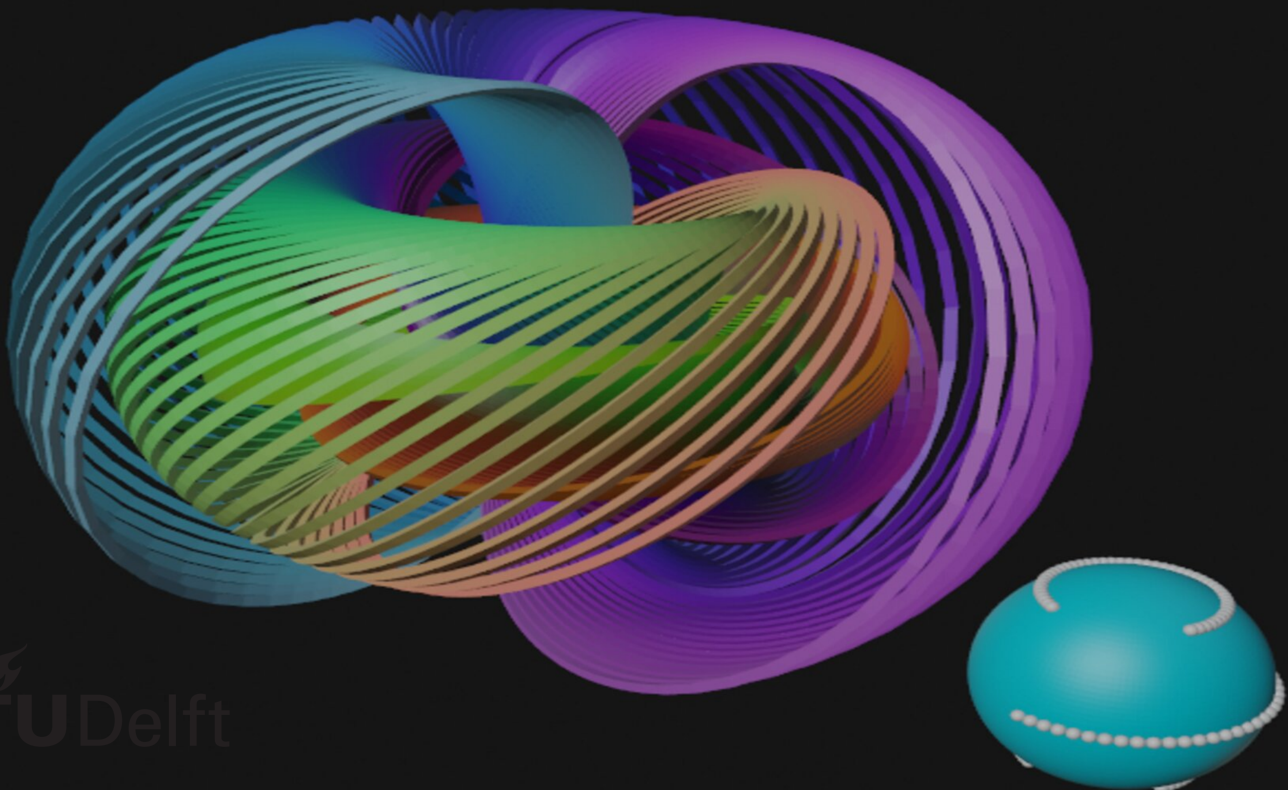
Bootstrap-Based Hypothesis Testing

Miltiadis Galanis



TU Delft

Bootstrap-Based Hypothesis Testing

by

Miltiadis Galanis

to obtain the degree of Master of Science in Applied Mathematics
at the Delft University of Technology,

Student number:	5608473	
Thesis committee:	Prof. Aad Van Der Vaart,	TU Delft
	Dr. Dorota Kurowicka,	TU Delft
	Dr. Alexis Derumigny,	TU Delft, supervisor
	Dr. Gianluca Finocchio,	University of Vienna, co-supervisor

TU Delft

Abstract

In this thesis, we explore the structure of consistent bootstrap statistics in hypothesis testing. Bootstrap, as a very useful technique when theoretical distributions are not available or when the sample size is small, enjoys a lot of interest from applied statisticians. Historically, guidelines for performing Bootstrap have been proposed. One of the guidelines proposed is to center the bootstrap statistic around the true statistic, calculated from the original sample. The second, is to perform resampling in a way such that the new sample reflects the hypothesis tested. However, both of the guidelines are proposed based mostly on an empirical point of view. In this project, we show that the calculation of the bootstrap statistic is directly related to the way the new sample is generated. We describe the specific conditions under which the Bootstrap statistic should or should not be centered around the true. As mentioned the resampling scheme that is picked directly influences this choice. The motivation is derived from the independence test and the same arguments apply to the regression slope test. Finally, we provide a generalized setting where a consistent bootstrap statistic is provided, based on the resampling scheme that is picked.

