phase II before the identification of the drug-responders subgroup.

4.2 The designs

In all RCTs discussed in this chapter we consider any treatment group at any stage to have a corresponding control group of the same size. Let Z_1^T, \dots, Z_n^T and Z_1^C, \dots, Z_n^C be the responses from the treatment and the control groups, respectively. For convenience we define normalized responses as

$$Y^C = \frac{Z^C - \mu^C}{\sigma} \sim \mathcal{N}(0, 1), \quad Y^T = \frac{Z^T - \mu^C}{\sigma} \sim (1-p)\mathcal{N}(0, 1) + p\mathcal{N}(\mu, 1),$$

where $\mu = \frac{\mu^T - \mu^C}{\sigma}$. Since we do not know μ^C and σ , it is not possible to observe Y^T . Instead, we take its estimate, $X_i = \frac{Z_i^T - \hat{\mu}^C}{\sigma}$, where $\hat{\mu}^C = \frac{\sum_{i=1}^n Z_i^C}{n}$. For the design development we assume σ to be known, but in practice we may use the approximation $\hat{\sigma}^C = \sqrt{\frac{\sum_{i=1}^n (Z_i^C - \hat{\mu}^C)^2}{n-1}}$. In simulations shown later, we show that the approximation error has a negligible effect on the results.

In a simple RCT with a single center, the test is based on the mean value statistic,

$$\bar{X} = \frac{X_1 + \dots + X_n}{n} = \overline{Y^T} - \frac{\hat{\mu}^C - \mu^C}{\sigma}.$$

Note that

$$\overline{Y^T} \sim \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} \mathcal{N}\left(\frac{k}{n}\mu, \frac{1}{n}\right),$$

$\frac{\hat{\mu}^C - \mu^C}{\sigma} \sim \mathcal{N}\left(0, \frac{1}{n}\right)$, and the distribution of \bar{X} is

$$\bar{X} \sim \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} \mathcal{N}\left(\frac{k}{n}\mu, \frac{2}{n}\right).$$

4.2.1 Hypothesis testing

The goal of a trial is to determine if there is a sensitive subpopulation. The decision is based on a statistical test of

$$\begin{aligned} \mathcal{H}_0 : Y^T \sim \mathcal{N}(0, 1) & \quad \text{versus} \\ \mathcal{H}_1 : Y^T \sim (1-p)\mathcal{N}(0, 1) + p\mathcal{N}(\mu, 1), & \quad p \in (0, 1], \mu > 0. \end{aligned} \tag{4.1}$$

Of course, not all pairs (μ, p) are of equal interest for the drug developer. It is of small interest if \mathcal{H}_0 is rejected when \mathcal{H}_1 is true with some small p and/or μ . Therefore, it is important to define *interesting* values of (μ, p) , for which one wants to reject the null.

Definition 4.1. *The region \mathcal{E} of strong effect (se) is the set of interesting pairs (μ, p) such that $\mu \geq \mu_i$ if $p \in [p_i, p_{i+1}]$, where $\mu_s < \mu_{s-1} < \dots < \mu_1$ and $0 < p_1 < \dots < p_{s+1} = 1$ (see Fig. 4.2).*

For large values of p , average or even small effect sizes, μ , can be of interest. Conversely, for smaller values of p , μ would need to be larger. For each magnitude of the treatment effect we set the smallest desirable fraction p of the sensitive subpopulation. In Fig. 4.2 we show the example of the set of possible pairs (μ, p) . In the simplest case the staircase consists of one pair (μ_1, p_1) . We suppose the region of this type to be provided by the user.

In the next two subsections, we introduce the single and the multicenter RCT designs, for which type I and II errors are defined as follows.

Definition 4.2. *For a simple study, a type I error occurs when a true null is rejected. A type II error occurs if a false null is not rejected in the presence of a strong effect. The corresponding*

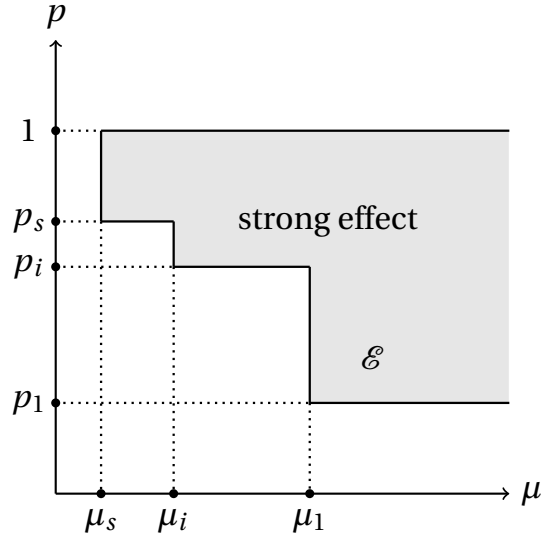


Figure 4.2 – The region of strong effect, \mathcal{E} , is a subset of the (μ, p) plane satisfying $\mu \geq \mu_i$ if $p \in [p_i, p_{i+1}]$ where $\mu_s < \mu_{s-1} < \dots < \mu_1$ and $0 < p_1 < p_2 < \dots < p_{s+1} = 1$.

type I error rate will be denoted as α and the maximum of the type II error rate over \mathcal{E} is denoted as β^{se} . In the case of a multicenter study with M centers a type I error occurs if at least one of the true nulls is rejected, and a type II error occurs if at least m false nulls among the M_1 centers with strong effects are not rejected. Note that $m \leq M_1 \leq M$. The type I error is a family-wise error, while the type II error depends on M_1 and m . The type II error rate maximized over \mathcal{E} will be denoted by β^{se} for the single stage study and $\beta_{fw}^{se}(M_1, m)$ of the multicenter study.

4.2.2 Simple RCT designs - single center RCT

We start with the simplest design: single center, one stage. For the mean value statistic the hypothesis testing problem is

$$\begin{aligned} \mathcal{H}_0 : \bar{X} &\sim \mathcal{N}\left(0, \frac{2}{n}\right); \\ \mathcal{H}_1 : \bar{X} &\sim \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} \mathcal{N}\left(\frac{k}{n}\mu, \frac{2}{n}\right). \end{aligned} \tag{4.2}$$

The design of a trial is defined by the sample size n and some positive threshold η , upon which the decision about a subgroup's existence is made,

$$\text{Reject } \mathcal{H}_0 \text{ if } \bar{X} > \eta. \tag{D1}$$

The chosen test is justified by the following lemma:

Lemma 4.1. *The tests that reject \mathcal{H}_0 for $\bar{X} > \eta$ are uniformly most powerful.*

Proof. We denote the pdf of \bar{X} under \mathcal{H}_0 as $f(x)$ and the pdf of \bar{X} under an alternative as $g(x)$. The function $\frac{g(x)}{f(x)}$ is non-decreasing in x if $\mu > 0$, since

$$\frac{g(x)}{f(x)} = \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} \frac{f\left(x - \frac{k}{n}\mu\right)}{f(x)} = \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} e^{\frac{n}{2}\left(x - \frac{k}{n}\mu - \frac{1}{2}\left(\frac{k}{n}\mu\right)^2\right)}.$$

From the Karlin–Rubin theorem it follows that the test $\bar{X} > \eta$ is UMP. \square

The equivalent formulation of the design is

$$\text{Reject } \mathcal{H}_0 \text{ if the p-value } \left(\bar{X}\right) < \alpha, \text{ where p-value } \left(\bar{X}\right) = 1 - \Phi\left(\bar{X}\sqrt{\frac{n}{2}}\right) \text{ and } \alpha < 0.5. \quad (\text{D1})$$

Here, α is a type I error rate, and $\eta = z_{1-\alpha}\sqrt{\frac{2}{n}}$, where $z_{1-\alpha} = \Phi^{-1}(1-\alpha)$ stands for the standard Gaussian quantile. The probability of a false negative for a fixed alternative is

$$\beta(n, \eta, \mu, p) = \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} \Phi\left(\left(\eta - \frac{k}{n}\mu\right)\sqrt{\frac{n}{2}}\right). \quad (4.3)$$

By Definition 4.2,

$$\beta^{se}(n, \alpha) = \max_{(\mu, p) \in \mathcal{E}} \beta\left(n, z_{1-\alpha}\sqrt{\frac{2}{n}}, \mu, p\right). \quad (4.4)$$

As the following lemma shows, one can easily compute $\beta^{se}(n, \alpha)$ for the regions \mathcal{E} of the form we defined.

Lemma 4.2. *For the region of strong effect defined as in Definition 4.1, the maximum type II error rate for the one-stage RCT is*

$$\beta^{se}(n, \alpha) = \max_{i=1, \dots, s} \beta\left(n, z_{1-\alpha}\sqrt{\frac{2}{n}}, \mu_i, p_i\right).$$

Proof. The probability of a false negative is

$$\beta(n, \eta, \mu, p) = \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} \Phi\left(\sqrt{\frac{n}{2}}\left(\eta - \frac{k}{n}\mu\right)\right).$$

The maximum value is attained in one of the corners of \mathcal{E} . Indeed, this follows from

$$\frac{\partial \beta}{\partial \mu}(n, \eta, \mu, p) < 0 \quad \text{and} \quad \frac{\partial \beta}{\partial p}(n, \eta, \mu, p) < 0.$$

$$\frac{\partial \beta}{\partial \mu}(n, \eta, \mu, p) = - \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} \frac{k}{\sqrt{2n}} \varphi\left(\sqrt{\frac{n}{2}}\left(\eta - \frac{k}{n}\mu\right)\right) < 0; \quad (4.5)$$

$$\begin{aligned}
 \frac{\partial \beta}{\partial p}(n, \eta, \mu, p) &= n \sum_{k=1}^n \binom{n-1}{k-1} p^{k-1} (1-p)^{n-k} \Phi\left(\sqrt{\frac{n}{2}}\left(\eta - \frac{k}{n}\mu\right)\right) \\
 &\quad - n \sum_{k=0}^{n-1} \binom{n-1}{k} p^k (1-p)^{n-k-1} \Phi\left(\sqrt{\frac{n}{2}}\left(\eta - \frac{k}{n}\mu\right)\right) \\
 &= n \sum_{k=0}^{n-1} \binom{n-1}{k} p^k (1-p)^{n-k-1} \left(\Phi\left(\sqrt{\frac{n}{2}}\left(\eta - \frac{k+1}{n}\mu\right)\right) - \Phi\left(\sqrt{\frac{n}{2}}\left(\eta - \frac{k}{n}\mu\right)\right)\right) < 0.
 \end{aligned} \tag{4.6}$$

□

Two stage designs

By Wald's theory of sequential analysis, there are modifications of the RCT design enabling significant reduction in the required minimum number of the participants. The idea of sequential analysis is to split the testing process into several steps and update the inference based on the information accumulated from the previous steps. In the context of RCT, the steps are the stages of the trial.

The first sequential design was introduced by Gehan [1961] for anti-cancer drugs where toxicity is the major concern and the goal of the trial is to make conclusions about the efficiency with the minimum number of patients involved. The end point in these designs is the tumor response, which is the proportion π of patients that have tumors' shrink by at least 50% during some time period. In Gehan's design, 14 patients are enrolled at the first stage. With the response rate of $\pi \geq 0.2$, which is considered as promising for further investigations, 14 patients is chosen to ensure the type I error rate of $(1 - 0.2)^{14} = 0.044$, close to 0.05. A second stage is conducted to guarantee the standard error rate of the response with certain precision. The main drawback of this design is the fixed value of n_1 that might allow drugs with low anti-tumor activity to pass the first stage.

In Simon [1989] two other designs were proposed. One minimizes the maximum sample size $n_1 + n_2$ under \mathcal{H}_0 (minimax criterion) and the other minimizes the average sample size under \mathcal{H}_0 . Simon's designs were adapted to early stopping for efficacy in Jones and Holmgren [2007]. However, for many indications an early acceptance even of a highly effective drug is rarely possible, and most of the time a confirmatory trial is needed. On the contrary, in our case of sensitive subgroup identification, the design admits an early acceptance (or an early rejection of \mathcal{H}_0). The sooner one finds evidence in favour of the existence of sensitive population, the more gain in terms of the drug development time she acquires.

Two-stage designs for continuous responses are discussed in Jennison and Turnbull [1999]. Optimality criteria along with corresponding designs were proposed in Whitehead et al. [2009], Wason and Mander [2012] with individual responses distributed as $\mathcal{N}(\delta_C, \sigma_C)$ and $\mathcal{N}(\delta_T, \sigma_T)$ in the control and treatment groups, respectively.

Below we will show by how much the introduction of a second stage decreases the total ex-

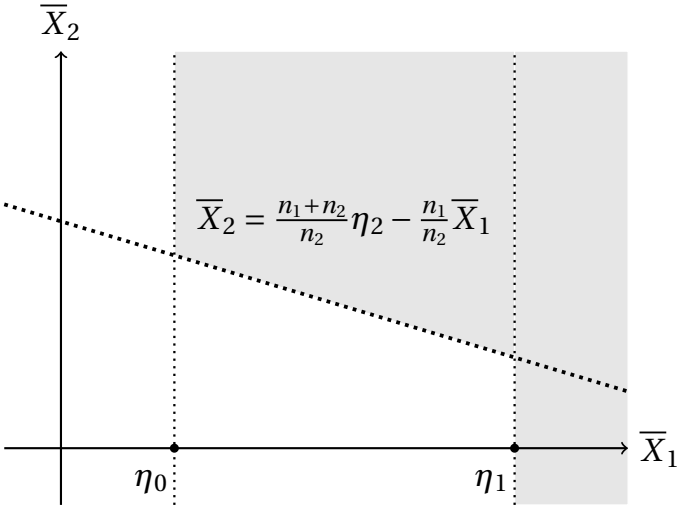


Figure 4.3 – Scheme of a trial performed in two stages. The rejection region is depicted in grey.

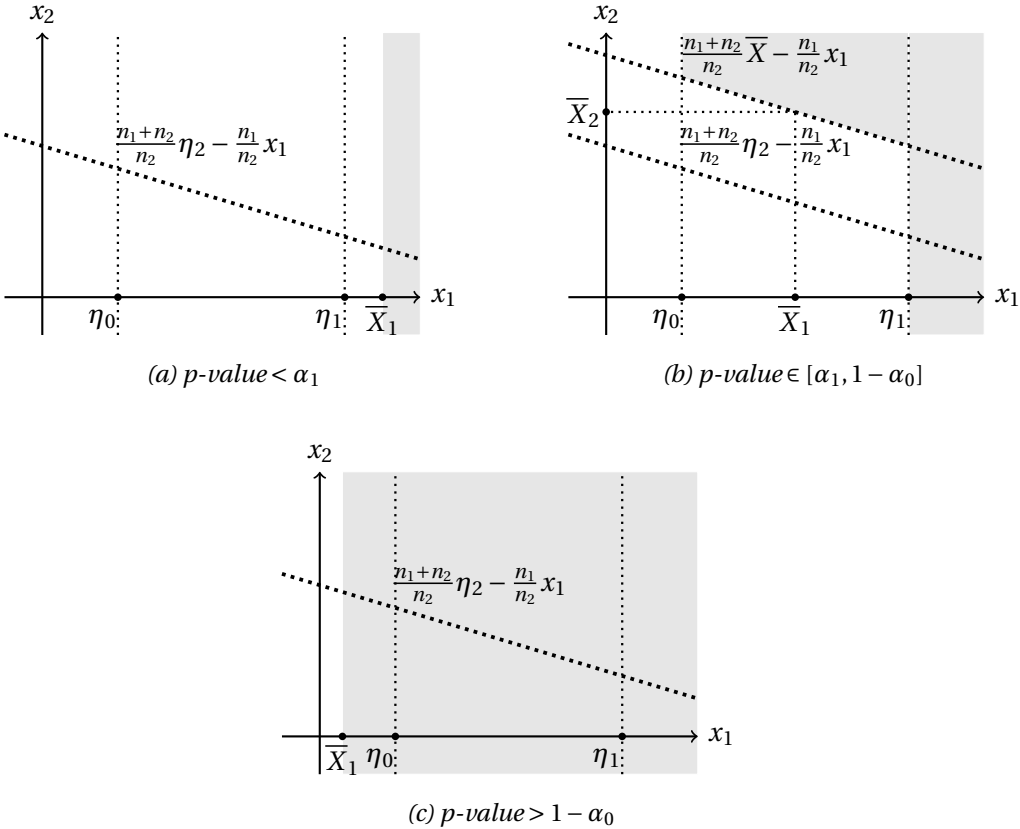


Figure 4.4 – Computation of the p-value for a trial performed in two stages.

pected number of enrolled patients in comparison with a one-stage design, while maintaining the same error rates. For the first stage we propose two stopping rules. First, if the evidence in favour of the alternative is very strong after the first stage, there is no need to conduct the second stage. Second, if the evidence in favour of the null is strong enough at the first stage, one can stop the trial early, claiming the absence of an effect.

To start with, consider some fixed sample sizes with n_1 and n_2 participants at the first and second stages, respectively. We also assume that there are two different control groups at each stage with the same number of participants as in the treatment groups. Denote the mean observed values after the first and the second stages by \bar{X}_1 and \bar{X}_2 , and the mean of all observations as \bar{X} . The design is described as follows:

- If $\bar{X}_1 < \eta_0$: \mathcal{H}_0 is not rejected and the trial is stopped
(strong evidence of an absence of a positive treatment effect);
- If $\bar{X}_1 > \eta_1$: reject \mathcal{H}_0 and stop the trial (strong evidence in favour of \mathcal{H}_1); (D2)
- If $\eta_0 \leq \bar{X}_1 \leq \eta_1$: conduct the second stage and reject \mathcal{H}_0 if $\bar{X} > \eta_2$
(strong evidence in favour of \mathcal{H}_1 after two steps).

The rejection region is depicted in Fig. 4.3 (grey area). For the thresholds, η_0, η_1 , denote the corresponding probabilities $\mathbb{P}(\bar{X}_1 < \eta_0)$ and $\mathbb{P}(\bar{X}_1 > \eta_1)$ under the null as $\alpha_0 = \Phi\left(\eta_0\sqrt{\frac{n_1}{2}}\right)$ and $\alpha_1 = 1 - \Phi\left(\eta_1\sqrt{\frac{n_1}{2}}\right)$. In terms of p-values, this design becomes:

Reject \mathcal{H}_0 if p-value $(\bar{X}_1, \bar{X}_2) < \alpha$, where

$$\text{p-value}(\bar{X}_1, \bar{X}_2) = \begin{cases} 1 - \Phi\left(\bar{X}_1\sqrt{\frac{n_1}{2}}\right), & \bar{X}_1 \notin [\eta_0, \eta_1], \\ \int_{\eta_0}^{\eta_1} \sqrt{\frac{n_1}{2}} \varphi\left(x_1\sqrt{\frac{n_1}{2}}\right) \left(1 - \Phi\left(\frac{\sqrt{n_2}\bar{X}_2 + \frac{n_1}{\sqrt{n_2}}(\bar{X}_1 - x_1)}{\sqrt{2}}\right)\right) dx_1 + \alpha_1, & \bar{X}_1 \in [\eta_0, \eta_1]. \end{cases} \quad (\text{D2})$$

The computation of the p-value is complicated by the fact that the grey region in Fig. 4.4(b) is not simple. The p-value is an integral of the joint density function $f_{(\bar{X}_1, \bar{X}_2)}(x_1, x_2)$ under \mathcal{H}_0 over the grey area in Fig. 4.4. If $\bar{X}_1 \in [\eta_0, \eta_1]$, we integrate above the line

$$\frac{n_1 + n_2}{n_2} \bar{X} - \frac{n_1}{n_2} x_1 = \bar{X}_2 + \frac{n_1}{n_2} (\bar{X}_1 - x_1)$$

and for the fixed x_1 the probability that x_2 is above this line equals

$$1 - \Phi\left(\frac{\sqrt{n_2}\bar{X}_2 + \frac{n_1}{\sqrt{n_2}}(\bar{X}_1 - x_1)}{\sqrt{2}}\right).$$

Under \mathcal{H}_0 this p-value is uniformly distributed on the interval $(0, 1)$, because for any two different realizations of (\bar{X}_1, \bar{X}_2) , one of the grey areas (see Fig. 4.4) contains the other one.