Contents

0

List of Notations

- **Probability Theory**

$\overset{i.i.d.}{\sim}$	Independent and Identically Distributed
$\overset{d}{\rightarrow}$	Convergence in distribution
$\overset{\mathbb{P}}{\rightarrow}$	Convergence in probability w.r.t. probability measure $\mathbb{P}$
$\mathbb{B}(\mathbb{D})$	Borel sets of a separable metric space $\mathbb{D}$
$\mathbb{B}(\mathbb{R}^d)$	Borel sets of $\mathbb{R}^d$
$\overset{d}{=}$	Equality in Distribution
$d_{BL}(P,Q)$	Bounded Lipschitz Distance between measures P,Q
$d_K(P,Q)$	Kolmogorov Distance between measures P,Q
δ	Dirac Measure
Pf	$\int f dP$ for a probability measure P
$\mathbb{P}_n$	Empirical measure of a given sample from $\mathbb{P}$
$\hat{\mathbb{P}}_n$	Empirical distribution of a bootstrap sample
$\mathbb{G}_{\mathbb{P}}$	$\mathbb{P}$-Brownian Bridge
$Pr(\cdot\|\mathcal{C})$	Regular Conditional Probability
$\mu_{X\|\mathcal{C}}(\cdot\|\mathcal{C})$	Regular Conditional Distribution

- **Hypothesis Testing**

$\mathbf{X}_n$	A finite sample $X_1, \ldots, X_n$
H_0	Null Hypothesis
H_1	Alternative Hypothesis
B	Number of bootstrap replications
$\hat{F}_n$	Empirical distribution of multivariate random sample
$\hat{F}_n^*$	Empirical distribution of a Bootstrap random sample
T_n	The true statistic
T_n^{*eq}	Bootstrap Equivalent statistic
T_n^{*c}	Bootstrap Centered statistic
pvalue_n	Bootstrap p-value
$\text{pvalue}_{n,B}$	Approximated bootstrap p-value
$\pi(\mathbb{P}; T_n, T_n^*)$	Power of the Bootstrap test (for any Type of Bootstrap statistic T_n^*)
GR	General Resampling Bootstrap Scheme
NHR	Null Hypothesis Resampling Bootstrap Scheme

1

Introduction

1.1. Motivation

Testing is crucial in various areas of life, including scientific research, industry, and everyday scenarios. Experiments or studies provide answers to questions such as whether a certain diet is effective for weight loss or if one brand of laptop is more reliable than another. However, the results of experiments can be ambiguous, and many conflicting hypotheses may appear viable. The theory of testing aims to formalize the decision-making process between these hypotheses.

Bootstrap hypothesis testing is a powerful and versatile non-parametric technique for assessing the validity of hypotheses based on observed data. The method relies on resampling techniques to estimate the sampling distribution of a test statistic and provides a robust alternative to classical hypothesis testing, particularly when no parametric assumptions are made for the observations or when the limiting variance of a test statistic is intractable.

Bootstrap testing can be summarized in two steps. One collects a random sample from a distribution and calculates a test statistic, referred to as the "true test statistic", based on this sample. Then the statistician generates a new sample, called "Bootstrap sample", from the original. From this bootstrap sample is computed a "Bootstrap test statistic", corresponding to the true test statistic. A large distance between the bootstrap test statistic and the true test statistic constitutes indication against a null hypothesis. However, the way the Bootstrap sample is generated or the formula used to calculate a Bootstrap statistic directly influences the "appropriateness" of the statistical test. The extensive interest in Bootstrap in various applications, for instance in the medical sector [12][11], presents the necessity to provide a solid mathematical framework, in order to assist applied scientists and transfuse confidence in the experiments' results.

Various suggestions have been proposed to efficiently calibrate a Bootstrap hypothesis test. In [7] the importance of ensuring that the resampling process reflects the null hypothesis, even when the data might not be drawn from a population that satisfies the null hypothesis, is emphasized. The first guideline that is proposed in performing Bootstrap testing, in this work, suggest to center the Bootstrap statistic around the true, in order for the Bootstrap statistic to reflect the null hypothesis. In other words, given a true statistic, calculated as a functional on the original sample, and its Bootstrap equivalent calculated as the same functional on the Bootstrap sample, the Bootstrap statistic proposed to perform the test is the true statistic substracted from its Bootstrap equivalent. The argument for this suggestion is that the test performed without centering does not lead to a correct decision rule accordingly with the statistical testing framework. Indeed, in the setting discussed in this work, failing to center the Bootstrap statistic results to a poor hypothesis test. The empirical results in this paper make an argument for their suggestion. However, in this project only the standard non parametric Bootstrap was used. Hence a systematic comparison between different kinds of resampling schemes deems necessary, as well as the probabilistic framework that explains the empirical observations.

Various resampling schemes are known in the literature, the most common of which being the nonparametric bootstrap. It is a process that picks observations uniformly from the original sample. In [3], the Wild Bootstrap for regression with heteroskedastic noise is introduced, while in [2] a similar

Multiplier (Wild) Bootstrap is used to estimate the maximum of a sum of random vectors. Remillard and Genest [6] prove the efficiency of the parametric Bootstrap in the parametric Goodness of Fit test. In [10, Section 3.8], another type of resampling is discussed for the independence test: from a bivariate dataset, the statistician resamples independently the two variables using the nonparametric Bootstrap, and concatenates the two univariate Bootstrap samples to create a new bivariate Bootstrap sample. The results from [10] imply that, in this setting, the equivalent Bootstrap statistic without centering leads to a valid hypothesis test.