The design D is fully determined by the sample sizes in the first and second stages, n_1 and n_2 , and the probabilities α_0 and α_1 ; the overall level of the procedure is α . These values often satisfy $1 - \alpha_0 > \alpha > \alpha_1$. If this does not hold, the second stage is not needed. The thresholds η_0, η_1, η_2 can be computed as

$$\begin{aligned}\eta_0 &= z_{\alpha_0} \sqrt{\frac{2}{n_1}}, \\ \eta_1 &= z_{1-\alpha_1} \sqrt{\frac{2}{n_1}}, \\ \eta_2 &\text{ such that } \int_{\eta_0}^{\eta_1} \sqrt{\frac{n_1}{2}} \varphi\left(x_1 \sqrt{\frac{n_1}{2}}\right) \left(1 - \Phi\left(\left(\frac{n_1 + n_2}{n_2} \eta_2 - \frac{n_1}{n_2} x_1\right) \sqrt{\frac{n_2}{2}}\right)\right) dx_1 + \alpha_1 = \alpha.\end{aligned}\tag{4.7}$$

Denote the maximum type II error rate for the two-stage design as $\beta_2^{se}(D, \alpha)$. To compute it, we use an expression for the probability of the false negative:

$$\begin{aligned}\beta_2(D, \alpha, \mu, p) &= \mathbb{P}\left(\text{p-value}(\bar{X}_1, \bar{X}_2) \geq \alpha \mid \mu, p\right) = \\ &= \begin{cases} \beta\left(n_1, z_{1-\alpha_1} \sqrt{\frac{2}{n_1}}, \mu, p\right), & \alpha \notin (\alpha_1, 1 - \alpha_0), \\ \beta(n_1, \eta_0, \mu, p) + \int_{\eta_0}^{\eta_1} \beta_\eta(n_1, x_1, \mu, p) \beta\left(n_2, \frac{n_1 + n_2}{n_2} \eta_2(D, \alpha) - \frac{n_1}{n_2} x_1, \mu, p\right) dx_1, & \alpha \in (\alpha_1, 1 - \alpha_0). \end{cases}\end{aligned}\tag{4.8}$$

We introduce the case $\alpha \notin (\alpha_1, 1 - \alpha_0)$ in (4.8), because we will need it later for the multicenter two-stage design. Similarly to the one-stage trial, the following holds.

Lemma 4.3. *For the region of strong effect defined as in Definition 4.1, the maximum type II error rate for the two-stage trial is*

$$\beta_2^{se}(D, \alpha) = \max_{i=1, \dots, s} \beta_2(D, \alpha, \mu_i, p_i).$$

Proof. This can be shown by proving that

$$\frac{\partial \beta_2}{\partial \mu}(n_1, \eta_0, \eta_1, n_2, \alpha, \mu, p) < 0 \text{ and } \frac{\partial \beta_2}{\partial p}(n_1, \eta_0, \eta_1, n_2, \alpha, \mu, p) < 0.$$

We will use inequalities $\frac{\partial \beta}{\partial \mu}(n, \eta, \mu, p) < 0$ and $\frac{\partial \beta}{\partial p}(n, \eta, \mu, p) < 0$ from Lemma 4.2 and inequality $\frac{\partial \beta}{\partial \eta}(n, \eta, \mu, p) > 0$, which is a trivial consequence from (4.3). There are two different cases:

1. $\alpha \notin (\alpha_1, 1 - \alpha_0)$:

$$\begin{aligned}\frac{\partial \beta_2}{\partial \mu}(n_1, \eta_0, \eta_1, n_2, \alpha, \mu, p) &= \frac{\partial \beta}{\partial \mu}(n_1, z_{1-\alpha} \sqrt{\frac{2}{n_1}}, \mu, p) < 0, \\ \frac{\partial \beta_2}{\partial p}(n_1, \eta_0, \eta_1, n_2, \alpha, \mu, p) &= \frac{\partial \beta}{\partial p}(n_1, z_{1-\alpha} \sqrt{\frac{2}{n_1}}, \mu, p) < 0.\end{aligned}$$

2. $\alpha \in (\alpha_1, 1 - \alpha_0)$:

$$\begin{aligned}\frac{\partial \beta_2}{\partial \mu}(n_1, \eta_0, \eta_1, n_2, \alpha, \mu, p) &= \frac{\partial \beta}{\partial \mu}(n_1, \eta_0, \mu, p) + \int_{\eta_0}^{\eta_1} \frac{\partial \beta}{\partial \eta \partial \mu}(n_1, x_1, \mu, p) \beta(n_2, x_2(x_1), \mu, p) dx_1 + \\ &+ \int_{\eta_0}^{\eta_1} \frac{\partial \beta}{\partial \eta}(n_1, x_1, \mu, p) \frac{\partial \beta}{\partial \mu}(n_2, x_2(x_1), \mu, p) dx_1 = \frac{\partial \beta}{\partial \mu}(n_1, \eta_0, \mu, p) + \frac{\partial \beta}{\partial \mu}(n_1, x_1, \mu, p) \beta(n_2, x_2(x_1), \mu, p) \Big|_{\eta_0}^{\eta_1} + \\ &+ \frac{n_1}{n_2} \int_{\eta_0}^{\eta_1} \frac{\partial \beta}{\partial \mu}(n_1, x_1, \mu, p) \frac{\partial \beta}{\partial \eta}(n_2, x_2(x_1), \mu, p) dx_1 + \int_{\eta_0}^{\eta_1} \frac{\partial \beta}{\partial \eta}(n_1, x_1, \mu, p) \frac{\partial \beta}{\partial \mu}(n_2, x_2(x_1), \mu, p) dx_1 = \\ &= \frac{\partial \beta}{\partial \mu}(n_1, \eta_0, \mu, p) (1 - \beta(n_2, x_2(\eta_0), \mu, p)) + \frac{\partial \beta}{\partial \mu}(n_1, \eta_1, \mu, p) \beta(n_2, x_2(\eta_1), \mu, p) + \\ &+ \frac{n_1}{n_2} \int_{\eta_0}^{\eta_1} \frac{\partial \beta}{\partial \mu}(n_1, x_1, \mu, p) \frac{\partial \beta}{\partial \eta}(n_2, x_2(x_1), \mu, p) dx_1 + \int_{\eta_0}^{\eta_1} \frac{\partial \beta}{\partial \eta}(n_1, x_1, \mu, p) \frac{\partial \beta}{\partial \mu}(n_2, x_2(x_1), \mu, p) dx_1 < 0; \\ \frac{\partial \beta_2}{\partial p}(n_1, \eta_0, \eta_1, n_2, \alpha, \mu, p) &= \frac{\partial \beta}{\partial p}(n_1, \eta_0, \mu, p) + \int_{\eta_0}^{\eta_1} \frac{\partial \beta}{\partial p \partial \eta}(n_1, x_1, \mu, p) \beta(n_2, x_2(x_1), \mu, p) dx_1 + \\ &+ \int_{\eta_0}^{\eta_1} \frac{\partial \beta}{\partial \eta}(n_1, x_1, \mu, p) \frac{\partial \beta}{\partial p}(n_2, x_2(x_1), \mu, p) dx_1 = \frac{\partial \beta}{\partial p}(n_1, \eta_0, \mu, p) + \frac{\partial \beta}{\partial p}(n_1, x_1, \mu, p) \beta(n_2, x_2(x_1), \mu, p) \Big|_{\eta_0}^{\eta_1} + \\ &+ \frac{n_1}{n_2} \int_{\eta_0}^{\eta_1} \frac{\partial \beta}{\partial p}(n_1, x_1, \mu, p) \frac{\partial \beta}{\partial \eta}(n_2, x_2(x_1), \mu, p) dx_1 + \int_{\eta_0}^{\eta_1} \frac{\partial \beta}{\partial \eta}(n_1, x_1, \mu, p) \frac{\partial \beta}{\partial p}(n_2, x_2(x_1), \mu, p) dx_1 = \\ &= \frac{\partial \beta}{\partial p}(n_1, \eta_0, \mu, p) (1 - \beta(n_2, x_2(\eta_0), \mu, p)) + \frac{\partial \beta}{\partial p}(n_1, \eta_1, \mu, p) \beta(n_2, x_2(\eta_1), \mu, p) + \\ &+ \frac{n_1}{n_2} \int_{\eta_0}^{\eta_1} \frac{\partial \beta}{\partial p}(n_1, x_1, \mu, p) \frac{\partial \beta}{\partial \eta}(n_2, x_2(x_1), \mu, p) dx_1 + \int_{\eta_0}^{\eta_1} \frac{\partial \beta}{\partial \eta}(n_1, x_1, \mu, p) \frac{\partial \beta}{\partial p}(n_2, x_2(x_1), \mu, p) dx_1 < 0,\end{aligned}$$

where $x_2(x_1) = \frac{n_1+n_2}{n_2} \eta_2(n_1, \eta_0, \eta_1, n_2, \alpha) - \frac{n_1}{n_2} x_1$. □

4.2.3 Multicenter RCT designs

In this section, we generalize the designs proposed above to multicenter RCTs. We suggest first to conduct single center RCTs with identical designs in all centers in parallel, with each center having its own control group. Following this initial stage, the results are analyzed together in order to make conclusions about the existence of the subpopulation in each center.

Combining the evidence across multiple centers requires p-values. For this reason, in the previous subsection, we introduced an equivalent designs formulation using p-values. We

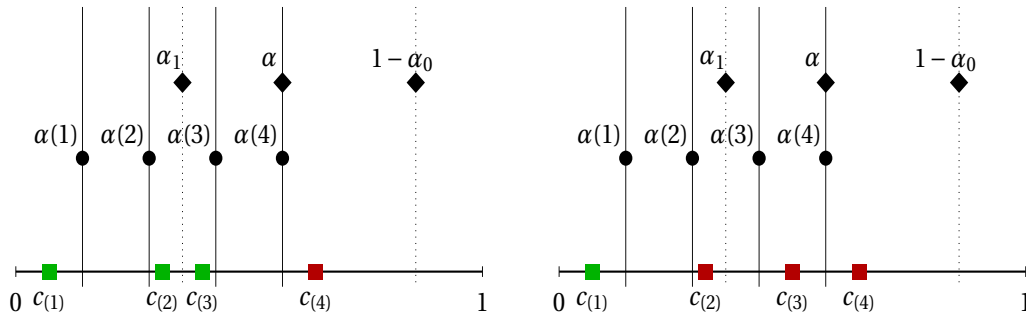


Figure 4.5 – The figure shows two examples. The green squares corresponds to the centers for which \mathcal{H}_0 was rejected and the red ones to the centers where it was not rejected. In both $c_{(1)}$ and $c_{(2)}$ the study was terminated after the first stage (their p-values are less than α_1). However, the decision to reject the null for $c_{(2)}$ depends on the results in the centers $c_{(3)}$ and $c_{(4)}$. On the left-hand side, the null is rejected at $c_{(2)}$ and $c_{(3)}$. On the right-hand side, decisions about $c_{(2)}$, $c_{(3)}$, $c_{(4)}$ are all non-rejections.

assume patient enrollment in all centres is done independently, and that all the p-values are independent.

In the single center design one rejects \mathcal{H}_0 if the p-value is below α , but to control the family wise type I error rate in the multi-center context at the level α , we need a multiple testing procedure, such as Hochberg’s step-up procedure Hochberg [1988]. The whole design is then described as follows:

- Let $p\text{-value}_1, p\text{-value}_2, \dots, p\text{-value}_M$ be the set of p-values at the end of each study (see Eq.D2). Put the p-values in an ascending order:
 $p\text{-value}(1) \leq p\text{-value}(2) \leq \dots \leq p\text{-value}(M)$.
- Denote the index of the center corresponding to $p\text{-value}(k)$ as $c_{(k)}$. (D3)
- Given the thresholds, $\alpha(1) \leq \alpha(2) \leq \dots \leq \alpha(M)$, find the index K corresponding to the largest k such that $p\text{-value}(k) \leq \alpha(k)$.

In Hochberg’s step-up procedure $\alpha(k) = \frac{\alpha}{M + 1 - k}$.

- The centers $c_{(1)}$ through $c_{(K)}$ are then judged to have sensitive subpopulations, and the null hypothesis is rejected for them.

There is a complication for the two-stage multicenter design. In the case of a single center, if the p-value $< \alpha_1$, then \mathcal{H}_0 is rejected after the first stage. When more than one center is involved, the condition $p\text{-value}_i < \alpha_1$, which only occurs when we stop after the first stage, does not necessarily mean that for this center the null is rejected. If the design parameter α_1 is such that $\alpha(1) > \alpha_1$, rejection always occurs. Otherwise, rejection depends on the other p-values or, more precisely, on their arrangement with regard to $\alpha(k)$, which is only known at the completion of the second stage.

Chapter 4. Randomized controlled trial (RCT) designs

Consider two p-value realisations shown in Fig. 4.5. There are two centers, $c_{(1)}, c_{(2)}$ where the trial was stopped after the first stage, and their p-values are less than α_1 . In two other centers the second stage was conducted and their p-values are in the interval $(\alpha_1, 1 - \alpha_0)$. For $c_{(1)}$ the null is rejected after the first stage in both situations, because $\text{p-value}_1 < \alpha(1)$, but the decision about $c_{(2)}$ depends on the results in the centers where the second stage is conducted. To summarize, if for some center after the first stage $\alpha(k) < \text{p-value}(k) < \alpha_1$, the final decision about this center might depend on the results from the centers where the trial continues to the second stage.

Recall that the type II error is defined as "at least m false nulls with strong effect among M_1 centers ($m \leq M_1 \leq M$) are not rejected". Denote the maximum type II error rates as $\beta_{\text{fw}}^{\text{se}}(M_1, m)$. It is complicated to calculate $\beta_{\text{fw}}^{\text{se}}(M_1, m)$ exactly for arbitrary couples (M_1, m) , but the upper bound can be obtained rather straightforwardly. To do this, we introduce auxiliary values

$$\beta_j^{\text{se}} = \begin{cases} \beta^{\text{se}}(n, \alpha(j)), & \text{one-stage design,} \\ \beta_2^{\text{se}}(D, \alpha(j)), & \text{two-stage design.} \end{cases} \quad (4.9)$$

Lemma 4.4. *The following inequality always holds:*

$$1 - \beta_{\text{fw}}^{\text{se}}(M_1, m) \geq \left(1 - \beta_{M_1+1-m}^{\text{se}}\right)^{M_1+1-m}. \quad (4.10)$$

If $M_1 = M$ and $m = 1$, equality is achieved.

Proof. Consider an arbitrary set of p-values: $\text{p-value}_1, \dots, \text{p-value}_M$ of which M_1 correspond to the alternatives with the strong effect. Let $\beta_{\text{fw}}^{\text{se}}(M_1, m)$ be the type II error rate for these p-values. We will show that

$$1 - \beta_{\text{fw}}^{\text{se}}(M_1, m) \geq \left(1 - \beta_{M_1+1-m}^{\text{se}}\right)^{M_1+1-m}.$$

Then, for the set of p-values where the maximum type II error rate is achieved, the statement in the lemma holds.

Let S_1 be the set of indexes for the centres with the strong effect. Define \tilde{k} as the rank of the m -th highest p-value among S_1 ,

$$\sum_{c(k) \in S_1} \mathbb{1}(k \geq \tilde{k}) = m. \quad (4.11)$$

First, we will prove that

$$1 - \beta_{\text{fw}}^{\text{se}}(M_1, m) \geq \mathbb{P}(\text{p-value}(\tilde{k}) \leq \alpha(M_1 + 1 - m)). \quad (4.12)$$

By definition, $\tilde{k} \geq M_1 + 1 - m$. The condition $\text{p-value}(\tilde{k}) \leq \alpha(M_1 + 1 - m)$ is sufficient to reject at least $M_1 + 1 - m$ null hypotheses for the centers with strong effect and therefore to avoid

type II error. The second inequality we have to prove is

$$\mathbb{P}(\text{p-value}(\tilde{k}) \leq \alpha(M_1 + 1 - m)) \geq \left(1 - \beta_{M_1+1-m}^{se}\right)^{M_1+1-m}. \quad (4.13)$$

For the proof we will use S_2 , an arbitrary subset of centers with strong effects such that $|S_2| = M_1 + 1 - m$:

$$\begin{aligned} \mathbb{P}(\text{p-value}(\tilde{k}) \leq \alpha(M_1 + 1 - m)) &= \mathbb{P}\left(\sum_{c_k \in S_1} \mathbb{1}(\text{p-value}(k) \leq \alpha(M_1 + 1 - m)) \geq M_1 + 1 - m\right) \\ &\geq \mathbb{P}\left(\sum_{i \in S_2} \mathbb{1}(\text{p-value}_i \leq \alpha(M_1 + 1 - m)) \geq M_1 + 1 - m\right) \geq \left(1 - \beta_{M_1+1-m}^{se}\right)^{M_1+1-m}. \end{aligned} \quad (4.14)$$

If $M_1 = M$ and $m = 1$ then the inequality (4.12) becomes an equality, because in this case $\tilde{k} = M$ and condition $\text{p-value}(M) \leq \alpha(M)$ is necessary and sufficient to reject all null hypotheses. Also if $\mathbb{P}(\text{p-value}_i \leq \alpha(M)) = 1 - \beta_M^{se}$ for all i then the inequality (4.13) becomes an equality :

$$\mathbb{P}(\text{p-value}(M) \leq \alpha(M)) = \mathbb{P}(\text{p-value}_i \leq \alpha(M), \forall i = 1, \dots, M) = \left(1 - \beta_M^{se}\right)^M. \quad (4.15)$$

□

Comment: If $M_1 < M$ or $m > 1$, equality in (4.12) cannot be attained, because there is a non-zero probability of having

$$\tilde{k} = M_1 + 1 - m, \text{ p-value}(\tilde{k}), \text{p-value}(\tilde{k} + 1) \in (\alpha(M_1 + 1 - m), \alpha(M_1 + 2 - m)],$$

in this case there is no type II error.

From this lemma, we conclude that

$$\beta_{\text{fw}}^{se}(M_1, m) \leq 1 - \left(1 - \beta_{M_1+1-m}^{se}\right)^{M_1+1-m},$$

and

$$\beta_{\text{fw}}^{se}(M, 1) = 1 - \left(1 - \beta_M^{se}\right)^M.$$

4.3 Planning the trial

In this section we present a framework which helps the user to choose the design parameters in order to control the type II error rate below some given level β_{\max} or the power above $(1 - \beta_{\max})$. For the multi-centre trial we suggest to control $\beta_{\text{fw}}^{se}(M, 1)$. Here we assume $\beta_{\max} < 0.5$ for the single center RCT and $\beta_{\max} < 1 - 0.5^M$ for the multicenter RCT. At the beginning the user should carefully define the region of strong effect as in Definition 4.1. Next, the user chooses β_{\max} and the design parameters. We also provide additional information that may influence

the user's choice of β_{\max} . In the case of changing β_{\max} , one should recalculate the design parameters. To illustrate the planning procedure, we use the region of strong effect as depicted in Fig. 4.6 and $\beta_{\max} = 0.2$.

4.3.1 Planning a single center RCT

One-stage design: choosing the sample size

The planning of a single center trial proceeds by choosing α and computing the sample size that ensures the desired minimal power. After α is chosen (here and elsewhere $\alpha = 0.05$), n is computed as

$$n = \min(n : \beta^{se}(n, \alpha) \leq \beta_{\max}). \quad (4.16)$$

To compute $\beta^{se}(n, \alpha)$, we use an asymptotic approximation of $\beta(n, \eta, \mu, p)$ based on the central limit theorem. Notice that Y_1^T, \dots, Y_n^T are i.i.d. random variables with $\mathbb{E}(Y_i^T) = \mu p$ and $\text{Var}(Y_i^T) = 1 + (1 - p)p\mu^2$. Hence,

$$\overline{Y^T} \sim \mathcal{N}\left(\mu p, \frac{1 + (1 - p)p\mu^2}{n}\right), \quad \overline{X} \sim \mathcal{N}\left(\mu p, \frac{1 + (1 - p)p\mu^2}{n} + \frac{1}{n}\right).$$

With Barry–Essen's upper bound (Berry [1941]) on the absolute deviation of the probability of a false negative from its normal approximation, we have

$$\left| \beta(n, \eta, \mu, p) - \Phi\left(\frac{\sqrt{n}(\eta - \mu p)}{\sqrt{2 + (1 - p)p\mu^2}}\right) \right| \leq \frac{C}{\sqrt{n}}. \quad (4.17)$$

We use this normal approximation in the following computations, because the real difference in (4.17) is negligible even for small sample sizes. From Lemma 4.2,

$$\beta^{se}(n, \alpha) \approx \max_{i=1, \dots, s} \Phi\left(\frac{\sqrt{2}z_{1-\alpha} - \sqrt{n}\mu_i p_i}{\sqrt{2 + (1 - p_i)p_i\mu_i^2}}\right)$$

and

$$n \approx \max_{i=1, \dots, s} \left(\frac{\sqrt{2}z_{1-\alpha} + z_{1-\beta_{\max}} \sqrt{2 + (1 - p_i)p_i\mu_i^2}}{\mu_i p_i} \right)^2. \quad (4.18)$$

Example

Let the region \mathcal{E} be as in Fig. 4.6 and let the maximum type II error rate be 0.2. The minimum sample size for the trial with these parameters from (4.18) is $n \approx 85.3$, whereas the exact value from (4.16) is 86. For $n = 86$ and $\alpha = 0.05$, the rejection threshold $\eta = 0.251$ and it is

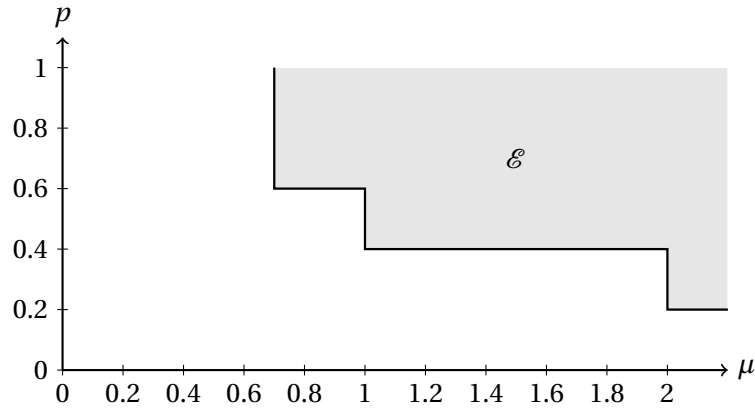


Figure 4.6 – Region of strong effect, characterised by $\bar{\mu} = (2, 1, 0.7)$, $\bar{p} = (0.2, 0.4, 0.6)$.

straightforward to conduct the trial (D1).

Two-stage design

The design is determined by $D = \{n_1, \alpha_0, \alpha_1, n_2\}$. After α is chosen, the sample size at the second stage is computed as

$$n_2 = \min \left(n : \beta_2^{se}(D, \alpha) \leq \beta_{\max} \right). \quad (4.19)$$

To approximate n_2 , we use the following approximation for the type two error for the two-stage design:

$$\beta_2(D, \alpha, \mu, p) \approx \Phi \left(\frac{\sqrt{n_1}(\eta_0 - \mu p)}{\sqrt{2 + (1-p)p\mu^2}} \right) + \int_{\eta_0}^{\eta_1} \frac{\varphi \left(\frac{\sqrt{n_1}(x_1 - \mu p)}{\sqrt{2 + (1-p)p\mu^2}} \right)}{\sqrt{\frac{2 + (1-p)p\mu^2}{n_1}}} \Phi \left(\frac{\frac{n_1 + n_2}{n_2} \eta_2(D, \alpha) - \frac{n_1}{n_2} x_1 - \mu p}{\sqrt{\frac{2 + (1-p)p\mu^2}{n_2}}} \right) dx_1.$$

We next compute n_2 numerically. Note that it is possible that no design fulfills the condition on β_{\max} . In the following we will address this problem.

For planing purposes we next consider the expected sample size under the null,

$$q_0(D) = n_1 + (1 - \alpha_0 - \alpha_1) n_2 \quad (4.20)$$

and the analogous formula for the maximum expected sample size under the alternative is

$$q_1(D) = n_1 + \max_{\mu > 0, p > 0} \left(\beta \left(n_1, \frac{\sqrt{2} z_{1-\alpha_1}}{\sqrt{n_1}}, \mu, p \right) - \beta \left(n_1, \frac{\sqrt{2} z_{\alpha_0}}{\sqrt{n_1}}, \mu, p \right) \right) n_2. \quad (4.21)$$

Figure 4.7 depicts all of these values in a schematic way.

We set $\alpha_0 > 0.5$ as a condition for stopping the trail early in the case of the absence of the

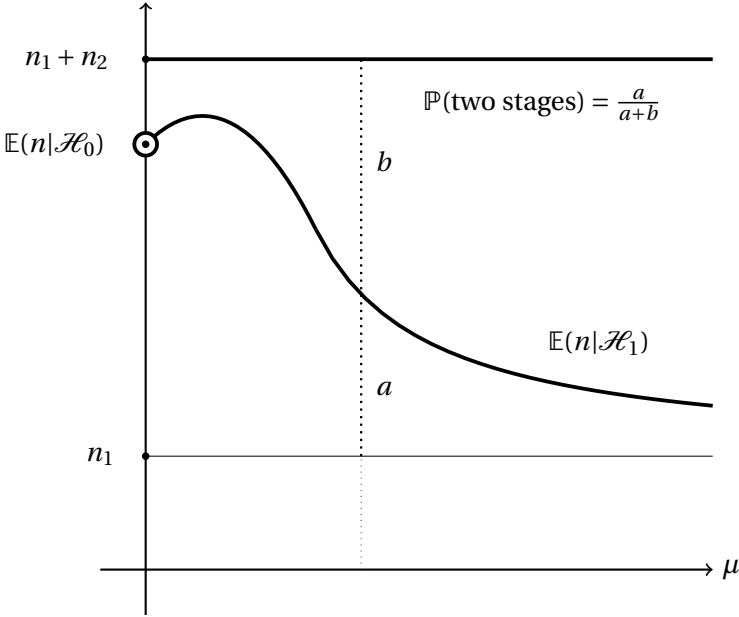


Figure 4.7 – Expected sample sizes under the null and under the alternative for constant p .

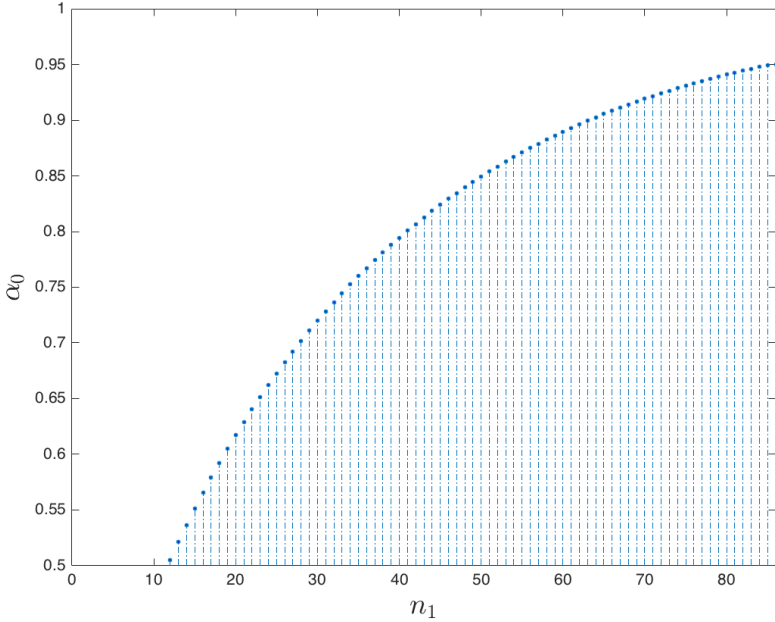


Figure 4.8 – Possible pairs of (n_1, α_0) for the two-stage design. The parameters are $\alpha = 0.05$, $\beta_{\max} = 0.2$, $n = 86$; $12 \leq n_1 \leq 86$.

effect. With this choice, both $z_{1-\alpha_1}$ and z_{α_0} are positive.

Lemma 4.5. *If $0 < z_{\alpha_0} < z_{1-\alpha_1}$, the maximum probability of conducting the second stage in (4.21) approximately satisfies*

$$\max_{\mu>0, p>0} \left(\beta \left(n_1, \frac{\sqrt{2} z_{1-\alpha_1}}{\sqrt{n_1}}, \mu, p \right) - \beta \left(n_1, \frac{\sqrt{2} z_{\alpha_0}}{\sqrt{n_1}}, \mu, p \right) \right) \approx 2\Phi \left(\frac{z_{1-\alpha_1} - z_{\alpha_0}}{2} \right) - 1. \quad (4.22)$$

The maximum is achieved for $\mu = \frac{z_{1-\alpha_1} + z_{\alpha_0}}{\sqrt{2n_1}}$, $p = 1$.

Proof. To show this, we bound the probability of conduction the second stage as follows:

$$\begin{aligned} & \max_{\mu>0, p>0} \left(\beta \left(n_1, \frac{\sqrt{2} z_{1-\alpha_1}}{\sqrt{n_1}}, \mu, p \right) - \beta \left(n_1, \frac{\sqrt{2} z_{\alpha_0}}{\sqrt{n_1}}, \mu, p \right) \right) \approx \Phi \left(\frac{\sqrt{2} z_{1-\alpha_1} - \sqrt{n_1} \mu p}{\sqrt{2 + (1-p)p\mu^2}} \right) - \Phi \left(\frac{\sqrt{2} z_{\alpha_0} - \sqrt{n_1} \mu p}{\sqrt{2 + (1-p)p\mu^2}} \right) \\ & \leq \Phi \left(\frac{\sqrt{2} z_{1-\alpha_1} - (\sqrt{2} z_{1-\alpha_1} + \sqrt{2} z_{\alpha_0})/2}{\sqrt{2 + (1-p)p\mu^2}} \right) - \Phi \left(\frac{\sqrt{2} z_{\alpha_0} - (\sqrt{2} z_{1-\alpha_1} + \sqrt{2} z_{\alpha_0})/2}{\sqrt{2 + (1-p)p\mu^2}} \right) \\ & = 2\Phi \left(\frac{\sqrt{2} z_{1-\alpha_1} - \sqrt{2} z_{\alpha_0}}{\sqrt{2 + (1-p)p\mu^2}} \right) - 1 \leq 2\Phi \left(\frac{z_{1-\alpha_1} - z_{\alpha_0}}{2} \right) - 1. \end{aligned}$$

Notice, that an equality is achieved for $\mu = \frac{z_{1-\alpha_1} + z_{\alpha_0}}{\sqrt{2n_1}} > 0$ and $p = 1$. This ends the proof. \square

From Lemma 4.5, it follows that $q_1(D) \approx n_1 + \left(2\Phi \left(\frac{z_{1-\alpha_1} - z_{\alpha_0}}{2} \right) - 1 \right) n_2$.

We let the user choose n_1 and α_0 , but we suggest setting α_1 such that it minimizes $q_1(D)$ for given n_1, α_0 . This leads to the choice $\alpha_1 = \underset{\alpha_1}{\operatorname{argmin}} q_1(D)$. Once n_1, α_0, α_1 have been fixed, all the other quantities including n_2, q_0, q_1 can be computed.

In the following text we discuss guidance for the user on the choice of n_1 and α_0 . We suppose that the choice of n_1, α_0 is based on q_0, q_1, n_2 and n . We note that the design does not exist for every pair (n_1, α_0) . With $0.5 < \alpha_0 < 1 - \alpha$ and (4.8), we have $\beta^{se}(n_1, 1 - \alpha_0) < \beta_2^{se}(D, \alpha) \leq \beta_{\max}$, and with an approximation for $\beta^{se}(n_1, 1 - \alpha_0)$, mentioned above,

$$\alpha_0 < \min_{i=1, \dots, s} \Phi \left(\sqrt{\frac{n_1}{2}} \mu_i p_i - z_{1-\beta_{\max}} \sqrt{1 + \frac{(1-p_i)p_i}{2} \mu_i^2} \right). \quad (4.23)$$

This in turn implies that

$$n_1 > \left(\max_{i=1, \dots, s} \frac{z_{1-\beta_{\max}} \sqrt{2 + (1-p_i)p_i \mu_i^2}}{\mu_i p_i} \right)^2. \quad (4.24)$$

Keep in mind also that the first stage sample size should satisfy $n_1 < n$, otherwise the two-stage scheme is less efficient than the one-stage design. Based on (4.23) and (4.24), the region

of possible (n_1, α_0) provided to the user is shown in Fig. 4.8.

Among all possible pairs (n_1, α_0) an optimal selection would make q_0 , q_1 and $n_1 + n_2$ min small. However, these criteria cannot be achieved simultaneously. This trade-off is left for the user to resolve, and in Fig. 4.9 we provide an example of a graphical aid for decision making. Recall that the design is built to ensure the type II error rate is not greater than β_{\max} in all the corners of \mathcal{E} . However, this probability can be very sensitive to μ and/or p . We assume that given the probability of a false negative for all (μ, p) , for example, as in Fig. 4.10, the user may decide to change β_{\max} . If these rates are small, the trial size may be reduced by using a larger β_{\max} , and if β_{\max} is considered to be reduced, then the larger trial is necessary. If he decides to change β_{\max} , the design needs to be elaborated from the beginning.

4.3.2 Planning a multicenter RCT

The design is determined by the set $\{M, n_1, \alpha_0, \alpha_1, n_2\}$. For a multicenter trial we build the design to control $\beta_{\text{fw}}^{\text{se}}(M, 1)$. The type II error rates $\beta_{\text{fw}}^{\text{se}}(M_1, m)$ where $M_1 < M$ or $m > 1$ are not controlled. From Lemma 4.4, the condition $\beta_{\text{fw}}^{\text{se}}(M, 1) \leq \beta_{\max}$ is equivalent to $\beta_M^{\text{se}} \leq 1 - (1 - \beta_{\max})^{\frac{1}{M}}$. This means that if we consider M centers separately, the design in each of the centers will correspond to the single center RCT with the type I and type II errors controlled at levels $\alpha(M)$, the largest threshold of the multiple rejection procedure, and $\beta_M^{\text{se}} = 1 - (1 - \beta_{\max})^{\frac{1}{M}}$, respectively.

To illustrate the planning for a RCT with four centers, suppose that FWER is controlled at 0.05 and the family-wise type II error rate is controlled at $\beta_{\max} = 0.2$. With Hochberg's multiple testing procedure the design is then determined by $\alpha(4) = 0.05$ and $\beta_M^{\text{se}} = 1 - (1 - 0.2)^{\frac{1}{4}} \approx 0.054$. As in the single center RCT, the choice of n_1 and α_0 is made by the user based on the information provided in Fig. 4.11.

Nonetheless, maximum type II error rates for $M_1 < M$ or $m > 1$ can be also of interest for the study. Even if an effect exists, it is not necessarily strong in all centers. Furthermore, a trial missing a few centers with strong effect can still be useful if in the other centers the strong effect is detected. Hence, given the information about $\beta_{\text{fw}}^{\text{se}}(M_1, m)$, the user might change β_{\max} and recalculate the design. In Lemma 4.4, we show an upper bound on these errors,

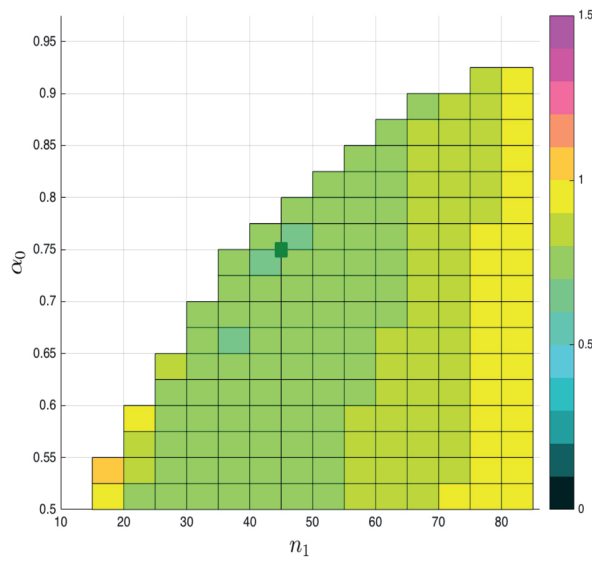
$$\beta_{\text{fw}}^{\text{se}}(M_1, m) \leq 1 - \left(1 - \beta_{M_1+1-m}^{\text{se}}\right)^{M_1+1-m}.$$

Once the parameters n_1, α_0 have been chosen, the user is supplied with Table 4.1(a).

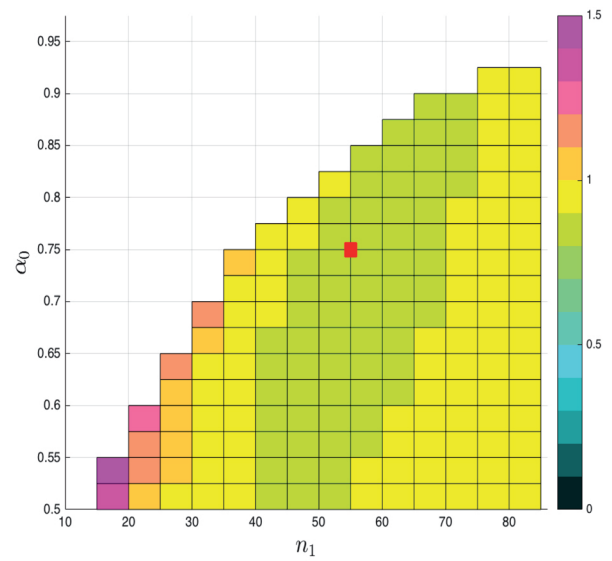
Type II error rate for the multi-center design computation

We do not know an exact value of $\beta_{\text{fw}}^{\text{se}}(M_1, m)$, but we can compute the type II error rate $\beta_{\text{fw}}(M_1, m)$ for the given responses. Let $\alpha(1) < \alpha(2) < \dots < \alpha(M)$ be the p-value thresholds for the rejection procedure. For convenience put $\alpha(0) = 0$, $\alpha(M+1) = 1$. To compute $\beta_{\text{fw}}(M_1, m)$,

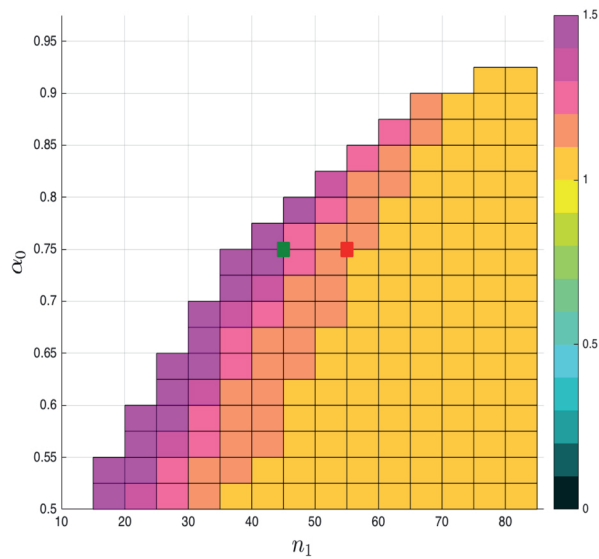
4.3. Planning the trial



(a) $\frac{q_0}{n}$. $\min q_0 = 60$ for $n_1 = 45$, $\alpha_0 = 0.75$.



(b) $\frac{q_1}{n}$. $\min q_1 = 75$ for $n_1 = 55$, $\alpha_0 = 0.75$



(c) $\frac{n_1 + n_2}{n}$

Figure 4.9 – The diagnostic plot for choosing the design. The parameters are $\alpha = 0.05$, $\beta_{\max} = 0.2$, $\vec{\mu} = (2, 1, 0.7)$, $\vec{p} = (0.2, 0.4, 0.6)$, $n = 86$. The color coding indicates the percentage gain in sample size compared to the one stage design. Values below one indicate a gain, and values above 1 indicate a loss. The green square in (a) is a pair of parameters that attains the minimum of q_0 , whereas the red square in (b) attains the minimum for q_1 . These points are also shown in (c). For the choice $n_1 = 55$, $\alpha_0 = 0.7$, $\alpha_1 = 0.026$, $n_2 = 38$, $\eta_0 = 0.10$, $\eta_1 = 0.37$, $\eta_2 = 0.26$, we have $q_0 = 66$, $q_1 = 75$, $n_1 + n_2 = 93$.