For the unsuspecting eye, this seems to contradict the guideline from [7]. The apparent contradiction motivates the interest to discuss and research how to properly resample and calculate a bootstrap test statistic. Hence, the impetus to perform simulations with various resampling schemes and bootstrap statistics in order to detect any possible pattern. For different resampling schemes, the statistician may need to use a different bootstrap test statistic. More specifically, we experiment with two types of Bootstrap schemes, where the first type of Bootstrap, mentioned as "General Resampling (GR)" throughout this project, generates a sample that is "very similar" in distribution to the original regardless of whether the null hypothesis holds or not, while the second type of Bootstrap, mentioned as "Null Hypothesis Resampling (NHR)", generates samples that always satisfy the null hypothesis. First, by doing simulations, we explore the relationships between each of these type of resampling schemes and the centering/lack of centering of the bootstrap test statistic. In [8] the NHR scheme is assessed with respect to a resampling scheme the authors refer to as "alternative" resampling. The latter involves generating Bootstrap samples that satisfy the setting's alternative hypothesis. The authors present some cases of NHR in the settings of correlation coefficient test, variance test and goodness of fit test. In [9] the same NHR scheme as in [8] is discovered for the goodness of fit test. Finally, in [4] Bootstrap schemes for conditional copulas tests are compared and evaluated, where a natural NHR scheme is derived and proved to consist a proper option for the test.

Our first simulation results, detailed in Section 1.6, indicate an interesting implication. When performing a GR Bootstrap the centered statistic seems to lead to a hypothesis test with good level, while the un-centered statistic does not. And vice versa, when performing a NHR Bootstrap the un-centered statistic seems to lead to a correct test, while the centered does not. The simulations are performed in the settings of the independence test and the regression slope test, where one would like to test whether the slope between the dependent and explanatory variable is zero. A common GR Bootstrap used, in both cases, is the non-parametric Bootstrap. The NHR Boostrap scheme used in the independence test is the one mentioned earlier from [10], which generates independent joint samples that directly satisfy the null hypothesis. The nature of the regression slope test, and its deeper complexity, provides an environment where one could easily think of many ways to generate new samples, hence there is a greater variety of Bootstrap schemes. A NHR Bootstrap in the regression is the same independent resampling as the NHR in the independence test, which happens to satisfy the null hypothesis in this setting too.

The natural conjecture that arose is that the two specific combinations mentioned, of Bootstrap statistics and Bootstrap schemes, perform well with each other, while other combinations fail. In this work, we prove this claim first in the independence test. Then we proceed with the proof of the equivalent argument in the regression slope test. Finally, we present a more general argument in bootstrap hypothesis testing, of which all the previous arguments become special cases.

1.2. Hypothesis Testing

Let $\mathcal{P}$ denote the set of all probability measures that could potentially describe the distribution of a sample, which serves as the parameter space for the hypothesis test. The parameter space can be partitioned into two disjoint subsets, namely $\mathcal{P}_0$ and $\mathcal{P}\backslash\mathcal{P}_0$. We wish to test the null hypothesis $H_0 : \mathbb{P} \in \mathcal{P}_0$ against the alternative hypothesis $H_1 : \mathbb{P} \in \mathcal{P}\backslash\mathcal{P}_0$, where $\mathcal{P}_0$ and is a subset of the space of possible distribution $\mathcal{P}$.

Definition 1.2.1. *Let X be a random variable with $X \sim \mathbb{P}$ and let $\mathcal{P}$ be a family of distributions such that $\mathbb{P} \in \mathcal{P}$. Let $\mathcal{P}_0$ and $\mathcal{P}\backslash\mathcal{P}_0$ be a partition of $\mathcal{P}$.*

The null hypothesis is a statement about the distribution of X. It is denoted by H_0 and defined as:

$$H_0 : \mathbb{P} \in \mathcal{P}_0$$

On the other hand, the alternative hypothesis is a statement that contradicts the null hypothesis and is denoted by H_1. It is defined as:

$$H_1 : \mathbb{P} \in \mathcal{P} \backslash \mathcal{P}_0$$

Definition 1.2.2. *Given a null hypothesis H_0, a statistical test consists of a set K of possible values for the observation X, the critical region; this can potentially be a random set. Let X be a random variable, that represents an observation. H_0 is rejected, if $X \in K$; H_0 is not rejected, if $X \notin K$.*

In hypothesis testing, the possible mistakes can be classified into two types based on our conclusions drawn from the observations:

- Type I error occurs when we reject the null hypothesis (H_0) while it is true;
- Type II error occurs when we fail to reject H_0 while it is false.

Since we prioritize the null hypothesis in hypothesis testing, it is crucial to choose H_0 and H_1 carefully. The statement we want to demonstrate is formulated as the alternative hypothesis, and we argue for H_0 unless there is strong evidence against it.

The quality of a test is measured by the function $\mathbb{P} \mapsto \mathbb{P}(X \in K)$, also known as the power function. The following definition introduces the power function as the function $\pi(\mathbb{P}; K) = \mathbb{P}(X \in K)$ for a test with critical region K. The power function should have small values when $\mathbb{P} \in \mathcal{P}_0$ and large values when $\mathbb{P} \in \mathcal{P} \backslash \mathcal{P}_0$.

Definition 1.2.3. *The power function of a test with critical region K is $\mathbb{P} \to \pi(\mathbb{P}; K) = \mathbb{P}(X \in K)$.*

We seek a critical region for which the power function takes on ”small values” (close to 0) when $\mathbb{P} \in \mathcal{P}_0$ and ”large values” (close to 1) when $\mathbb{P} \in \mathcal{P} \backslash \mathcal{P}_0$.

Definition 1.2.4. *We say that the test achieves good level when the power function is small for $\mathbb{P} \in \mathcal{P}_0$ and that the test achieves good power when the power function converges to 1 for $\mathbb{P} \in \mathcal{P} \backslash \mathcal{P}_0$.*

Remark 1.2.5. *The level of the test, which means the acceptable value of the power function under the null hypothesis, is defined arbitrarily and denoted by α. So a hypothesis test with a level α is a test that falsely rejects the null hypothesis with probability at most α.*

1.3. Preliminaries

In this section, we introduce the general probabilistic framework to define Bootstrap schemes and probability measures of Bootstrap samples. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and $\mathbb{D}$ a metric space.

Definition 1.3.1. *Let $(\mathbb{D}, d)$ be a metric space. $\mathbb{D}$ is said to be a separable space if there exists a countable dense subset U of $\mathbb{D}$, namely for every $x \in \mathbb{D}$ there exists a sequence $(x_n)_{n=1}^{\infty}$ such that for every $n \in \mathbb{N}$, $x_n \in U$ and $x_n \to x$.*

For the rest of this project, $(\mathbb{D}, d)$ represents a metric space and $\mathbb{B}(\mathbb{D})$ denotes the Borel sets of $\mathbb{D}$.

Definition 1.3.2. *Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of $\mathbb{D}$-valued random variables defined on $(\Omega, \mathcal{A}, \mathbb{P})$.*

1. *We say that the sequence (X_n) converges in distribution to the random variable X ($X_n \xrightarrow{d} X$) if for every $A \in \mathbb{B}(\mathbb{D})$:*

$$\mathbb{P}(X_n \in A) \to \mathbb{P}(X \in A) \text{ as } n \to +\infty.$$

2. *We say that the sequence (X_n) converges in probability to the random variable X ($X_n \xrightarrow{\mathbb{P}} X$) if for every $\epsilon > 0$:*

$$\mathbb{P}(d(X_n, X) > \epsilon) \to 0 \text{ as } n \to +\infty.$$

In the following, we will need the concepts of “version” of a random variable and “regular conditional distribution”. For completeness, we state these definitions.

Definition 1.3.3. *Let $X : \Omega \to \mathbb{D}$ be a random variable in the probability space $(\Omega, \mathcal{A}, \mathbb{P})$. The random variable $Y : \Omega \to \mathbb{D}$ is said to be a version of X if $\mathbb{P}(B) = 0$ where B is the set defined by $B := \{\omega \in \Omega : X(\omega) \neq Y(\omega)\}$.*

By definition of the set B, whenever Y is a version of X we have $X|_{\Omega\setminus B} = Y|_{\Omega\setminus B}$, where for a set $A \subseteq \Omega$, $X|_A$ denotes the restriction of the mapping X in A, namely $X|_A : A \mapsto \mathbb{R}$ such that $\forall\omega \in A, X|_A(\omega) = X(\omega)$.

Definition 1.3.4 (Section 5.1.3 in [5]). *Let X be a random variable, $X : (\Omega, \mathcal{A}, \mathbb{P}) \mapsto (\mathbb{D}, \mathbb{B}(\mathbb{D}))$. For each $B \in \mathbb{B}(\mathbb{D})$, let:*

$$\mu_{X|\mathcal{C}}(B|\mathcal{C})(\omega) := \mathbb{P}(X^{-1}(B)|\mathcal{C})(\omega).$$

Then the mapping $\mu_{X|\mathcal{C}}(\cdot|\mathcal{C})(\cdot) : \mathbb{B}(\mathbb{D}) \times \Omega \mapsto [0, 1]$ is a regular conditional distribution of X given $\mathcal{C}$ if the following hold:

1. *$\forall B \in \mathbb{B}(\mathbb{D})$, $\mu_{X|\mathcal{C}}(B|\mathcal{C})(\cdot)$ is a version of $\mathbb{E}(1_{X\in B}|\mathcal{C})$.*
2. *For $\mathbb{P}$-a.s. $\omega \in \Omega$, $\mu_{X|\mathcal{C}}(\cdot|\mathcal{C})(\omega)$ is a probability measure on $(\mathbb{D}, \mathbb{B}(\mathbb{D}))$.*

1.4. The Bootstrap

Bootstrapping is a technique that involves resampling data to perform inference about a population. For instance, one would like to estimate the variance of an estimator, as the true error in a sample statistic against its population value is usually unknown. In bootstrapping, the "population" is the sample itself, which is known, allowing for the measurement of the quality of inference of the true sample from resampled data. This method estimates the distribution of an estimator or test statistic and is comparable in accuracy to first-order asymptotic theory, making it a practical alternative in cases where calculating the asymptotic distribution is challenging.