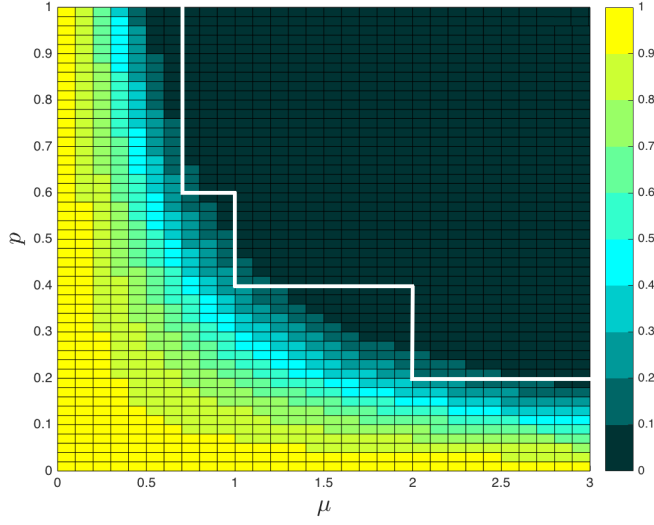


Figure 4.10 – Probability of a false negative computed for $n_1 = 55$, $\alpha_0 = 0.7$, $\alpha_1 = 0.026$, $n_2 = 38$.

one should look at the disposition of the p-values inside the intervals $(\alpha(j-1), \alpha(j)]$, $j = 1, 2, \dots, M+1$.

Define $\beta_{i,j} = \mathbb{P}(\text{p-value}_i \geq \alpha(j))$ to be the type II error for the i -th center evaluated at $\alpha = \alpha(j)$ (see Fig. 4.12). The probability that the p-value of the i -th center is in the interval $(\alpha(j-1), \alpha(j)]$ is

$$\mathbb{P}(\text{p-value}_i \in (\alpha(j-1), \alpha(j)]) = \beta_{i,j-1} - \beta_{i,j}, \quad i = 1, \dots, M; \quad j = 1, \dots, M+1. \quad (4.25)$$

Notice that if p-values for some centres are in the same interval $(\alpha(j), \alpha(j-1)]$, the decision about \mathcal{H}_0 there will be the same.

Let the random index j_i define an interval of the i -th center p-value,

$$j_i = (j \mid \text{p-value}_i \in (\alpha(j-1), \alpha(j)]), \quad j_i = 1, \dots, M+1. \quad (4.26)$$

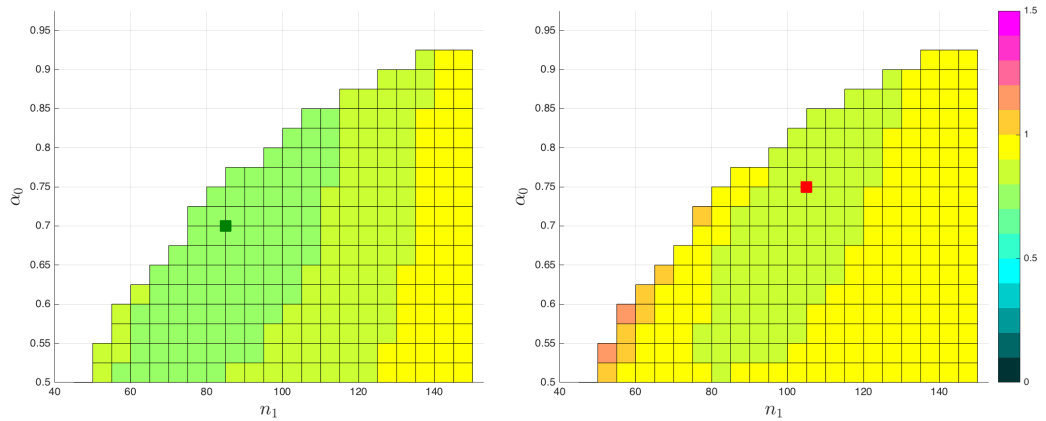
The probability of observing $\{j_1, \dots, j_M\}$ is then expressed as

$$\prod_{i=1}^M (\beta_{i,j_i-1} - \beta_{i,j_i}), \quad \text{where } 1 = \beta_{i,0} > \beta_{i,1} > \dots > \beta_{i,M} > \beta_{i,M+1} = 0. \quad (4.27)$$

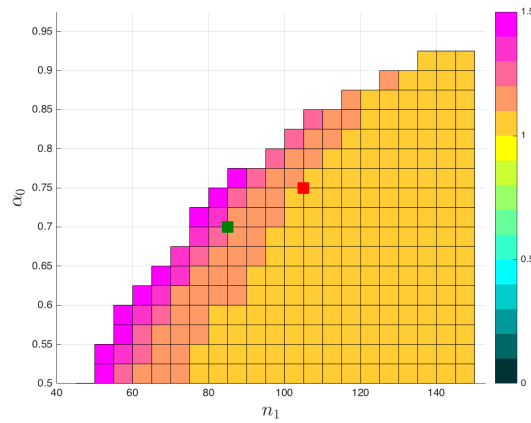
Define e_j to be the number of p-values falling in the interval $(0, \alpha(j)]$,

$$e_j = e_j(j_1, \dots, j_M) = \sum_{i=1}^M \mathbb{1}(j_i \leq j). \quad (4.28)$$

Using e_j , an alternative formulation of Hochberg's step-up procedure is as follows:



(a) $\frac{q_0}{n}$. $\min q_0 = 113$ for $n_1 = 85$, $\alpha_0 = 0.7$. (b) $\frac{q_1}{n}$. $\min q_1 = 134$ for $n_1 = 105$, $\alpha_0 = 0.75$



(c) $\frac{n_1+n_2}{n_{\min}}$

Figure 4.11 – The diagnostic plot for choosing the design. The parameters are $M = 4$, $\alpha = 0.05$, $\beta_{\max} = 0.2$, $\vec{\mu} = (2, 1, 0.7)$, $\vec{p} = (0.2, 0.4, 0.6)$, $n = 153$. For the parameters $n_1 = 100$, $\alpha_0 = 0.7$, $\alpha_1 = 0.026$, $n_2 = 65$, $\eta_0 = 0.07$, $\eta_1 = 0.28$, $\eta_2 = 0.19$, we have $q_0 = 118$, $q_1 = 134$, $n_1 + n_2 = 165$.

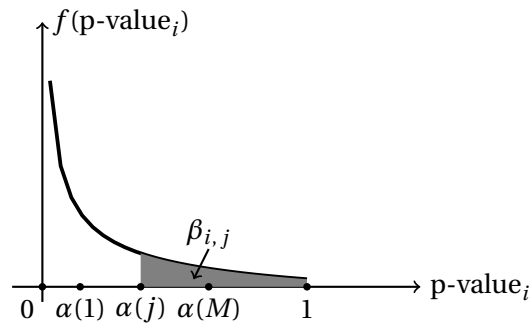


Figure 4.12 – $\beta_{i,j}$ for the multiple testing procedure.

Chapter 4. Randomized controlled trial (RCT) designs

Table 4.1 – $\beta_{fw}(M_1, m)$. The design parameters are $\alpha = 0.05$, $\beta_{\max} = 0.2$, $n_1 = 100$, $\alpha_0 = 0.7$, $\alpha_1 = 0.026$, $n_2 = 65$. Real parameters in the centers with strong effects are (μ^*, p^*) , for the other centers \mathcal{H}_0 is true.

		M_1			
		1	2	3	4
m	1	0.305	0.469	0.534	0.200
	2		0.305	0.469	0.534
	3			0.305	0.469
	4				0.305

(a) Upper bound on $\beta_{fw}^{se}(M_1, m)$.

		M_1						M_1			
		1	2	3	4			1	2	3	4
m	1	0.303	0.464	0.515	0.200	m	1	0.032	0.050	0.050	0.002
	2		0.091	0.170	0.099		2		0.001	0.002	< 0.001
	3			0.026	0.030		3			< 0.001	< 0.001
	4				0.004		4				< 0.001

(b) $\beta_{fw}(M_1, m)$ for $(\mu^*, p^*) = (2, 0.2)$.

(c) $\beta_{fw}(M_1, m)$ for $(\mu^*, p^*) = (1.2, 0.5)$.

- Compute e_j for each $j = 1, \dots, M$;
- Find the maximum J such that $e_j \geq J$;
- Reject \mathcal{H}_0 in centres c_k , $k = 1, \dots, e_J$.

Let S_1 be the set of indexes for the centers with the strong effects. Define

$$\tilde{j} = \min \left(j \mid \sum_{i \in S_1} \mathbb{1}(j_i > j) < m \right) \quad (4.29)$$

Now an expression for $\beta_{fw}(M_1, m)$ can be written as

$$\beta_{fw}(M_1, m) = \mathbb{P}(e_j < j, \text{ for all } j \geq \tilde{j}). \quad (4.30)$$

Consequently, one can compute the type II error rate as

$$\beta_{fw}(M_1, m) = \left(\sum_{j_1=1}^{M+1} \dots \sum_{j_M=1}^{M+1} \right) \prod_{i=1}^M (\beta_{i, j_{i-1}} - \beta_{i, j_i}), \quad (4.31)$$

$e_j < j \text{ for all } j \geq \tilde{j}$

which is the sum of the probabilities of all realisations $\{j_1, \dots, j_M\}$, such that $e_j < j$ for all $j \geq \tilde{j}$.

The user will have a better intuition about $\beta_{fw}^{se}(M_1, m)$ if we provide him with $\beta_{fw}(M_1, m)$

		M_1			
		1	2	3	4
m	1	0.303 0.294	0.464 0.498	0.515 0.523	0.200 0.185
	2		0.091 0.099	0.170 0.179	0.100 0.079
	3			0.026 0.032	0.030 0.032
	4				0.004 0.007

(a) For the centres with strong effect
 $(\mu^*, p^*) = (2, 0.2)$.

		M_1			
		1	2	3	4
m	1	0.032 0.041	0.050 0.044	0.050 0.043	0.002 0.005
	2		0.001 0.001	0.002 0	< 0.001 0
	3			< 0.001 0	< 0.001 0
	4				< 0.001 0

(b) For the centres with strong effect
 $(\mu^*, p^*) = (1.2, 0.5)$.

Table 4.2 – Theoretical (top) and empirical (bottom) $\beta_{fw}(M_1, m)$ for 1000 simulations. The parameters are $M = 4$, $\alpha = 0.05$, $\beta_{\max} = 0.2$, $n_1 = 100$, $\alpha_0 = 0.7$, $\alpha_1 = 0.026$, $n_2 = 65$.

computed for different sets of responses. We take M_1 centres with the same $(\mu^*, p^*) \in \mathcal{E}$, while in the remaining $M - M_1$ centres \mathcal{H}_0 is true (see Table 4.1 (b),(c)). We suggest that the user checks more than one pair (μ^*, p^*) to have a better understanding of the possible type II error rates. Based on these tables, the user decides whether to change β_{\max} .

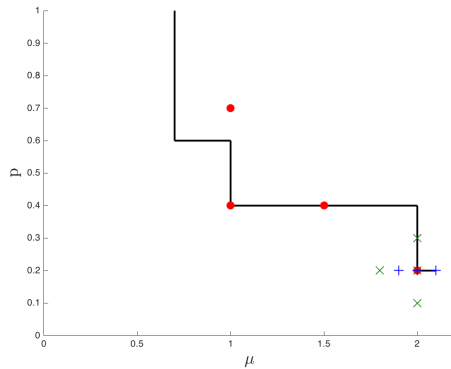
As illustrated, the values of $\beta_{fw}(M_1, m)$ are very sensitive to the true parameters (μ^*, p^*) . In Tab.4.1(b), we took $\mu^* = 2$, $p^* = 0.2$, which is the lowest corner of \mathcal{E} , while in Tab.4.1(c), $\mu^* = 1.2$, $p^* = 0.5$ is inside \mathcal{E} but not far away from the boundary. These numbers show that $\beta_{fw}(M_1, m)$ decreases rapidly when μ^* and/or p^* increases.

4.4 Examples

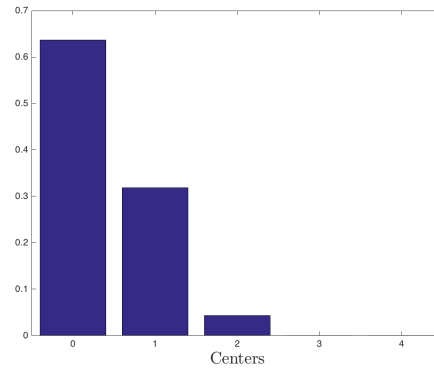
To validate the model, we take the parameters from the previous section: $M = 4$, $\alpha = 0.05$, $\beta_{\max} = 0.2$, $n_1 = 100$, $\alpha_0 = 0.7$, $\alpha_1 = 0.026$, $n_2 = 65$. For different M_1 and m we simulate the trial 1000 times. The responses in the control and treatment groups are, $Z^C \sim \mathcal{N}(0, 1)$, $Z^T \sim (1 - p^*)\mathcal{N}(0, 1) + p^*\mathcal{N}(\mu^*, 1)$. We estimate the standard deviation as $\hat{\sigma} = \sqrt{\frac{\sum_{i=1}^n (Z_i^C - \hat{\mu}^C)^2}{n-1}}$. The results are summarized in Tab.4.2.

Consider now three different sets of true parameters (μ, p) as in Fig. 4.13. For the optimal design, the probability of conducting the second stage (and, implicitly, the number of patients enrolled) is shown in Fig. 4.13. In Table 4.3 we show the results of the simulations for different *non-optimal* designs, which means that the values (n_1, α_0) are chosen without optimization, which assumed the worst case scenario under the alternative; see Section 4.3. The results of the simulated clinical trials and theoretical predictions agree. The scenario where all the centers have the same effect size attains the largest type II error rate. The last section in Table 4.3 demonstrates, indeed, the maximum type II error rate of 20%. Though these designs are not optimal in the sense that we declared before, they can still be applied if they satisfy any

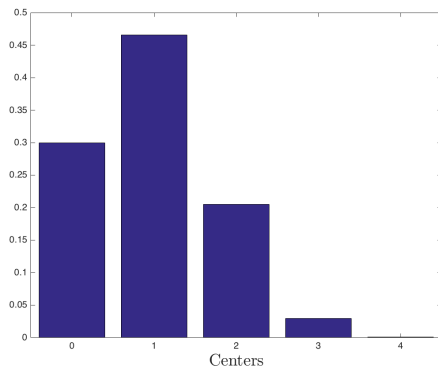
other restrictions or objectives of the trial.



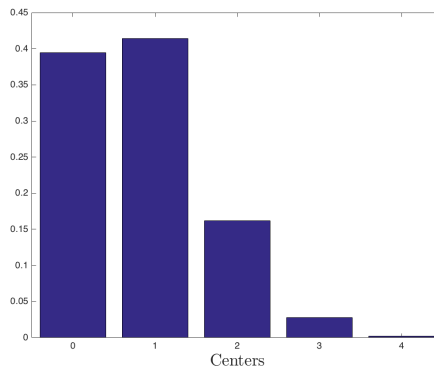
(a) Three parameters sets (PS).



(b) $PS_1 = \{(2, 0.2), (1.5, 0.4), (1, 0.4), (1, 0.7)\}$.



(c) $PS_2 = \{(2, 0.2), (2, 0.3), (2, 0.1), (1.8, 0.2)\}$.



(d) $PS_3 = \{(2, 0.2), (2.1, 0.2), (1.9, 0.2), (1.9, 0.2)\}$.

Figure 4.13 – In (a) we show three parameter sets employed for simulations. For them in (b), (c), (d) we present a frequency function of the number of groups with second stage conducted. The parameters are $M = 4$, $\alpha = 0.05$, $\beta_{\max} = 0.2$, $\vec{\mu} = (2, 1, 0.7)$, $\vec{p} = (0.2, 0.4, 0.6)$, $n_1 = 100$, $\alpha_0 = 0.7$, $\alpha_1 = 0.026$, $n_2 = 65$.

4.5 Discussion

Region of strong effect

The region of strong effect for the mixture response model was introduced as an instrument for decision making. Further standardization for regulatory reasons might be of interest. From our point of view, it may also assist in making the research on trial designs more coherent. In our framework we suggest to first determine the region of strong effect, while β_{\max} and other parameters may be tuned according to the needs of the study.

		n_1										
		70	75	80	85	90	95	100	105	110	115	120
		$PS_1 = \{(2, 0.2), (1.5, 0.4), (1, 0.4), (1, 0.7)\}$										
α_0	0.6	0.093 0.095	0.093	0.093	0.093	0.093	0.093	0.093	0.094	0.094	0.094	0.094
	0.7		0.094 0.082	0.093	0.093	0.093	0.093	0.092	0.093	0.093	0.093	0.094
	0.8						0.093 0.111	0.093	0.092	0.093	0.093	0.093
		$PS_2 = \{(2, 0.2), (2, 0.3), (2, 0.1), (1.8, 0.2)\}$										
α_0	0.6	0.113 0.105	0.237 0.246	0.143	0.204	0.134	0.175	0.107	0.111	0.139	0.129	0.120
	0.7		0.226 0.214	0.215	0.129	0.187	0.174	0.162	0.150	0.139	0.129	0.120
	0.8						0.169 0.177	0.160	0.149	0.138	0.129	0.120
		$PS_3 = \{(2, 0.2), (2.1, 0.2), (1.9, 0.2), (1.9, 0.2)\}$										
α_0	0.6	0.111 0.098	0.157	0.124	0.147	0.121	0.138	0.111	0.114	0.126	0.122	0.119
	0.7		0.156 0.147	0.152	0.119	0.142	0.138	0.134	0.130	0.126	0.122	0.119
	0.8						0.138 0.133	0.134	0.13	0.126	0.123	0.119
		All four centers have effect $\mu = 2, p = 0.2$										
α_0	0.6	0.200 0.201	0.200	0.200	0.200	0.199	0.199	0.199	0.2	0.199	0.200	0.200
	0.7		0.200 0.204	0.200	0.200	0.200	0.199	0.200	0.200	0.200	0.199	0.200
	0.8						0.200 0.192	0.200	0.200	0.199	0.200	0.199

Table 4.3 – Real, $\beta_{fw}(4, 1)$, (on the top of each cell) and empirical (on the bottom of each cell) type II error rates for different choices of (n_1, α_0) based on 1000 simulations.

Mean value statistic

The choice of the mean value statistic is not only motivated by its simplicity and its role in the classical RCT design. The mean value statistic also has some advantages for mixture models. If the responses in the control and treatment groups, Y^C and Y^T , have densities $f(y)$ and $(1-p)f(y) + pf(y-\mu)$, respectively, and $\text{Var}(Y^C) = 1$, then by the central limit theorem the following approximations hold

$$\overline{Y^T} \sim \mathcal{N}\left(\mu p, \frac{1 + (1-p)p\mu^2}{n}\right), \quad \overline{X} \sim \mathcal{N}\left(\mu p, \frac{1 + (1-p)p\mu^2}{n} + \frac{1}{n}\right).$$

This means that all the designs presented here may be used for non-normal responses.

In the two-stage design we use the criterion $\overline{X} > \eta_2$ to decide on rejection of the null after the second stage. It is not obvious why this test works well and why \overline{X}_1 and \overline{X}_2 are not combined in some other way, but the proposed test is close to the UMP. The likelihood ratio based on the joint distribution function of $(\overline{X}_1, \overline{X}_2)$ is

$$\begin{aligned} \frac{g_{(\overline{X}_1, \overline{X}_2)}(x_1, x_2)}{f_{(\overline{X}_1, \overline{X}_2)}(x_1, x_2)} &= \frac{\varphi\left(\sqrt{\frac{n_1}{2}} \frac{x_1 - \mu p}{\sqrt{1 + \frac{(1-p)p}{2}\mu^2}}\right) \varphi\left(\sqrt{\frac{n_2}{2}} \frac{x_2 - \mu p}{\sqrt{1 + \frac{(1-p)p}{2}\mu^2}}\right)}{\varphi\left(\sqrt{\frac{n_1}{2}} x_1\right) \varphi\left(\sqrt{\frac{n_2}{2}} x_2\right)} \approx \frac{\varphi\left(\sqrt{\frac{n_1}{2}}(x_1 - \mu p)\right) \varphi\left(\sqrt{\frac{n_2}{2}}(x_2 - \mu p)\right)}{\varphi\left(\sqrt{\frac{n_1}{2}} x_1\right) \varphi\left(\sqrt{\frac{n_2}{2}} x_2\right)} \\ &= e^{\frac{(x_1 n_1 + x_2 n_2)\mu p - \frac{1}{2}(\mu p)^2}{2}}, \end{aligned} \tag{4.32}$$

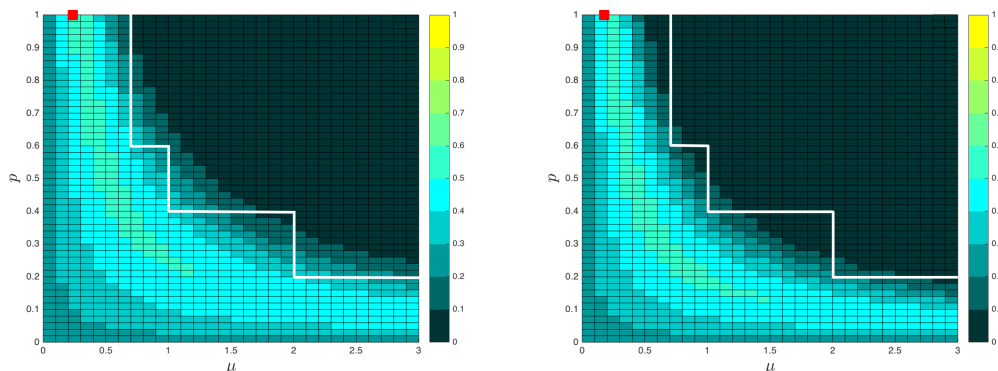
where $f(x_1, x_2)$ and $g(x_1, x_2)$ are pdfs under the null and alternative, respectively. The normal approximation in the likelihood works if the variance $1 + \frac{(1-p)p\mu^2}{2}$ is close to one. In this case likelihood ratio is the monotone non-decreasing function of $\frac{x_1 n_1 + x_2 n_2}{n_1 + n_2}$. The test becomes UMP when $p = 1$.

Expected sample size

As we have seen in Sec.4.3, the minimum value of the optimized q_1 is quite close to n . This value is reached for the 'worst' alternative $\mu = \frac{z_{1-\alpha_1} + z_{\alpha_0}}{\sqrt{2n_1}}$, $p = 1$, and for other values it can be substantially lower. Further in many situations we have much smaller expected sample sizes. To illustrate this, we plot the probability of conducting the second stage for single- and multicenter designs, see Fig. 4.14, presented as examples in Sec.4.3. This graph might be also of interest for the user during the design planning.

Rejection procedures for multicenter trials

For multicenter RCTs, we construct the design using Hochberg's step-up rejection procedure. However, one can use any step-up procedure with thresholds $\alpha(1), \dots, \alpha(M)$. For example, it could be Benjamini–Hochberg's procedure (Benjamini and Hochberg [1995]) with $\alpha(k) = \frac{k\alpha}{M}$



(a) Single center trial. The maximum probability is 0.52. (b) Multicenter trial. The maximum probability is 0.52.

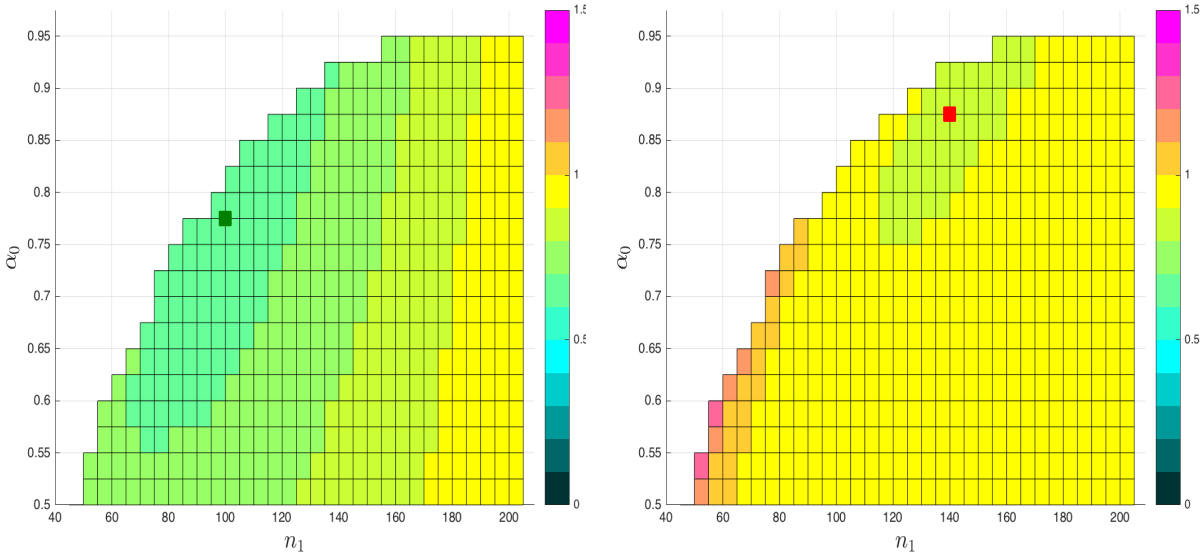
Figure 4.14 – The plot is in the (μ, p) space and shows the region of strong effect together with the probability of conducting the second stage. On the left is the case of a single center RCT and on the right is the case of 4 centers RCT. In each case the red dot, $(0.24, 1)$ on the left and $(0.17, 1)$ on the right, shows the worst combination of μ and p in terms of the expected sample size under the alternative. The design on the left is $n_1 = 55$, $\alpha_0 = 0.7$, $\alpha_1 = 0.026$, $n_2 = 38$, whereas on the right it has $M = 4$, $n_1 = 100$, $\alpha_0 = 0.7$, $\alpha_1 = 0.026$, $n_2 = 65$.

which controls the false discovery rate (FDR). There will be no difference in parameters $\{n_1, \alpha_0, \alpha_1, n_2\}$ for the designs using Hochberg's or Benjamini–Hochberg's procedures to control $\beta_{\text{fw}}^{\text{se}}(M, 1)$, because in each center the design corresponds to the single center RCT with the maximum type I error rate controlled at the same level $\alpha(M) = \alpha$ (see (4.9) and Lemma 4.4). If $M = 2$, these methods are fully equivalent, including the thresholds $\alpha(1) = \alpha/2$, $\alpha(2) = \alpha$. Nonetheless, if $M > 2$, $\beta_{\text{fw}}^{\text{se}}(M_1, m)$ is generally smaller for the Benjamini–Hochberg procedure (see Table 4.4).

To emphasize the importance of step-up procedures, consider the classical Bonferroni correction. For this procedure the decision is made for each center individually. However, the method may be reformulated as a step-up procedure with $\alpha(k) = \frac{\alpha}{M}$ for $k = 1, \dots, M$. An additional constraint for the two-stage design will be $\alpha_1 < \frac{\alpha}{M}$. If this is not satisfied, the second stage is not needed. In Fig. 4.15, we give a diagnostic plot similar to those in Fig. 4.11 for the Bonferroni method. First, the sample size for the one-stage design is $n = 209$, while for Hochberg's and Benjamini–Hochberg's procedures $n = 153 < 209$.

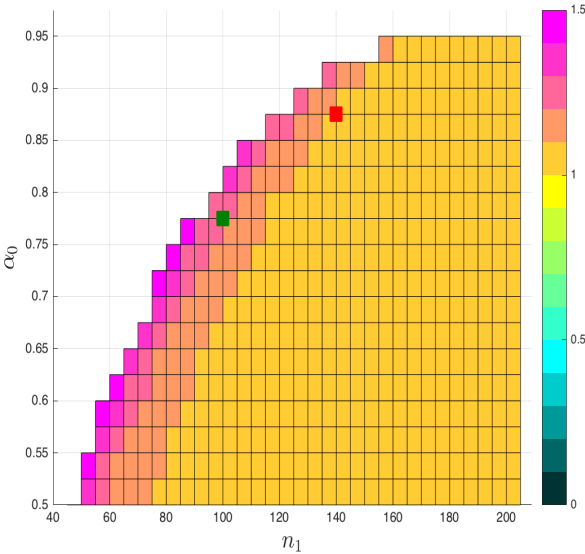
The minimum values of q_0 , q_1 , $n_1 + n_2$ for the Bonferroni method are also considerably higher than for the step-up procedures ($q_0 = 118$, $q_1 = 134$, $n_1 + n_2 = 165$). Therefore, the designs using Hochberg's step-up procedure control the same FWER as the Bonferroni and is clearly superior in terms of effort.

Chapter 4. Randomized controlled trial (RCT) designs



(a) $\frac{q_0}{n}$. $\min q_0 = 134$ for $n_1 = 100$, $\alpha_0 = 0.775$.

(b) $\frac{q_1}{n}$. $\min q_1 = 185$ for $n_1 = 140$, $\alpha_0 = 0.875$.



(c) $\frac{n_1+n_2}{n}$. $\min(n_1 + n_2) = 209$ for $n_1 = 205$, $\alpha_0 = 0.55$.

Figure 4.15 – The diagnostic plot for choosing the design. The parameters are $M = 4$, $\alpha = 0.05$, $\beta_{\max} = 0.2$, $\vec{\mu} = (2, 1, 0.7)$, $\vec{p} = (0.2, 0.4, 0.6)$. For the Bonferroni procedure $n = 209$.

		M_1			
		1	2	3	4
m	1	0.303 0.298	0.464 0.380	0.515 0.200	0.200 0.200
	2		0.091 0.082	0.170 0.069	0.099 0.027
	3			0.026 0.014	0.030 0.009
	4				0.004 0.002

(a) For the centres with strong effect
 $(\mu^*, p^*) = (2, 0.2)$.

		M_1			
		1	2	3	4
m	1	0.032 0.032	0.050 0.033	0.050 0.030	0.002 0.002
	2		0.001 < 0.001	0.002 < 0.001	< 0.001 < 0.001
	3			< 0.001 < 0.001	< 0.001 < 0.001
	4				< 0.001 < 0.001

(b) For the centres with strong effect
 $(\mu^*, p^*) = (1.2, 0.5)$.

Table 4.4 – $\beta_{fw}(M_1, m)$ for Hochberg’s (top) and Benjamini–Hochberg’s (bottom) rejection rules. The parameters are $M = 4$, $\alpha = 0.05$, $\beta_{\max} = 0.2$, $n_1 = 100$, $\alpha_0 = 0.7$, $\alpha_1 = 0.026$, $n_2 = 65$.

Random treatment effect

In our model the treatment-specific effect for the drug responders is a fixed value $\mu^T - \mu^C$. However, a more realistic model might consider it to be random. To address this issue, we present type II error rates for the chosen designs if the treatment-specific effect is a random variable from $\mathcal{N}(\mu^T - \mu^C, \delta^2)$. There, $Z^T \sim (1 - p)\mathcal{N}(\mu^C, \sigma^2) + p\mathcal{N}(\mu^T, \sigma^2 + \delta^2)$. In Table 4.5, average empirical type II error rates for the fixed and random treatment-specific effects are given. We observe bigger error rates when an effect is random. This might be taken into account in the development of the design.

		M_1			
		1	2	3	4
m	1	0.294 0.321	0.498 0.519	0.523 0.571	0.185 0.355
	2		0.099 0.111	0.179 0.22	0.079 0.193
	3			0.032 0.038	0.032 0.068
	4				0.007 0.009

(a) For the centres with strong effect
 $(\mu^*, p^*) = (2, 0.2)$.

		M_1			
		1	2	3	4
m	1	0.041 0.150	0.044 0.279	0.043 0.353	0.005 0.253
	2		0 0.038	0 0.075	0 0.072
	3			0 0.008	0 0.012
	4				0 0

(b) For the centres with strong effect
 $(\mu^*, p^*) = (1.2, 0.5)$.

Table 4.5 – Empirical type II error rates based on 1000 simulations. $Z^C \sim \mathcal{N}(0, 1)$, $Z^T \sim (1 - p^*)\mathcal{N}(0, 1) + p^*\mathcal{N}(\mu^*, 1)$ for fixed effect (top) and $Z^T \sim (1 - p^*)\mathcal{N}(0, 1) + p^*\mathcal{N}(\mu^*, 1 + 0.5^2)$ for random effect (bottom) effects. The parameters are $M = 4$, $\alpha = 0.05$, $\beta_{\max} = 0.2$, $n_1 = 100$, $\alpha_0 = 0.7$, $\alpha_1 = 0.026$, $n_2 = 65$.

Estimation of μ^T , μ^C , σ and p

For the classical sequential RCT design (no mixture response) the estimation of the parameters is discussed in Armitage et al. [1969], Siegmund [1978], O'Brien and Fleming [1979], and Tsiatis et al. [1984]. In the mixture framework, to characterize the sensitive subgroup, one would like to have estimators of μ^T , μ^C , σ and p . We do not further investigate this question, but there are some estimation procedures that can be applied to the data. A large literature covers various aspects of estimation of the mixture model parameters. In the case of a one-stage design, one can use the likelihood-based estimators of μ^T , μ^C , σ and p (e.g., the EM algorithm or the method proposed in Pavlic et al. [2001]). Regarding a two-stage trial, one can apply these algorithms to either the first or the second stage results.

4.6 The next step. Subgroup analysis

For a late-stage clinical trials, methods aimed at investigating the efficacy of the drug in pre-specified subgroups are referred to as a subgroup analysis. This rapidly growing area of research is of particular interest, for example, in oncology where tumor heterogeneity might not allow all patients to benefit from the molecularly targeted drugs. Given a large set of biomarkers, patient stratification is routinely done, and the testing of a new treatment is conducted given some information about the potential responders. One discriminates between a *confirmatory* and *exploratory* subgroup analysis. The former is used when the subgroups have already been identified and the goal is to build the design for the confirmatory trial controlling the type I error rate with regard to multiple hypotheses testing (for the details we address the reader to Li and Chan [2006], Brannath et al. [2009], Millen et al. [2012], Stallard et al. [2014], Wang and Hung [2014], and Kitcher and Hothorn [2014]). The latter focuses either on the discovery of the most promising subgroups or on the additional evaluation of the hypothesized subgroup and estimation of its treatment effect. Usually such an analysis is performed in a post-hoc manner employing machine learning algorithms. Examples include single gene logistic regression to extract markers with significant gene-treatment interactions (Freidlin et al. [2010]); voting methods (Breiman [1996]); regression trees (Su et al. [2009], Lipkovich and Dmitrienko [2014]) and penalized regressions (Gunter et al. [2011]).

However, adaptive schemes recently have been proposed where subgroup identification is made during the course of the trial, e.g., in Simon and Simon [2013], Xu et al. [2016], Gu et al. [2016], and Antoniou et al. [2016]. These represent two-stage adaptive designs, where in the first stage one classifies responders and non-responders, and the following enrollment is done only for patients that are in the "responders" subgroup. Given a new patient's results, one may increase the prediction accuracy and recalculate the model for classification. Targeted designs with patient selection are obviously more effective than sampling from the whole population. They enable a smaller sample size to be used for demonstrating the drug's efficacy and terminate the trial earlier, see Maitournam and Simon [2005].

4.7 Conclusion

In this chapter we presented an extension of the classical continuous response RCT design to the case when only a fraction of the treated patients responds to the tested treatment. The response in the treatment group is modeled as a two-component mixture, representing placebo and drug responders. We modified the conventional procedure that decides on the existence of a treatment-specific effect to a test that decides on the existence of a drug-responders subgroup. We characterize the latter by the fraction $p > 0$ and the treatment-specific effect $\mu > 0$. We assume that pairs (μ, p) of potential interest belong to the set \mathcal{E} , the region of strong effect. The trial protocol is then determined by the testing procedure that ensures a certain power of detection for any subgroup whose parameter is in \mathcal{E} . We developed one- and two-stage RCT designs along with exact and approximate expressions for the type II error rates. We also generalized these designs to the multicenter framework where we control the family-wise error rate. To decrease the sample size, we use Hochberg's step-up multiple test. If one wants to control FDR, one may simply use Benjamini–Hochberg's test or any other suitable thresholding of the p-values. We provide graphical aid to guide the user in choosing design parameters, where a few are determined by minimizing the required effort.

5 Classification procedures for real-data applications

5.1 Example 1: Classification models for gene expression data in cancer patients

5.1.1 Background and objectives

One of the most common applied medical research problems nowadays involves patient classification based on biomarker data. The classification may be done between disease stages, treatment outcomes, etc. In this chapter we present classification strategies for patients at risk for colon cancer. This type of cancer is the third most common after breast and lung cancers. Starting as adenoma polyps (benign tumors), the tumor's development into a malign form usually takes a long time with the first symptoms appearing at late stages. Early detection of the disease significantly increases 5-year survival rates, from 12% for stage IV to 35% for stage III, 71% for stage II and up to 90% for stage I (see National Cancer Institute (NCI). SEER 18 2008-2014). For years routine screening was done by colonoscopy, whose cost and side effects make patients less prone to agree to the procedure. However, recent studies, e.g., Ward et al. [2006], Marisa et al. [2013], Sideris and Papagrigoriadis [2014], Ciarloni et al. [2015], showed that gene expression measurements from blood cells can potentially be used for screening. If the predictive power is good enough, affordable, minimally-invasive diagnostics might become available in the future.

In the following sections we will compare the performance of linear and non-linear gene expression-based classifiers based on the subsets of biomarkers reported in Ciarloni et al. [2015]. We demonstrate that the projection pursuit method proposed in Anderson and Bahadur [1962], which we will refer to as *the most powerful linear projection method* (MPLP), has comparable results to those of the logistic regression and the support vector machine. Surprisingly, the MPLP is not often mentioned and used in applications though its simple implementation and the absence of tuning parameters make it worth considering for multivariate problems even if the model assumptions do not hold for the observed data.

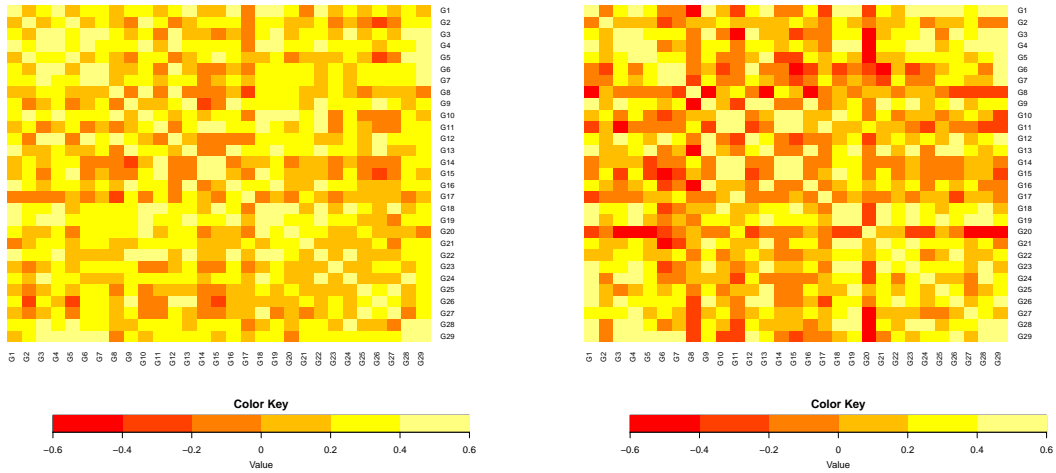


Figure 5.1 – The Pearson correlation matrices of 29 biomarkers for control (figure on the left) and cancer patients (figure on the right).

5.1.2 Data and Methodology

Data

The training dataset (TS), Fig. 5.1, consists of 50 control and 30 cancer patients measurements $\{(y_1, \mathbf{x}_1), \dots, (y_n, \mathbf{x}_n)\}$, where $\mathbf{y}_i = \{0, 1\}$, $i = 1, \dots, n$ is the vector of labels and $X = (\mathbf{x}_1, \dots, \mathbf{x}_n) \in \mathbb{R}^{n \times K}$ is the matrix of biomarkers. For each individual there are $K = 29$ biomarkers values that have been pre-selected by the Wilcoxon rank test and validated through penalized logistic regression as significant in discriminating between the patient populations. Three additional datasets, validation (VS), test (TestS) and extension (ES) sets (see Table 5.1) are used to validate the models.