Let $X : \Omega \to \mathbb{D}$ be a random variable following the (unknown) distribution $\mathbb{P}$ and let $X_1, \dots, X_n$ be a random sample from X, namely $\mathbf{X}_n = (X_1, \dots, X_n) \overset{i.i.d.}{\sim} \mathbb{P}$. The objective is to estimate the distribution of a $\mathbb{E}$-valued statistic, where $(\mathbb{E}, d)$ is a metric space, denoted by $T_n = T_n(\mathbf{X}_n)$. In the context of hypothesis testing, T_n represents the true test statistic and we need to estimate or approximate its distribution.

For this purpose, we aim to construct bootstrap replicates of T_n. First, we need to introduce the i.i.d. random variables $(W_n^{(b)})_{b\in\mathbb{N}}$ that introduce additional sources of randomness. They are assumed to be valued in some measurable space $\mathcal{W}$ and independent of the original sample $\mathbf{X}_n$. The bootstrapped observations $\mathbf{X}_n^*$ are generated from the original sample $(X_1, \dots, X_n)$ and some other source of randomness $W_n^{(b)}$. In this sense, we can write $\mathbf{X}_n^* = \mathbf{X}_n^*(X_1, \dots, X_n; W_n^{(b)})$. These bootstrapped observations are generated such that the law of $\mathbf{X}_n^*$ given the original sample $(X_1, \dots, X_n)$ is precisely the probability measure R_n. This is formalized in the following definitions.

Definition 1.4.1. *A bootstrapping scheme $(R_n)_{n=1}^{\infty}$ is a sequence of probability measures on $\mathcal{X}$ that depends on the observed sample $(X_1, \dots, X_n)$. In that sense, for a given sample $\mathbf{X}_n$, $R_n : \mathbb{D}^n \mapsto \tilde{\mathcal{P}}$, where $\tilde{\mathcal{P}}$ is the set of probability measures on $\mathbb{D}$.*

Definition 1.4.2. *For a Bootstrap scheme $(R_n)_{n=1}^{\infty}$, a bootstrap sample $\mathbf{X}_n^* = (X_1^*, \dots, X_n^*)$ is an i.i.d. sample of size n simulated from R_n, conditionally to the original sample $\mathbf{X}_n$.*

In this sense, we define resampling as a sequence of measures, mapping each finite sample of size n to a probability measure. This measure is used to simulate the new observations in the bootstrap sample. As a consequence, for any measurable set $A \subseteq \mathcal{X}$,

$$\text{Prob}(X_1^* \in A|\mathbf{X}_n) = R_n(\mathbf{X}_n)(A),$$

and for any measurable sets $A_1, \dots, A_n \subseteq X$,

$$\text{Prob}(\mathbf{X}_n^* \in A_1 \times \cdots \times A_n|\mathbf{X}_n) = \prod_{i=1}^{n} R_n(\mathbf{X}_n)(A_i).$$

Let $B > 0$ be the number of boostrap replications, and let $\mathbf{X}_n^{*(i)}, i = 1, \dots, B$ be B independent samples of size n from the bootstrap scheme $R_n(\mathbf{X}_n)$. For a given mapping $T^* : \mathcal{X}^n \times \mathcal{X}^n \to \mathbb{R}$, we define the bootstrapped statistics $T_n^{*(b)} := T^*\big(\mathbf{X}_n, \mathbf{X}_n^{*(b)}\big)$ for any $b \geq 0$. Therefore, the random variables $T_n^{*(1)}, T_n^{*(2)}, \ldots$ are conditionally independent given $\mathbf{X}_n$.

Definition 1.4.3. *With the previous notation, the p-value is defined by:*

$$pvalue_n := \text{Prob}\left(T^{*(1)} > T_n(\mathbf{X}_n) | \mathbf{X}_n\right), \tag{1.1}$$

and the (bootstrapped approximated) p-value is defined by

$$pvalue_{n,B} := \frac{\left|i \in \{1, \ldots, B\} : T_n^{*(i)} > T(X_1, \ldots, X_n)\right|}{B} = \frac{1}{B}\sum_{i=1}^{B} 1\left(T_n^{*(i)} > T_n\right), \tag{1.2}$$

and as usual, for a given level $\alpha \in (0,1)$, we reject if and only if the p-value is smaller than α.

Remark 1.4.4. *In the particular case where the original test statistic $T_n(X_1, \ldots, X_n)$ is bigger than all the bootstrapped test statistics $T^{*(i)}$, then the (bootstrapped approximated) p-value is 0 and we reject at all levels.*

1.5. Bootstrap Consistency

In order to determine the consistency of a resampling scheme, it is necessary to show that the distance between the conditional distribution of a bootstrap replicate of a statistic T_n^*, given the available observations, and the distribution of T_n itself, converges to zero in probability. Interestingly, under minimal assumptions, this convergence of conditional laws is equivalent to the unconditional weak convergence of T_n^* jointly with two bootstrap replicates to independent copies of the same limit. Additionally, the distance between the empirical distribution of the bootstrap replicates and the unobservable distribution of T_n^* should converge in probability to zero as the number of replicates and the sample size increase.

Note that there are various metrics that can potentially be used to measure the distance between two distributions. In this framework we will make use of the Bounded-Lipschitz distance d_{BL} and the Kolmogorov distance d_K, defined below.

Definition 1.5.1. *Let $(\mathbb{D}, d)$ be a metric space and $BL = \{h : \mathbb{D} \to [-1,1]$ such that $|h(x) - h(y)| \leq d(x,y), \forall x, y \in \mathbb{D}\}$. Then for any probability measures P, Q on $\mathbb{D}$, the Bounded Lipschitz metric is defined as:*

$$d_{BL}(P,Q) := \sup_{f \in BL} \left| \int f dP - \int f dQ \right|.$$

The Kolmogorov distance between two measures on $\mathbb{R}^d$ is defined as:

$$d_K(P,Q) := \sup_{x \in \mathbb{R}^d} \left| P((-\infty, x]) - Q((-\infty, x]) \right|.$$

Definition 1.5.2. *Let $\mathbf{X}_n$ be a random sample from a distribution function F on $\mathbb{R}^d$, for some $d > 0$. If we denote $X_i = (X_{i,1}, \ldots, X_{i,d})$, the empirical distribution function is defined as:*

$$F_n(\mathbf{t}) := \frac{1}{n}\sum_{i=1}^{n} 1(X_{i,1} \leq t_1, \ldots, X_{i,d} \leq t_d), \tag{1.3}$$

where $\mathbf{t} = (t_1, \ldots, t_d) \in \mathbb{R}^d$.

Definition 1.5.3. *Let $\mathbf{X}_n$ be a random sample from a probability measure $\mathbb{P}$ on a measurable space $(\Omega, \mathcal{A})$. The empirical measure is the discrete uniform measure on the observations, denoted by:*

$$\mathbb{P}_n := \frac{1}{n}\sum_{i=1}^{n} \delta(X_i), \tag{1.4}$$

where $\delta(x)$ is the Dirac measure:

$$\delta(x)(A) := \begin{cases} 1, & \text{if } x \in A, \\ 0, & \text{if } x \notin A, \end{cases}$$

for any $x \in \mathcal{X}$.

Remark 1.5.4. *Assume that $\mathbb{D} = \mathbb{R}^d$ for a given $d > 0$. In this space, the empirical distribution can be seen as a special case of the empirical measure. Indeed pick $A = ((-\infty, \dots, -\infty), \mathbf{t}]$ and we get:*

$$\delta(\mathbf{X})(A) = 1(X \leq \mathbf{t}).$$

In order to achieve good power in hypothesis testing, it is necessary to construct statistics that when the null hypothesis holds, both the true and the Bootstrap statistics converge to the same distribution. In the context of hypothesis testing, the Bounded Lipschitz Distance between the distributions of those statistics converging to 0 is the sufficient condition to check, in order for this to hold. When this is achieved we say that the Bootstrap statistic is consistent with the true statistic. Before stating this important result, we state a condition on the regularity of the conditional distribution of a bootstrap statistic.

Condition 1. *[Condition 2.1 in [1]] The conditional distribution $\mathbb{P}\left(T_n^{*(1)} \in \bullet \,|\mathbf{X}_n\right)$ admits a regular version, denoted by $\mathbb{P}^{T_n^{*(1)}|\mathbf{X}_n}$.*

Note that if $\mathbb{E}$ is complete and separable, then Condition 1 is satisfied. This can be proved by combining Theorems 2.1.15 and 5.1.9 of [5]. This is the case in particular if the statistic $T_n^{*(1)}$ is real-valued. The following Lemma from [1] presents the equivalent conditions of Bootstrap consistency.

Lemma 1.5.5. *[Lemma 2.2 in [1]] Let $\mathbf{X}_n$ be a sample of size n from $\mathbb{P}$. For a sequence of Bootstrap samples of size B, and the respective statistics $T_n^{*(1)}, \dots, T_n^{*(B)}$, denote the empirical measure of the Bootstrap statistics by*

$$\hat{\mathbb{P}}_B^{T_n^*} := \frac{1}{B} \sum_{i=1}^{B} \delta(T_n^{*(i)}). \tag{1.5}$$

Assume that Condition 1 is met and that the true statistic T_n converges weakly to a random variable T in $\mathbb{D}$. Then the following are equivalent:

1. $\mathbb{P}^{\left(T_n, T_n^{*(1)}, T_n^{*(2)}\right)} \xrightarrow{d} \mathbb{P}^T \otimes \mathbb{P}^T \otimes \mathbb{P}^T$ *as* $n \to \infty$.
2. $\mathbb{P}^{\left(T_n,\, T_n^{*(1)},\, \dots,\, T_n^{*(B)}\right)} \xrightarrow{d} \left(\mathbb{P}^T\right)^{\otimes(B+1)}$ *as* $n \to \infty$ *and for* $B \geq 2$.
3. $d_{BL}\left(\mathbb{P}^{T_n^{*(1)}|\mathbf{X}_n}, \mathbb{P}^{T_n}\right) \xrightarrow{\mathbb{P}} 0$ *as* $n \to \infty$.
4. $d_{BL}\left(\hat{\mathbb{P}}_B^{T_n^*}, \mathbb{P}^{T_n}\right) \xrightarrow{\mathbb{P}} 0$ *as* $n, B \to \infty$.