	Training set	Validation set	Test set	Extension set
# of controls	50	25	48	26
# of cases	30	15	39	41

Table 5.1 – The datasets for training and validation.

Methodology

Feature selection

- **Robust PCA.** The biomarker correlation structure (see Fig. 5.1) suggests that the feature selection might be done based on the principal components. Denote the eigenvectors and the corresponding eigenvalues of the covariance matrix as $(\mathbf{a}_k, \lambda_k)$, $k = 1, \dots, 29$. The total variance of the data projected onto the space spanned by the principal components is $\sum_{i=1}^K \lambda_i$. In order to reduce the initial set of biomarkers, we search for variables in the original space that contribute to the variance along the i -th principal component more

5.1. Example 1: Classification models for gene expression data in cancer patients

than an average value of $\frac{1}{K}\lambda_i$. The heuristic selection rule is then

$$\mathbb{1}\left(K \frac{\sum_{i=1}^K w_{ij}^2 \lambda_i}{\sum_{i=1}^K \lambda_i} \geq 1\right), \quad (5.1)$$

where w_{ij} is the weight of the j -th variable of the original space in the vector of loadings \mathbf{a}_i .

Outliers can strongly affect the results of PCA. Thus, it is safer to consider the robust PCA, which is a modification of the standard PCA procedure based on a robust measure of variance, the Median Absolute Deviation,

$$\text{MAD}(x_1, \dots, x_n) = 1.48 \cdot \text{med}_j |x_j - \text{med}_i x_i|. \quad (5.2)$$

Finding the principal components then becomes a nonlinear eigenvector problem, with k -th eigenvector

$$\mathbf{a}_k = \underset{\|\mathbf{a}\|=1}{\text{argmax}} \{\text{MAD}(\hat{X}_k \mathbf{a})\},$$

where $\hat{X}_k = X - \sum_{s=1}^{k-1} X \mathbf{a}_s \mathbf{a}_s^T$.

- **Data & biologically driven selection.** Here, feature selection was guided by possible disease-driven changes. First, the standardized changes in the biomarker mean values, second, the loss of correlations (positive or negative) that have been observed among healthy individuals but changed in the case of the disease.
- **Aggregated set of biomarkers.** These are the significant predictors of the penalized logistic regression, the main model used in the production for clinical patients.

	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11	G12	G13	G14
robust PCA		X	X		X	X	X	X	X		X	X	X	X
data & biology		X	X		X	X			X				X	X
aggregated	X		X			X			X	X		X	X	X

Table 5.2 – Reduced sets of biomarkers

Remark: If any of methods for feature selection or for classification required the estimation of the covariance matrix Σ , we used a shrinkage covariance estimator proposed in Schäfer and Strimmer [2005], $\Sigma^* = \lambda T + (1 - \lambda)U$, where T is the *target* covariance. In genetic applications it is usually considered to be diagonal with unequal diagonal elements, and U is the MLE of the covariance matrix. The shrinkage parameter in this case is $\hat{\lambda} = \frac{\sum_{i \neq j} \widehat{\text{Var}}(s_{ij})}{\sum_{i \neq j} s_{ij}^2}$, where $s_{ij} = \frac{1}{n-1} \sum_{k=1}^n (x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j)$.

Classification algorithms

- **Penalized logistic regression (PLR).** The log-odds ratio of class memberships is modeled as a linear function of biomarkers,

$$\log\left(\frac{\mathbb{P}(y = 1|\mathbf{x})}{\mathbb{P}(y = 0|\mathbf{x})}\right) = \beta_0 + \beta^T \mathbf{x}.$$

To fit the model one minimizes the penalized negative binomial likelihood,

$$-l(\beta_0, \beta|(y_1, x_1), \dots, (y_n, x_n)) + \lambda\left(\frac{1-\alpha}{2}\|\beta\|_2 + \alpha\|\beta\|_1\right).$$

The tuning parameters λ and α are determined by cross-validation and bootstrap.

- **Linear projection.** When two classes are modeled by multivariate normal distributions $\sim \mathcal{N}(\boldsymbol{\mu}_i, \Sigma)$ with prior probabilities π_i , $i = 0, 1$, applying Bayes' theorem shows that the *a posteriori* log odds ratio $\log\left(\frac{\mathbb{P}(y=1|\mathbf{x})}{\mathbb{P}(y=0|\mathbf{x})}\right)$ equals Fisher's linear discriminant function $\lambda(\mathbf{x}) = \beta_0 + \beta^T \mathbf{x}$, where

$$\begin{aligned}\beta_0 &= \log\frac{\pi_1}{\pi_0} - \frac{1}{2}(\boldsymbol{\mu}_1^T \Sigma^{-1} \boldsymbol{\mu}_1 - \boldsymbol{\mu}_0^T \Sigma^{-1} \boldsymbol{\mu}_0), \\ \beta^T &= (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)^T \Sigma^{-1}.\end{aligned}$$

Therefore, the LR and the Fisher's LDA are closely linked. In the given clinical data, neither multivariate normality nor equality of the covariance matrices holds. Nevertheless, it has been claimed (Hastie et al. [2009], Hosmer Jr et al. [2013], McLachlan [2004]) that if the assumptions of the LDA are not severely violated, it still might be successfully applied. To eliminate the assumption of equal covariance matrices, we decided to use the method proposed in Anderson and Bahadur [1962], which is a projection pursuit method searching for the most powerful discriminant direction under the parametric assumptions of Fisher's LDA except for the condition on the equality of the covariance matrices. Surprisingly, we have not seen many examples of its application and it is not listed among the most useful machine learning techniques, though it is an excellent example of an intuitive and efficient method with low computational burden.

Let the class models be $\mathcal{N}(\boldsymbol{\mu}_0, \Sigma_0)$ and $\mathcal{N}(\boldsymbol{\mu}_1, \Sigma_1)$, respectively. For the predefined type I error rate $\alpha < 0.5$, one searches the hyperplane $\mathbf{c}^T \mathbf{x} = b$ such that $\mathbb{P}(\mathbf{c}^T \mathbf{x} > b | \mathcal{N}(\boldsymbol{\mu}_0, \Sigma_0)) = \alpha$ and $\beta(\mathbf{c}, b) = \mathbb{P}(\mathbf{c}^T \mathbf{x} < b | \mathcal{N}(\boldsymbol{\mu}_1, \Sigma_1))$ is minimized. Having the estimates of models' parameters, we shift the first group mean to zero and successively rotate the data with the eigenvectors of Σ_0 and then with the eigenvectors of Σ_1 . The new parameters are $\widetilde{\boldsymbol{\mu}}_0 = \mathbf{0}$, $\widetilde{\boldsymbol{\mu}}_1 = (\mu_1, \dots, \mu_K)$, $\mu_i \geq 0$, $\widetilde{\Sigma}_0 = I$, $\widetilde{\Sigma}_1 = \text{diag}(\sigma_1^2, \dots, \sigma_K^2)$, $i = 1, \dots, K$. Denote also $\widetilde{\mathbf{c}} = \mathbf{e} = (e_1, \dots, e_K)$ and $\widetilde{b} = a$. Without loss of generality consider $\sum_{i=1}^K e_i^2 = 1$. Denote the mean and the standard deviation of the projected data as $\mu = \sum_i \mu_i e_i$ and $\sigma = \sqrt{\sum_i \sigma_i^2 e_i^2}$, respectively, put $a = \Phi^{-1}(1 - \alpha) = z_{1-\alpha} > 0$. The optimization problem is

5.1. Example 1: Classification models for gene expression data in cancer patients

then reformulated as

$$G(\mathbf{e}) = \frac{\sum_i \mu_i e_i - a}{\sqrt{\sum_i \sigma_i^2 e_i^2}} \longrightarrow \max \quad \text{given that} \quad \sum_{i=1}^K e_i^2 = 1. \quad (5.3)$$

Let $\mathbf{e}^* = \{e_1^*, e_2^*, \dots, e_K^*\}$ be the point that solves (5.3). Denote $\mu^* = \sum_i \mu_i e_i^*$, $\sigma^* = \sqrt{\sum_i \sigma_i^2 e_i^{*2}}$ and the corresponding type II error rate $\beta^* = \Phi\left(\frac{a - \mu^*}{\sigma^*}\right)$. Then the following holds

$$\beta^* < 0.5 \text{ if and only if } \sum_i \mu_i^2 > a^2,$$

$$\beta^* = 0.5 \text{ if and only if } \sum_i \mu_i^2 = a^2,$$

$$\beta^* > 0.5 \text{ if and only if } \sum_i \mu_i^2 < a^2.$$

An algorithm to find \mathbf{e}^*

We will focus now on the case $\sum_i \mu_i^2 > a^2$, as it is the one of practical interest.

Lemma 5.1. *The components of the vector \mathbf{e}^* satisfy*

$$e_i^* = \frac{\mu_i}{a + \frac{\mu^* - a}{\sigma^{*2}} \sigma_i^2}, \quad i = 1, \dots, K. \quad (5.4)$$

Proof.

The point \mathbf{e}^* should satisfy the condition $\nabla G(\mathbf{e}^*) \sim \mathbf{e}^*$, i.e.,

$$\exists \lambda > 0 : \frac{\mu_i}{\sigma^*} - \frac{\mu^* - a}{\sigma^{*2}} \frac{\sigma_i^2 e_i^*}{\sigma^*} = \lambda e_i^*, \quad i = 1, \dots, K.$$

Multiplying the equation above by e_i^* and summing for all i , we get

$$\frac{\mu^*}{\sigma^*} - \frac{\mu^* - a}{\sigma^{*2}} \sigma^* = \lambda \implies \lambda = \frac{a}{\sigma^*}.$$

Therefore, $\mu_i = e_i^* \left(\frac{(\mu^* - a) \sigma_i^2}{\sigma^{*2}} + a \right)$. Recall that for $\sum_i \mu_i^2 > a^2$, one has $G(\mathbf{e}^*) = \frac{\mu^* - a}{\sigma^*} > 0$.

Consequently, $\frac{(\mu^* - a) \sigma_i^2}{\sigma^{*2}} > 0$ and $e_i^* = \frac{\mu_i}{a + \frac{\mu^* - a}{\sigma^{*2}} \sigma_i^2}$. \square

Note that if $\sum_i \mu_i^2 < a^2$, it is possible that for some i : $a + \frac{(\mu^* - a) \sigma_i^2}{\sigma^{*2}} = 0$, and the formula (5.4) is not valid.

The vector \mathbf{e}^* can be found numerically with Newton's iterative procedure. First we prove that there is a unique \mathbf{e} : $e_i = \frac{\mu_i}{k \sigma_i^2 + a}$ for some $k > 0$. Indeed, the function $f(k) = \sum_i \left(\frac{\mu_i}{k \sigma_i^2 + a} \right)^2 - 1$ is decreasing on $[0, +\infty)$. Since $f(0) = \sum_i \left(\frac{\mu_i}{a} \right)^2 - 1 > 0$ and $f(k) \xrightarrow{k \rightarrow \infty} -1$,

there exists a unique $k > 0$: $f(k) = 0$, which gives a unique vector \mathbf{e} with the components $e_i = \frac{\mu_i}{k\sigma_i^2 + a}$, $i = 1, \dots, K$.

For the point \mathbf{e}^* where the global maximum of $G(\mathbf{e})$ is attained, $e_i^* = \frac{\mu_i}{k^*\sigma_i^2 + a}$ and $k^* = \frac{\mu^* - a}{\sigma^{*2}} > 0$. Therefore, we conclude that \mathbf{e}^* is obtained from the unique solution of the equation $f(k^*) = 0$.

Notice that $f'(k) < 0$ and $f''(k) > 0$. Therefore $f(k)$ is decreasing and convex. Starting from some k_0 : $f(k_0) > 0$ (for example, $k_0 = 0$), the sequence

$$k_m = k_{m-1} - \frac{f(k_{m-1})}{f'(k_{m-1})}$$

converges to the solution k^* .

- **Support Vector Machines.** Finally, we add to the analysis one of the most widely used non-parametric generalized linear classifiers, support vector machines (SVM). For the purposes of the method, the class of an observation is $y_i = \pm 1$. In the case of perfect separability, SVM find a hyperplane $\mathbf{w}^T \mathbf{x} - b = 0$ such that $y_i[\mathbf{w}^T \mathbf{x}_i - b] \geq 1$ for all $i = 1, \dots, n$ and the margin $\frac{2}{\|\mathbf{w}\|}$ between $\mathbf{w}^T \mathbf{x} - b = -1$ and $\mathbf{w}^T \mathbf{x} - b = 1$, the hyperplanes that completely separate two classes, is maximized. Perfect separability is, however, an unrealistic assumption and the optimization problem is modified with the hinge loss function $\max(0, 1 - y_i(\mathbf{w}^T \mathbf{x}_i - b))$ as the following

$$\min_{\mathbf{w}, \xi} \left(\frac{1}{2} \|\mathbf{w}\|^2 + \frac{C}{n} \sum_{i=1}^n \xi_i \right) \text{ such that } y_i(\mathbf{w}^T \mathbf{x}_i - b) \geq 1 - \xi_i \quad \forall i = 1, \dots, n. \quad (5.5)$$

Here, $\xi_i \geq 0$ is a slack variable indicating the location of the i -th observation with regard to the margin:

- If $\xi_i = 0$, then the i -th observation is on the correct side of the margin;
- If $0 < \xi_i \leq 1$, then the i -th observation is inside the margin, but on the correct side of the hyperplane;
- If $\xi_i > 1$, then the i -th observation is on the wrong side of the hyperplane.

C is the penalty for violating the margins.

The solution for (5.5) is of the type

$$\mathbf{w} = \sum_{i=1}^n \alpha_i y_i \mathbf{x}_i \quad (5.6)$$

with $\alpha_i > 0$ only for the points (they are called the support vectors) inside the margin or on the wrong side of the discriminating hyperplanes. A decision about a new observation \mathbf{x}' is made based on its inner products with the support vectors. This property enables one to generalize the optimization to the nonlinear cases using kernels. In our analysis we took linear and radial kernels,

5.1. Example 1: Classification models for gene expression data in cancer patients

- Linear: $k(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i^T \mathbf{x}_j$;
- Radial: $k(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2)$, for $\gamma > 0$.

The SVM decision function is determined by $\text{sign}((\sum_{i=1}^n \alpha_i y_i k(\mathbf{x}_i, \mathbf{x}')) + b)$.

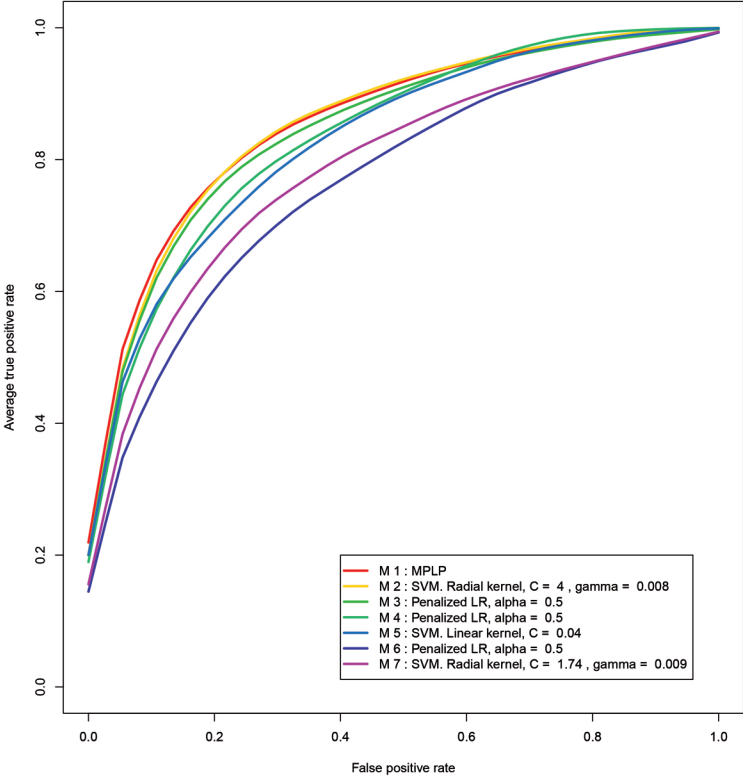
5.1.3 Results and conclusions

The best models were chosen based on 500 bootstraps from the Training Set. In Tables 5.3 and 5.4, we list the top seven models along with:

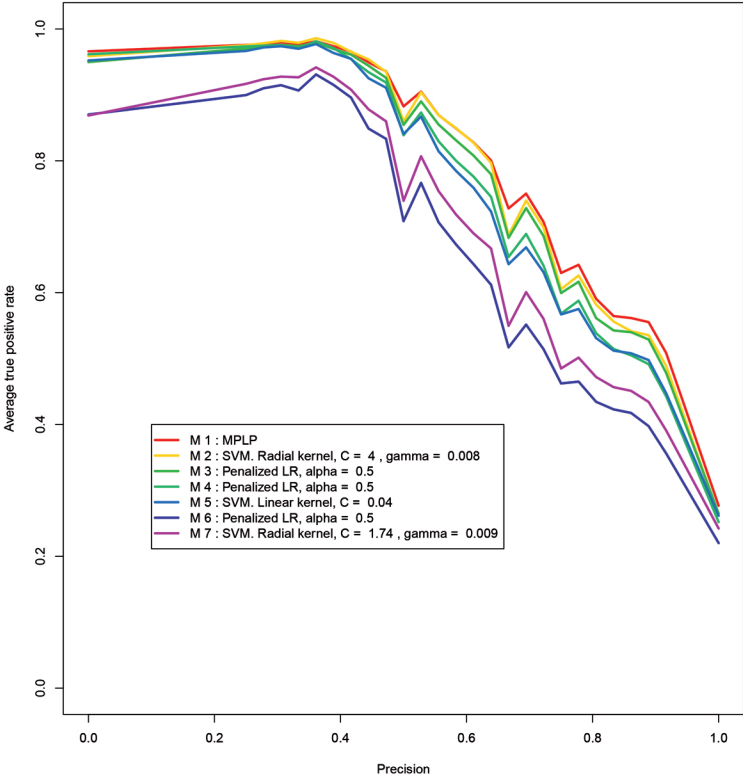
- Tuning parameters: the penalization parameter **cost**(C) for SVM or α for the logistic regression (for fixed α , λ is chosen based on the 5-fold cross validation); the parameter of the radial kernel in SVM, γ ;
- **Sensitivity** (Se) and **specificity** (Sp) of a method for a given dataset and the average value for bootstraps;
- The average area under ROC curve (**AUC**) on bootstraps;
- The L_1 measure of a model's stability, **Diff**= |Sensitivity TS–Sensitivity Boot|+|Specificity TS–Specificity Boot|.

The ROC and the Precision-Recall curves for the corresponding models are given in Fig. 5.2.

Dimensionality reduction significantly improves the performance of any of the trained classifiers. The biomarkers selected based on biological evidence and those extracted from the data given some strong external assumptions are reasonably good as predictors (the first four models in Table 5.3). The sets of biomarkers chosen solely by the data-dependent methods (here we mention the robust PCA, the only method with a satisfactory performance) are mostly non-informative. Except for the Validation Set where both sensitivity and specificity drop for all models (which is, probably, because of the small sample size), the results are in agreement with the bootstraps of the Training Set, which represent the 'average' data. The linear projection method and SVM with linear kernel showed the best overall results. Note that the difference between the TS bootstraps and Test or Extension Sets sensitivity and specificity is less for the linear projection method, which means that for the given data, it is more robust. In conclusion we note that with the reduced dimensionality the performance of the linear classifiers can be comparable with the non-linear methods. In particular, the projection pursuit method, an extension of Fisher's LDA for unequal covariance structures, might be a good alternative before applying more complex classifiers. With a careful feature selection, it gives tractable results with no additional tuning parameters needed.



(a) ROC curves for the classification models.



(b) Precision-Recall curves for the classification models.

Figure 5.2 – The classification characteristics for the chosen models

5.1. Example 1: Classification models for gene expression data in cancer patients

Table 5.3 – Models trained for discriminating between control versus cancer patients. Performance on the Training Set. The decision thresholds are chosen to reach 80% sensitivity on the bootstraps of the training set.