   If additionally $\mathbb{D} = \mathbb{R}^d$ and the cumulative distribution function of T is continuous then the preceding conditions are also equivalent to:

5. $d_K\left(\mathbb{P}^{T_n^{*(1)}|\mathbf{X}_n}, \mathbb{P}^{T_n}\right) \xrightarrow{\mathbb{P}} 0$ *as* $n \to \infty$.
6. $d_K\left(\hat{\mathbb{P}}_B^{T_n^*}, \mathbb{P}^{T_n}\right) \xrightarrow{\mathbb{P}} 0$ *as* $n, B \to \infty$.

For a consistent Bootstrap statistic T_n^* with respect to a true statistic T_n, we will be showing that T_n^* conditionally on $\mathbf{X}_n$ and T_n converge in distribution to the same limit. This means that equivalent condition 3 of Lemma 1.5.5 holds. Instead of the empirical measures, we denote the Bounded-Lipschitz distance by the statistics themselves. Hence to make the notation lighter, in the rest of the text $d_{BL}\,(T_n\,,\, T_n^*|\mathbf{X}_n)$ represents $d_{BL}\left(\mathbb{P}^{T_n}, \mathbb{P}^{T_n^*|\mathbf{X}_n}\right)$.

The following Corollary from [1] is crucial in recognizing which bootstrap statistics lead to good power and getting a sense of the central claim of this thesis. The corollary states that the p-values p_n of a consistent Bootstrap converge in distribution to a uniform random variable. This property leads to controlling the Type I error of the hypothesis test, meaning that for a test of significance α, we falsely reject the null hypothesis only α-% of the time.

Corollary 1.5.6 (Corollary 4.3 in [1]). *Assume that $\mathbb{D} = \mathbb{R}$ and that T has a continuous distribution function. Assume that one of the equivalent conditions in Lemma 1.5.5 is met. Then*

$$pvalue_n \xrightarrow{d} U(0,1), \text{ and } pvalue_{n,B} \xrightarrow{d} U(0,1) \text{ as } n, B \to \infty, \tag{1.6}$$

where the p-values are as defined in Equation (1.2).

When a proper Bootstrap scheme and statistic is picked, the Bootstrap p-value, as defined in Equation (1.1), holds its level asymptotically, in the sense that, under H_0,

$$\lim_{n\to\infty} \text{Prob}\left(T_n^{*(1)} > T_n\right) = 1 - \alpha,$$

or equivalently:

$$\lim_{n\to\infty} \text{Prob}\left(T_n^{*(1)} \leq T_n\right) = \alpha,$$

as a consequence of Corollary 1.5.6.

1.6. Introductory examples

As mentioned before, a consistent Bootstrap statistic leads to asymptotically uniform p-values, when the null hypothesis holds. In our simulations, the p-values histogram for each combination of Bootstrap schemes (GR,NHR) and Bootstrap statistics (T_n^{*eq}, T_n^{*c}), is enlightening. Corollary 1.5.6 leads us to search for the uniform, or better "approximately" uniform, histogram of p-values. From the four combinations of Bootstrap schemes and Bootstrap statistics, four different histograms are obtained but only two of them indicate convergence to the uniform distribution.

We initially focus on examining the behavior of two types of bootstrap statistics in the settings of the independence test and the regression slope test. Given an original sample $\mathbf{X}_n$, a true statistic T_n and a Bootstrap sample $\mathbf{X}_n^*$, denote by T_n^{*eq} the Bootstrap equivalent of T_n and T_n^{*c} the Bootstrap centered statistic.

1.6.1. Regression Slope test

Let X, Y be random variables in $\mathbb{R}$ such that $(X, Y) \sim \mathbb{H}$, $X \sim \mathbb{P}, Y \sim \mathbb{Q}$ where $\mathbb{P}, \mathbb{Q}, \mathbb{H}$ are restricted so that they satisfy the following setting. First, X has a finite second moment, namely $\mathbb{E}_{\mathbb{P}}(X^2) < \infty$. The variable ϵ, called "noise" is independent of X and Y and satisfies $\mathbb{E}(\epsilon^2) = \sigma_\epsilon^2$ for some $\sigma_\epsilon^2 > 0$. Then let $a, b \in \mathbb{R}$ such that conditionally on an observation X, Y is described as:

$$Y = a + bX + \epsilon. \tag{1.7}$$

The probability measure $\mathbb{H}$ is called the joint probability measure, while $\mathbb{P}, \mathbb{Q}$ are called the marginals of X, Y respectively. Note that $\mathbb{Q}$ does not describe Equation (1.7), since this equation describes the conditional distribution of Y given X. The definition of b in the regression setting is:

$$b = \frac{\mathbb{E}_{\mathbb{H}}(XY) - \mathbb{E}_{\mathbb{H}}(X)\mathbb{E}_{\mathbb{H}}(Y)}{\mathbb{E}_{\mathbb{H}}(X^2) - \mathbb{E}_{\mathbb{H}}(X)^2} = \frac{Cov_{\mathbb{H}}(X,Y)}{Var_{\mathbb{H}}(X)}.$$