	# of biomarkers	Model	C/α	γ	Se TS	Se Boot.	Sp TS	Sp Boot.	AUC	L_1
1	7	MPLP			0.80	0.80	0.88	0.77	0.87	0.11
2	7	SVM. Radial kernel	4	0.01	0.80	0.80	0.86	0.77	0.86	0.09
3	7	PLR	0.5		0.90	0.80	0.78	0.74	0.85	0.14
4	8	PLR	0.5		0.87	0.80	0.72	0.69	0.84	0.10
5	11	SVM. Linear kernel	0.04		0.83	0.79	0.76	0.68	0.84	0.12
6	29	PLR	0.5		1	0.80	0.48	0.53	0.77	0.26
7	29	SVM. Radial kernel	1.74	0.01	0.90	0.80	0.82	0.58	0.79	0.34

Table 5.4 – Results on the Validation, Test and Extension sets.

Models	Validation Set				Test Set				Test Set + Extension set						
	Se CRC	Sp	Se.POL	Se CRC I.II	Se CRC I.IV	Se CRC	Sp	Se POL	Se CRC I.II	Se CRC I.IV	Se CRC	Sp	Se POL	Se CRC I.II	Se CRC I.IV
1	0.73	0.72	0.43	0.63	0.86	0.76	0.77	0.46	0.79	0.73	0.78	0.80	0.50	0.74	0.82
2	0.73	0.76	0.52	0.62	0.86	0.79	0.73	0.44	0.86	0.73	0.82	0.74	0.45	0.83	0.82
3	0.80	0.68	0.48	0.75	0.86	0.79	0.73	0.49	0.86	0.73	0.80	0.74	0.52	0.78	0.82
4	0.80	0.68	0.57	0.75	0.86	0.86	0.62	0.54	0.86	0.87	0.86	0.65	0.55	0.83	0.89
5	0.80	0.72	0.48	0.75	0.86	0.83	0.71	0.49	0.93	0.73	0.84	0.72	0.50	0.87	0.82
6	0.87	0.40	0.71	0.88	0.86	0.90	0.35	0.67	0.93	0.87	0.90	0.34	0.69	0.91	0.89
7	0.80	0.60	0.52	0.75	0.86	0.86	0.56	0.64	0.86	0.87	0.86	0.57	0.62	0.83	0.89

5.2 Example 2: Linking non-annotated genes to existing gene sets

Genes involved in the same biological or biochemical pathway can be grouped into gene sets. These sets can then be used to make sense of microarray or RNASeq data by linking genes to a function. We describe a method for assigning non-annotated genes to existing gene sets (GS) based on the statistical similarity.

5.2.1 Background and objectives

Instead of analyzing differences in the expression level of a single gene between given conditions, gene set analysis focuses on a collection of genes that share a common biological function, e.g., participating in a certain biological pathway. The conclusion about the strength of an association between gene set and experimental target is usually based on the ranks of some relevant to the problem statistic values corresponding to genes in a query set. A similar approach is applied to classifying non-annotated genes. It is believed that equally co-expressed genes share a similar biological functions. Thus, to infer potential membership of a gene, one explores its co-expression with known pathways, e.g., Subramanian et al. [2005], Adler et al. [2009], Baughman et al. [2009], Zhu et al. [2015], Li et al. [2017].

One of the major problems in this type of analysis is inter-gene correlation in a gene set which inflates the type I error rate of the statistical tests (if they ignore the correlation). Many solutions to this problem have been proposed, e.g., Wang et al. [2008, 2009], Nam [2010], Wu and Smyth [2012], Yaari et al. [2013]. In particular, there are three main strategies. If in the study one measures gene activities in two or more groups, e.g., phenotypic categories, control/treatment, etc., the correction for the inter-gene dependence can be handled using the variance inflation factor (VIF). After the correction, the univariate statistics reflecting gene set associations are tested for significance. If there is only one available dataset, which is our case, one can either decorrelate genes in the gene set or permute sample labels and estimate the p-values of the resulted statistics based on the empirical null distribution. The decorrelation requires accurate estimate of the covariance matrix, which is computationally expensive in small size/high dimension problems. The permutation approach, on the other hand, besides carrying the same computational burden, was criticized for changing the hypothesis being tested, see Irizarry et al. [2009], Tamayo et al. [2016].

We propose a simple and comprehensible method that picks a non-annotated gene as a candidate for a given gene set based on its correlation with a so-called "gene set representative". This enables us to avoid correcting for the inter-correlation in the gene set while keeping necessary information about the gene similarities.

5.2.2 Data and Methods

The numbered set of all genes is $G = \{1, 2, \dots, k\}$ and a gene set of size m is a subset $GS = \{g_1, g_2, \dots, g_m\} \subset G$. We want to construct a method for deciding whether a target gene $t \in G$ ($t \notin GS$) should be considered linked to or associated with the gene set GS . To decide this, we dispose of data measuring the activity of all genes in various organs. Our method is based on the pairwise correlation observed in the data, which for a single data set is defined as

$$r(g_1, g_2) = \frac{\sum_{i=1}^n (x_{i,g_1} - \bar{x}_{g_1})(x_{i,g_2} - \bar{x}_{g_2})}{\sqrt{\sum_{i=1}^n (x_{i,g_1} - \bar{x}_{g_1})^2 \sum_{i=1}^n (x_{i,g_2} - \bar{x}_{g_2})^2}}. \quad (5.7)$$

In this formula, $(x_{1,g}, x_{2,g}, \dots, x_{n,g})$ are the n observed activities of gene g measured in comparable instances and \bar{x}_g is the observed average activity of gene g .

To represent the gene set $GS = \{g_1, \dots, g_m\}$, we construct an artificial gene with observed activities (s_1, \dots, s_n) representing the set. We choose the values of s_i to maximize the average correlation or equivalently the sum of the correlations $\sum_{g \in GS} r(g, s)$ with all genes in the gene set. This leads to

$$s_i = \frac{1}{m} \sum_{g \in GS} \frac{x_{i,g} - \bar{x}_g}{\sqrt{\sum_{j=1}^n (x_{j,g} - \bar{x}_g)^2}}, \quad i = 1, \dots, n,$$

that is, the average of activities of the genes in the gene set. If we correlate a non-annotated target gene with the representor, we obtain a correlation which is proportional to the average correlation of the target gene with the members of the gene set. The proportionality constant is equal to one over the square root of the average entry in the inter-gene set correlation matrix. In particular, for an equi-correlated structure with $r(g_1, g_2) = r_0 > 0$ for all pairs of (different) genes $g_1, g_2 \in GS$, then the correlation between the representor and any one of the genes in the gene set can be shown to be equal to

$$\sqrt{r_0 + \frac{1}{m}(1 - r_0)} > \sqrt{r_0} > r_0. \quad (5.8)$$

Even when the average gene-gene correlation in a GS is small, the average correlation between representor and the genes in the GS is high, which makes it possible to use the correlation with the representor as an indicator. An example is shown in Fig. 5.3, where the observed correlation matrix of the reactome respiratory electron transport gene set is depicted. On average the correlation between the representor and the genes in this gene set is 0.4, while the average inter-gene set correlation is only 0.145.

Statistics

Given a dataset of measured gene activities, we scan across unannotated target genes outside a given gene set in order to test the null hypothesis that the target gene is not associated with the gene set. As a test statistic, we use the correlation $r(t, s)$, where t is a target gene and s is

5.2. Example 2: Linking non-annotated genes to existing gene sets

conclude that the target gene t is an associate of the gene set GS, if

$$D = \frac{1}{2} \ln \left(\frac{1 + r(t, s)}{1 - r(t, s)} \right) - \frac{1}{2} \ln \left(\frac{1 + \rho_{\text{base}}}{1 - \rho_{\text{base}}} \right) \geq \frac{1}{2} \ln \left(\frac{1 + r_{\beta}}{1 - r_{\beta}} \right) - \frac{1}{2} \ln \left(\frac{1 + \rho_{\text{base}}}{1 - \rho_{\text{base}}} \right). \quad (5.9)$$

This choice ensures that if it were applied to the genes in GS, it would detect a fraction of $1 - \beta$ of them.

If the null hypothesis is true, we approximately have

$$D \sim N \left(0, \frac{1}{n-3} \right),$$

where n is the number of samples in the data set. It follows that the rate of false positives corresponding to the cutoff r_{β} is

$$\alpha = 1 - \Phi \left(\sqrt{n-3} \left(\frac{1}{2} \ln \left(\frac{1 + r_{\beta}}{1 - r_{\beta}} \right) - \frac{1}{2} \ln \left(\frac{1 + \rho_{\text{base}}}{1 - \rho_{\text{base}}} \right) \right) \right),$$

where $\Phi(z)$ is the standard normal cumulative distribution function. We will use this quantity to estimate the expected number of false positive annotations by multiplying α with the number of target genes to be tested.

Combining across multiple data sets

By combining the findings from several independent data sets (in our case these are different tissues), we can increase the power of the test. To combine, we first use the procedure described above to compute the type I error α_j for data set $j = 1, \dots, J$. Then we scan through all the target genes, that is, the genes whose activities were observed and are non-annotated with the given gene set. Suppose there are N_j such potential targets. Next, we estimate the expected number of false positives and control its total value to be less than 1. Finally, we combine the evidence only over the $J' \leq J$ data sets with the smallest number of expected false positives $N_j \alpha_j$, where J' is the largest number of data sets which still satisfies

$$\sum_{j'=1}^{J'} \alpha_{j'} \cdot N_{j'} < 1. \quad (5.10)$$

In this formula, j' is the rank of a data set with regard to the expected false positives, that is, $j' = 1$ is the data set with the smallest value of $N\alpha$, while $j' = J$ is the data set with the largest such value.

The combined statistic for the target gene t employs the inverse variance weighted score, which is based on the test statistics computed separately for each data set $D_{j'}$ (5.9). The

combined score is

$$D_{\text{combined}} = \frac{\sum_{j'=1}^{J'} (n_{j'} - 3) D_{j'}}{\sum_{j'=1}^{J'} (n_{j'} - 3)}, \quad (5.11)$$

where $n_{j'}$ is the sample size in the corresponding sample. If the null hypothesis of no association is true we have approximately

$$D_{\text{combined}} \sim \mathcal{N}\left(0, \frac{1}{\sum_{j'=1}^{J'} (n_{j'} - 3)}\right),$$

and the combined p-value is

$$p_{\text{combined}} = 1 - \Phi\left(D_{\text{combined}} \sqrt{\sum_{j'=1}^{J'} (n_{j'} - 3)}\right).$$

We will, however, not make use of the above null distribution, because it would lead to a high number of genes associated with a gene set. Instead, we refer to the bound r_β , which we calculated for each data set. In the analysis based on the combined correlation, we reject the null hypothesis if

$$D_{\text{combined}} > \frac{\sum_{j'=1}^{J'} (n_{j'} - 3) (z(r_{j',\beta}) - z(\rho_{j',\text{base}}))}{\sum_{j'=1}^{J'} (n_{j'} - 3)}, \quad (5.12)$$

where $z(r) = \ln((1+r)/(1-r))/2$ is the Fisher transformation. The rejection boundary in (5.12) is one among many possible choices. The proposed one is consistent with (5.9) for one data set. In case of many data sets the rejection rule selects as candidates genes that are strongly correlated with a representor of a gene set in every tissue explored. Not all genes in the gene set pass this bound, since their expression levels may vary between the tissues. For the combined procedure the pooled type I error is

$$\alpha_{\text{pooled}} = 1 - \Phi\left(\frac{\sum_{j'=1}^{J'} (n_{j'} - 3) (z(r_{j',\beta}) - z(\rho_{j',\text{base}}))}{\sqrt{\sum_{j'=1}^{J'} (n_{j'} - 3)}}\right).$$

Based on corresponding p-values one can apply further multiple testing corrections. In the next chapter we compare two multiple testing procedures with (5.12).

5.2.3 Example. Reactome respiratory electron transport gene set enrichment analysis

As an example of the procedure, we take the reactome respiratory electron transport pathway. If $\beta = 0.2$ and $\rho_{\text{base}} = 0$, eight tissues out of 52 are left out of the analysis due to the number of false positives they bring. For consistency we scan only the genes that are present in all analyzed data sets. Having the set of *supposed* and *not-supposed* genes for this gene set, we

5.2. Example 2: Linking non-annotated genes to existing gene sets

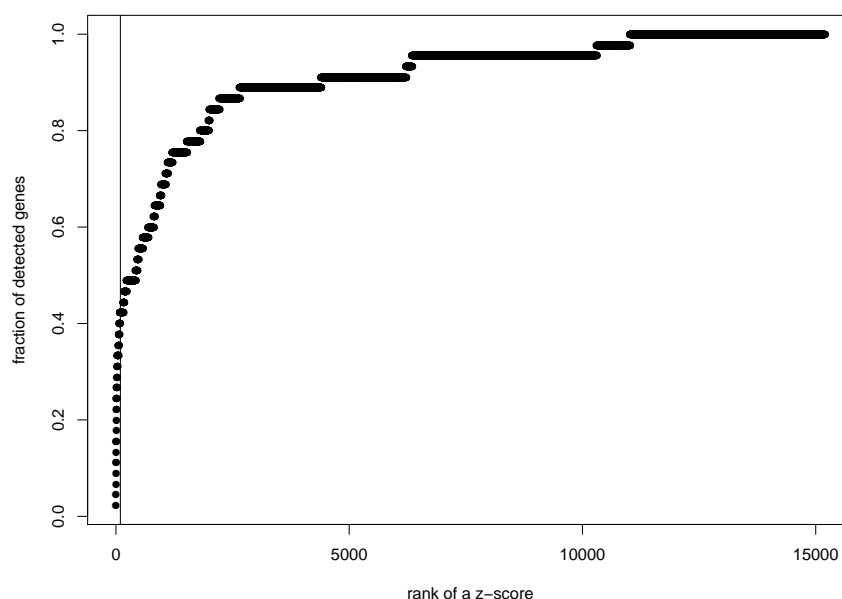


Figure 5.4 – The ROC curve for the rank based classifier of unannotated genes for the reactome respiratory electron transport gene set. The vertical line defines the significance cutoff. Here it is the 80 % quantile of correlations of genes in the set with its representor.

validate the procedure, plotting in Fig. 5.4 the chance of detecting *supposed* genes along the rank of the combined score from (5.11). The plot shows a good predictive performance with the top genes being among the *supposed* ones. In Tables 5.5, 5.6 results for different experimental setups are given. As is shown in Table 5.5, ρ_{base} does not have a marked effect on the number of discoveries, while it varies between 0 and 0.2. We think that this is a realistic range for baseline correlation, though one should decide about it based on the average correlations across the investigated gene sets. Table 5.6, on the contrary, shows that changing the power for the genes in the set directly modulates the number of discoveries. The value of α_{pooled} is so small that making inference based on procedures like the Bonferroni correction or the FDR leads to many more possible candidates.

ρ_{base}	# of discoveries	# of scanned genes	# of discoveries. Bonferroni	# of discoveries. FDR
0	106	34614	4170	4191
0.1	105	34644	1578	1579
0.2	105	34686	758	758

Table 5.5 – An example of the method for the reactome respiratory electron transport pathway. Thresholds are determined by $\beta = 0.2$

Chapter 5. Classification procedures for real-data applications

$1 - \beta$	# of discoveries	# of scanned genes	# of discoveries. Bonferroni	# of discoveries. FDR
0.7	42	34335	754	755
0.6	29	34309	754	754
0.5	14	34309	754	754

Table 5.6 – An example of the method for the reactome respiratory electron transport pathway. $\rho_{base} = 0.2$

5.2.4 Conclusion

The GSEA method proposed here is an alternative to existing approaches designed to correct for a gene-gene interactions. Our method combines the gene set expression values into a single gene, the set representative. We investigate the relatedness between the set representative and unannotated gene candidates and compare it to the gene set members. Depending on the gene set, we compute the cutoff value, which determines the extent of relatedness that is considered as an indicator of the membership in this gene set. This allows us to compute the type I error rate of the assignment using only the the log transformed correlations and the cutoff value. Hence, for the given gene set in a given dataset, the false discovery rate is computed in seconds without additional sampling. The explicit type I error rates for the given gene set could be employed when one needs to combine the evidence across multiple datasets. As a consequence, if in some dataset the investigated gene set is not coherent, meaning the absence of a strong linkage, the resulting type I error rate will be big and the dataset most likely will be excluded from the meta-analysis. As for the final inference, we claim that rather than considering p-values, which we also provide, the selection of new candidates should be based on the cutoff value which reveals the consistency of the scanned candidates with the gene set members. The advantage of this method is in its interpretability, i.e., rather than operating on the p-values as an argument for the gene membership, one can claim that the selected genes have been picked due to performing like the top $(1 - \beta)\%$ percent of the genes in the gene set.

One can criticize the method for being too conservative, as we require the selected genes to be in agreement with the most of the genes in the set in most of the datasets. However, changing the parameters β , ρ_{base} and the number of false positives in (5.10), will weaken the selection criteria and uncover many more new candidates.

Bibliography

- Priit Adler, Raivo Kolde, Meelis Kull, Aleksandr Tkachenko, Hedi Peterson, Jüri Reimand, and Jaak Vilo. Mining for coexpression across hundreds of datasets using novel rank aggregation and visualization methods. *Genome Biology*, 10(12):R139, Dec 2009. ISSN 1474-760X.
- Ofer Agid, Cynthia O Siu, Steven G Potkin, Shitij Kapur, Eric Watsky, Douglas Vanderburg, Robert B Zipursky, and Gary Remington. Meta-regression analysis of placebo response in antipsychotic trials, 1970–2010. *American Journal of Psychiatry*, 170(11):1335–1344, 2013.
- Theodore W Anderson and Raghu Raj Bahadur. Classification into two multivariate normal distributions with different covariance matrices. *The Annals of Mathematical Statistics*, 33(2):420–431, 1962.
- Joshua D Angrist, Guido W Imbens, and Donald B Rubin. Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, 91(434):444–455, 1996.
- Miranta Antoniou, Andrea L Jorgensen, and Ruwanthi Kolamunnage-Dona. Biomarker-guided adaptive trial designs in phase II and phase III: a methodological review. *PLoS One*, 11(2): e0149803, 2016.
- Peter Armitage, CK McPherson, and BC Rowe. Repeated significance tests on accumulating data. *Journal of the Royal Statistical Society. Series A (General)*, pages 235–244, 1969.
- Joshua M. Baughman, Roland Nilsson, Vishal M. Gohil, Daniel H. Arlow, Zareen Gauhar, and Vamsi K. Mootha. A computational screen for regulators of oxidative phosphorylation implicates slirp in mitochondrial rna homeostasis. *PLOS Genetics*, 5(8):1–12, 08 2009.
- Beecher. The powerful placebo. *Journal of the American Medical Association*, 159(17):1602–1606, 12 1955.
- Yoav Benjamini and Yosef Hochberg. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 289–300, 1995.
- Andrew C Berry. The accuracy of the Gaussian approximation to the sum of independent variates. *Transactions of the American Mathematical Society*, 49(1):122–136, 1941.

Bibliography

- Katja Boehm, Bettina Berger, Ulrich Weger, and Peter Heusser. Does the model of additive effect in placebo research still hold true? a narrative review. *JRSM open*, 8(3): 2054270416681434, 2017.
- Remy Boussageon, Francois Gueyffier, Theodora Bejan-Angoulvant, and Geraldine Felden-Dominiak. Critical of the additive model of the randomized controlled trial. *Therapie*, 63(1): 29–35, 2008.
- Werner Brannath, Emmanuel Zuber, Michael Branson, Frank Bretz, Paul Gallo, Martin Posch, and Amy Racine-Poon. Confirmatory adaptive designs with Bayesian decision tools for a targeted therapy in oncology. *Statistics in Medicine*, 28(10):1445–1463, 2009.
- Leo Breiman. Bagging predictors. *Machine Learning*, 24(2):123–140, 1996.
- T. Tony Cai, Jiashun Jin, and Mark G. Low. Estimation and confidence sets for sparse normal mixtures. *Ann. Statist.*, 35(6):2421–2449, 2007.
- D Chant. On asymptotic tests of composite hypotheses in nonstandard conditions. *Biometrika*, 61(2):291–298, 1974.
- Herman Chernoff. A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *The Annals of Mathematical Statistics*, 23(4):493–507, 1952.
- Herman Chernoff. On the distribution of the likelihood ratio. *The Annals of Mathematical Statistics*, pages 573–578, 1954.
- Laura Ciarloni, Sahar Hosseinian, Sylvain Monnier-Benoit, Natsuko Imaizumi, Gian Dorta, Curzio Ruegg, DGNP-COL-0310 Study Group, et al. Discovery of a 29-gene panel in peripheral blood mononuclear cells for the detection of colorectal cancer and adenomas using high throughput real-time pcr. *PLoS One*, 10(4):e0123904, 2015.
- David Donoho and Jiashun Jin. Higher criticism for detecting sparse heterogeneous mixtures. *The Annals of Statistics*, 32(3):962–994, 2004.
- David Donoho and Jiashun Jin. Higher criticism thresholding: Optimal feature selection when useful features are rare and weak. *Proceedings of the National Academy of Sciences*, 105(39): 14790–14795, 2008.
- David Donoho and Jiashun Jin. Feature selection by higher criticism thresholding achieves the optimal phase diagram. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 367(1906):4449–4470, 2009.
- David Donoho, Jiashun Jin, et al. Higher criticism for large-scale inference, especially for rare and weak effects. *Statistical Science*, 30(1):1–25, 2015.
- Carl-Gustaf Essen. On the liapounoff limit of error in the theory of probability. *Arkiv för Matematik, Astronomi Och Fysik*, 1942(A28):1–19, 1942.

- Barbara J Evans, David A Flockhart, and Eric M Meslin. Creating incentives for genomic research to improve targeting of therapies. *Nat Med*, 10(12):1289–1291, 12 2004.
- Constantine E Frangakis and Donald B Rubin. Principal stratification in causal inference. *Biometrics*, 58(1):21–29, 2002.
- Boris Freidlin, Wenyu Jiang, and Richard Simon. The cross-validated adaptive signature design. *Clinical Cancer Research*, pages 1078–0432, 2010.
- Edmund A Gehan. The determination of the number of patients required in a preliminary and a follow-up trial of a new chemotherapeutic agent. *Journal of chronic diseases*, 13(4): 346–353, 1961.
- Adolf Grünbaum. The placebo concept in medicine and psychiatry. *Psychological Medicine*, 16(1):19–38, 1986.
- Xuemin Gu, Nan Chen, Caimiao Wei, Suyu Liu, Vassiliki A. Papadimitrakopoulou, Roy S. Herbst, and J. Jack Lee. Bayesian two-stage biomarker-based adaptive design for targeted therapy development. *Statistics in Biosciences*, 8(1):99–128, 2016.
- Lacey Gunter, Ji Zhu, and Susan Murphy. Variable selection for qualitative interactions in personalized medicine while controlling the family-wise error rate. *Journal of Biopharmaceutical Statistics*, 21(6):1063–1078, 2011.
- Richard K Harrison. Phase II and phase III failures: 2013–2015, 2016.
- Trevor Hastie, Robert Tibshirani, and Jerome Friedman. Unsupervised learning. In *The Elements of Statistical Learning*, pages 485–585. Springer, 2009.
- Jiwei He and Richard Entsuah. A mixture model using likelihood-based and Bayesian approaches for identifying responders and non-responders in longitudinal clinical trials. *Pharmaceutical statistics*, 13(5):327–336, 2014.
- Yosef Hochberg. A sharper Bonferroni procedure for multiple tests of significance. *Biometrika*, 75(4):800–802, 12 1988.
- RD Holmes, AK Tiwari, and JL Kennedy. Mechanisms of the placebo effect in pain and psychiatric disorders. *The Pharmacogenomics Journal*, 16(6):491, 2016.
- David W Hosmer Jr, Stanley Lemeshow, and Rodney X Sturdivant. *Applied Logistic Regression*, volume 398. John Wiley & Sons, 2013.
- Yuri I Ingster, Christophe Pouet, and Alexandre B Tsybakov. Classification of sparse high-dimensional vectors. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 367(1906):4427–4448, 2009.
- Yuri Izmailovich Ingster and Irina A Suslina. On detection of a signal of known shape in multi-channel system. *Zapiski Nauchnykh Seminarov POMI*, 294:88–112, 2002.

Bibliography

- Rafael A Irizarry, Chi Wang, Yun Zhou, and Terence P Speed. Gene set enrichment analysis made simple. *Statistical Methods in Medical Research*, 18(6):565–575, 2018/08/16 2009.
- Farid Jamshidian, Alan E Hubbard, and Nicholas P Jewell. Accounting for perception, placebo and unmasking effects in estimating treatment effects in randomised clinical trials. *Statistical Methods in Medical Research*, 23(3):293–307, 2014.
- Christopher Jennison and Bruce W Turnbull. *Group Sequential Methods with Applications to Clinical Trials*. CRC Press, 1999. ISBN 158488858X.
- Jiashun Jin. Impossibility of successful classification when useful features are rare and weak. *Proceedings of the National Academy of Sciences*, 106(22):8859–8864, 2009.
- Jiashun Jin and T Tony Cai. Estimating the null and the proportion of nonnull effects in large-scale multiple comparisons. *Journal of the American Statistical Association*, 102(478):495–506, 2007.
- Cheryl L. Jones and Eric Holmgren. An adaptive simon two-stage design for phase 2 studies of targeted therapies. *Contemporary Clinical Trials*, 28(5):654–661, 9 2007.
- Arif Khan, Russell L Kolts, Michael E Thase, K Ranga Rama Krishnan, and Walter Brown. Research design features and patient characteristics associated with the outcome of antidepressant clinical trials. *American Journal of Psychiatry*, 161(11):2045–2049, 2004.
- Arif Khan, Amritha Bhat, Russell Kolts, Michael E Thase, and Walter Brown. Why has the antidepressant–placebo difference in antidepressant clinical trials diminished over the past three decades? *CNS Neuroscience & Therapeutics*, 16(4):217–226, 2010.
- Irving Kirsch. Are drug and placebo effects in depression additive? In *Clinical Trials in Mood Disorders: The Use of Placebo? Past, Present, and Future.*, Sep, 1999, Washington, DC, US; *Aspects of this work were presented at the aforementioned conference*. Elsevier Science, 2000.
- Irving Kirsch and Guy Sapirstein. Listening to prozac but hearing placebo: A meta-analysis of antidepressant medication. *Prevention & Treatment*, 1(2):2a, 1998.
- Andreas Kische and Ludwig A. Hothorn. Testing for qualitative interaction using ratios of treatment differences. *Statistics in Medicine*, 33(9):1477–1489, 2014.
- Jos Kleijnen, Anton JM de Craen, Jannes van Everdingen, and Leendert Krol. Placebo effect in double-blind clinical trials: a review of interactions with medications. *The Lancet*, 344(8933):1347–1349, 1994.
- Erich L Lehmann and Joseph P Romano. *Testing statistical hypotheses*. Springer Science & Business Media, 2006.
- Benjamin E Leiby, Mary D Sammel, Thomas R Ten Have, and Kevin G Lynch. Identification of multivariate responders and non-responders by using Bayesian growth curve latent class models. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 58(4):505–524, 2009.

- Jianjun Li and Ivan S. F. Chan. Detecting qualitative interactions in clinical trials: An extension of range test. *Journal of Biopharmaceutical Statistics*, 16(6):831–841, 12 2006.
- Yang Li, Alexis A. Jourdain, Sarah E. Calvo, Jun S. Liu, and Vamsi K. Mootha. Clic, a tool for expanding biological pathways based on co-expression across thousands of datasets. *PLOS Computational Biology*, 13(7):1–29, 07 2017.
- Ilya Lipkovich and Alex Dmitrienko. Strategies for identifying predictive biomarkers and subgroups with enhanced treatment effect in clinical trials using sides. *Journal of Biopharmaceutical Statistics*, 24(1):130–153, 01 2014.
- A Maitournam and R Simon. On the efficiency of targeted clinical trials. *Statistics in Medicine*, 24(3):329–339, 2005.
- Laetitia Marisa, Aurélien de Reyniès, Alex Duval, Janick Selves, Marie Pierre Gaub, Laure Vescovo, Marie-Christine Etienne-Grimaldi, Renaud Schiappa, Dominique Guenot, and Mira Ayadi. Gene expression classification of colon cancer into molecular subtypes: characterization, validation, and prognostic value. *PLoS Medicine*, 10(5):e1001453, 2013.
- Geoffrey McLachlan. *Discriminant Analysis and Statistical Pattern Recognition*. Wiley, 2004. ISBN 0471691151.
- Nicolai Meinshausen and John Rice. Estimating the proportion of false null hypotheses among a large number of independently tested hypotheses. *The Annals of Statistics*, 34(1):373–393, 2006.
- Brian A. Millen, Alex Dmitrienko, Stephen Ruberg, and Lei Shen. A statistical framework for decision making in confirmatory multipopulation tailoring clinical trials. *Drug Information Journal*, 46(6):647–656, 2017/01/16 2012.
- Daniel E Moerman and Wayne B Jonas. Deconstructing the placebo effect and finding the meaning response. *Annals of Internal Medicine*, 136(6):471–476, 2002.
- Patrick AP Moran. Maximum-likelihood estimation in non-standard conditions. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 70, pages 441–450. Cambridge University Press, 1971.
- Bengt Muthén and Hendricks C Brown. Estimating drug effects in the presence of placebo response: causal inference using growth mixture modeling. *Statistics in Medicine*, 28(27):3363–3385, 2009.
- Bengt Muthén, C Hendricks Brown, Katherine Masyn, Booil Jo, Siek-Toon Khoo, Chih-Chien Yang, Chen-Pin Wang, Sheppard G Kellam, John B Carlin, and Jason Liao. General growth mixture modeling for randomized preventive interventions. *Biostatistics*, 3(4):459–475, 2002.
- Dougu Nam. De-correlating expression in gene-set analysis. *Bioinformatics*, 26(18):i511–i516, 2010.

Bibliography

- National Cancer Institute (NCI). SEER 18 2008-2014. Colon cancer. survival by stage. . [Online; accessed 3-September-2018].
- Peter C O'Brien and Thomas R Fleming. A multiple testing procedure for clinical trials. *Biometrics*, pages 549–556, 1979.
- George I Papakostas and Maurizio Fava. Does the probability of receiving placebo influence clinical trial outcome? a meta-regression of double-blind, randomized clinical trials in mdd. *European Neuropsychopharmacology*, 19(1):34–40, 2009.
- Maja Pavlic, Richard J Brand, and Steven R Cummings. Estimating probability of non-response to treatment using mixture distributions. *Statistics in Medicine*, 20(12):1739–1753, 2001.
- Eva Petkova, Thaddeus Tarpey, and Usha Govindarajulu. Predicting potential placebo effect in drug treated subjects. *The International Journal of Biostatistics*, 5(1), 2009.
- Donald B Rubin. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66(5):688, 1974.
- Sanat K Sarkar. Some probability inequalities for ordered mtp2 random variables: a proof of the simes conjecture. *Annals of Statistics*, pages 494–504, 1998.
- Juliane Schäfer and Korbinian Strimmer. A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics. *Statistical Applications in Genetics and Molecular Biology*, 4(1), 2005.
- Steven G Self and Kung-Yee Liang. Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *Journal of the American Statistical Association*, 82(398):605–610, 1987.
- Michail Sideris and Savvas Papagrigoriadis. Molecular biomarkers and classification models in the evaluation of the prognosis of colorectal cancer. *Anticancer Research*, 34(5):2061–2068, 2014.
- D Siegmund. Estimation following sequential tests. *Biometrika*, 65(2):341–349, 1978.
- Noah Simon and Richard Simon. Adaptive enrichment designs for clinical trials. *Biostatistics*, 14(4):613–625, 09 2013.
- Richard Simon. Optimal two-stage designs for phase II clinical trials. *Controlled clinical trials*, 10(1):1–10, 1989.
- Nigel Stallard, Thomas Hamborg, Nicholas Parsons, and Tim Friede. Adaptive designs for confirmatory clinical trials with subgroup selection. *Journal of Biopharmaceutical Statistics*, 24(1):168–187, 01 2014.

- Donald E Stull, Ingela Wiklund, Rupert Gale, Gorana Capkun-Niggli, Katherine Houghton, and Paul Jones. Application of latent growth and growth mixture modeling to identify and characterize differential responders to treatment for copd. *Contemporary Clinical Trials*, 32(6):818–828, 2011.
- Xiaogang Su, Chih-Ling Tsai, Hansheng Wang, David M Nickerson, and Bogong Li. Subgroup analysis via recursive partitioning. *Journal of Machine Learning Research*, 10(Feb):141–158, 2009.
- Aravind Subramanian, Pablo Tamayo, Vamsi K Mootha, Sayan Mukherjee, Benjamin L Ebert, Michael A Gillette, Amanda Paulovich, Scott L Pomeroy, Todd R Golub, Eric S Lander, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences*, 102(43):15545–15550, 2005.
- Wenguang Sun and T Tony Cai. Oracle and adaptive compound decision rules for false discovery rate control. *Journal of the American Statistical Association*, 102(479):901–912, 2007.
- Pablo Tamayo, George Steinhardt, Arthur Liberzon, and Jill P Mesirov. The limitations of simple gene set enrichment analysis assuming gene independence. *Statistical Methods in Medical Research*, 25(1):472–487, 2016.
- T Tony Cai, X Jessie Jeng, and Jiashun Jin. Optimal detection of heterogeneous and heteroscedastic mixtures. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 73(5):629–662, 2011.
- Anastasios A Tsiatis, Gary L Rosner, and Cyrus R Mehta. Exact confidence intervals following a group sequential test. *Biometrics*, pages 797–803, 1984.
- Alexander H Tuttle, Sarasa Tohyama, Tim Ramsay, Jonathan Kimmelman, Petra Schweinhardt, Gary J Bennett, and Jeffrey S Mogil. Increasing placebo responses over time in us clinical trials of neuropathic pain. *Pain*, 156(12):2616–2626, 2015.
- Nicolas Verzelen and Ery Arias-Castro. Detection and feature selection in sparse mixture models. *The Annals of Statistics*, 45(5):1920–1950, 2017.
- NG Wahlgren, KW Ranasinha, T Rosolacci, CL Franke, PMM Van Erven, T Ashwood, L Claesson, et al. Clomethiazole acute stroke study (class): results of a randomized, controlled trial of clomethiazole versus placebo in 1360 acute stroke patients. *Stroke*, 30(1):21–28, 1999.
- Harald Walach. The efficacy paradox in randomized controlled trials of cam and elsewhere: beware of the placebo trap. 2001.
- Harald Walach and Martin Loef. Using a matrix-analytical approach to synthesizing evidence solved incompatibility problem in the hierarchy of evidence. *Journal of Clinical Epidemiology*, 68(11):1251–1260, 2018/05/29 2015.

Bibliography

- Harald Walach, Catarina Sadaghiani, Cornelia Dehm, and Dick Bierman. The therapeutic effect of clinical trials: understanding placebo response rates in clinical trials –a secondary analysis. *BMC Medical Research Methodology*, 5(1):26, 2005.
- Lily Wang, Bing Zhang, Russell D. Wolfinger, and Xi Chen. An integrated approach for the analysis of biological pathways using mixed models. *PLOS Genetics*, 4(7):1–9, 07 2008.
- Lily Wang, Xi Chen, Russell D Wolfinger, Jeffrey L Franklin, Robert J Coffey, and Bing Zhang. A unified mixed effects model for gene set analysis of time course microarray experiments. *Statistical applications in genetics and molecular biology*, 8(1):1–18, 2009.
- Sue-Jane Wang and H. M. James Hung. A regulatory perspective on essential considerations in design and analysis of subgroups when correctly classified. *Journal of Biopharmaceutical Statistics*, 24(1):19–41, 01 2014.
- DG Ward, N Suggett, Y Cheng, W Wei, H Johnson, LJ Billingham, T Ismail, MJO Wakelam, PJ Johnson, and A Martin. Identification of serum biomarkers for colon cancer by proteomic analysis. *British Journal of Cancer*, 94(12):1898, 2006.
- James MS Wason and Adrian P Mander. Minimizing the maximum expected sample size in two-stage phase II clinical trials with continuous outcomes. *Journal of Biopharmaceutical Statistics*, 22(4):836–852, 2012.
- John Whitehead, Elsa Valdés-Márquez, and Agneta Lissmats. A simple two-stage design for quantitative responses with application to a study in diabetic neuropathic pain. *Pharmaceutical Statistics: The Journal of Applied Statistics in the Pharmaceutical Industry*, 8(2): 125–135, 2009.
- Chi Heem Wong, Kien Wei Siah, and Andrew W Lo. Estimation of clinical trial success rates and related parameters. 2017.
- Di Wu and Gordon K. Smyth. Camera: a competitive gene set test accounting for inter-gene correlation. *Nucleic Acids Research*, 40(17):e133, 2012.
- Yanxun Xu, Lorenzo Trippa, Peter Müller, and Yuan Ji. Subgroup-based adaptive (suba) designs for multi-arm biomarker trials. *Statistics in Biosciences*, 8(1):159–180, 2016.
- Gur Yaari, Christopher R. Bolen, Juilee Thakar, and Steven H. Kleinstein. Quantitative set analysis for gene expression: a method to quantify gene set differential expression including gene-gene correlations. *Nucleic Acids Research*, 41(18):e170, 2013.
- Zhiwei Zhang, Richard M Kotz, Chenguang Wang, Shiling Ruan, and Martin Ho. A causal model for joint evaluation of placebo and treatment-specific effects in clinical trials. *Biometrics*, 69 (2):318–327, 2013.
- Qian Zhu, Aaron K Wong, Arjun Krishnan, Miriam R Aure, Alicja Tadych, Ran Zhang, David C Corney, Casey S Greene, Lars A Bongo, Vessela N Kristensen, Moses Charikar, Kai Li, and

Olga G Troyanskaya. Targeted exploration and analysis of large cross-platform human transcriptomic compendia. *Nature Methods*, 12:211 EP –, 01 2015.

Daria Rukina, biostatistician

(+41) 78-791-3996

EDUCATION	<i>Doctor of Philosophy (Ph.D.)</i> in Applied Mathematics-Statistics École polytechnique fédérale de Lausanne (EPFL), Switzerland Advisor: Prof. Stephan Morgenthaler	2013 - 2018
	<i>Master of Science (M.Sc.)</i> in Applied Mathematics Moscow Institute of Physics and Technology (State University), Russia Advisor: Prof. Menshikov S.A.	2011-2013
	<i>Bachelor of Science (B.Sc.)</i> in Applied Mathematics and Physics Moscow Institute of Physics and Technology (State University), Russia Advisor: Prof. Menshikov S.A.	2007-2011
RESEARCH EXPERIENCE	<i>Research assistant at EPFL</i> - experience in analyzing gene expression data and statistical models development; - statistical consulting for the Laboratory of Integrative Systems Physiology, EPFL; - teaching assistance for Master courses of Theoretical Statistics and Biostatistics;	2013-present
	<i>Bachelor and Master projects at MIPT</i> - designing and conducting game-theoretic experiments in the Laboratory of Experimental Economics, MIPT;	2011, 2013
INDUSTRIAL EXPERIENCE	<i>Intern Biostatistician</i> Novigenix SA, Lausanne <ul style="list-style-type: none">• Performed statistical analysis of data.• Built classification strategies for patients at risk for cancer. <i>IT Business Analyst</i> CRIF, Moscow <ul style="list-style-type: none">• Identified and wrote associated business requirements, functional specifications and test cases.• Active participation in negotiations with internal development team in Italy.• Taking part part in pre-sales of the several company's products.• Organizing a group of developers during internal development and testing process.	05/2017-08/2017 11/2011-10/2013
SKILLS	Software: C/C++, Statistical Data Analysis in R, MATLAB and Python. Languages: Russian (native), English (full professional proficiency), French (Professional working proficiency), German (Elementary proficiency).	

**EXTRA-
CURRICULAR
ACTIVITIES &
HOBBIES**

- *Promotion of MassChallenge accelerator in Switzerland in behalf of Inartis Foundation team, presenting information at startup events and Life Science conferences in Switzerland.*
 - *Private tutor in Mathematics and Physics for high school children and students.*
 - *Chorister in Choeur Universitaire de Lausanne, piano player, ballroom dancer, and yoga addict.*
-