The number b is called the slope of Y given X. For a finite sample $(X_1, Y_1), \dots, (X_n, Y_n)$, the OLS estimator is calculated as:

$$\hat{b} = \frac{\sum_{i=1}^n X_i Y_i - \sum_{i=1}^n X_i \sum_{i=1}^n Y_i}{\sum_{i=1}^n X_i^2 - (\sum_{i=1}^n X_i)^2}.$$

Consider the test:

$$H_0 : b = 0 \text{ vs } H_1 : b \neq 0.$$

The true statistic is set as $T_n = |\hat{b}|$. Assume resmapling scheme $(R_n)_{n=1}^\infty$, where for each $n \in \mathbb{N}$, the number of Bootstrap replications generated is fixed $B = n$. For fixed $n \in \mathbb{N}$, let $(X_1^*, Y_1^*), \dots, (X_n^*, Y_n^*)$ bootstrap sample generated by R_n. The two Bootstrap statistics are $T_n^{*eq} = |\hat{b^*}|$ and $T_n^{*c} = |\hat{b}^* - \hat{b}|$, where:

$$\hat{b}^* = \frac{\sum_{i=1}^n X_i^* Y_i^* - \sum_{i=1}^n X_i^* \sum_{i=1}^n Y_i^*}{\sum_{i=1}^n (X_i^*)^2 - (\sum_{i=1}^n X_i^*)^2},$$

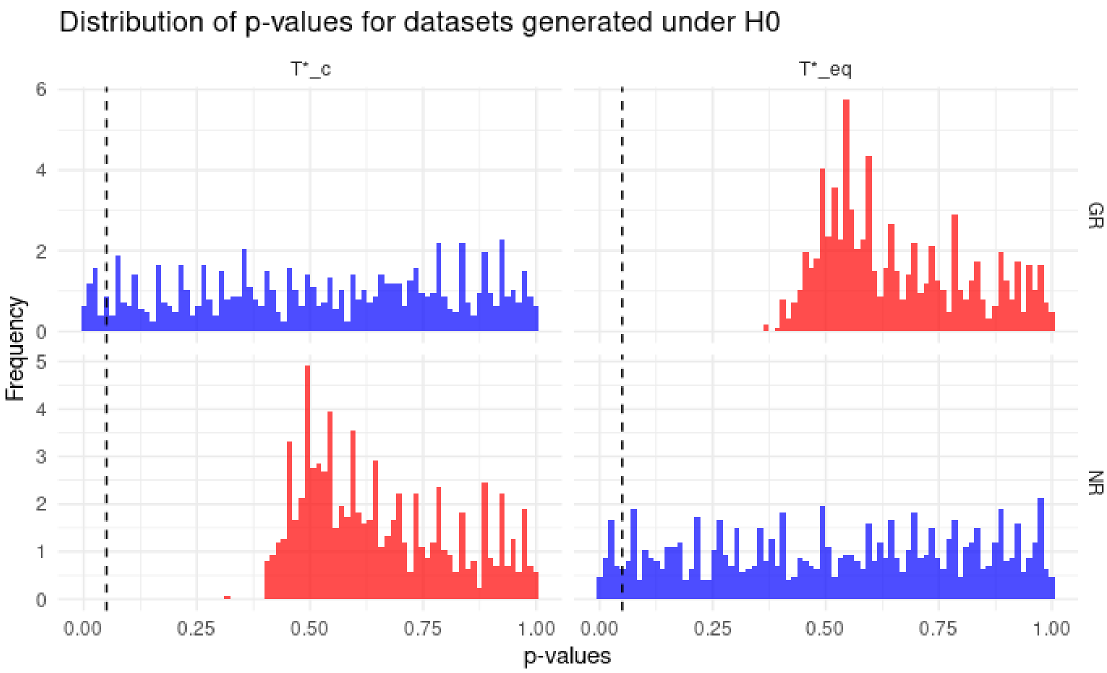
is the OLS estimator calculated on the Bootstrap sample. Although we can state various examples for each type of resampling scheme we provide the context of one from each type. A general resampling scheme is simulating from the joint empirical measure $\mathbb{H}_n = \frac{1}{n}\sum_{i=1}^n \delta(X_i, Y_i)$. An example from the null resampling category is generating samples from the empirical product $\mathbb{P}_n \otimes \mathbb{Q}_n$, where $\mathbb{P}_n, \mathbb{Q}_n$ are the empirical measures of the samples $\mathbf{X}_n, \mathbf{Y}_n$ respectively. For a sample generated by $\mathbb{P}_n \otimes \mathbb{Q}_n$ we see that $Cov_{\mathbb{P}_n \otimes \mathbb{Q}_n}(X, Y) = 0$. If one modelled the Bootstrap samples generated by this measure as a regression model, the theoretical slope would be zero and hence the "null" resampling characterization.

We perform $s = 1000$ simulations for each one of the combinations between Bootstrap statistics and Bootstrap schemes. On each iteration, we generate a random sample of size $n = 100$, of X_i, Y_i, where the X_i are generated from a uniform distribution between random lower and upper bounds. An intercept α is also generated by a uniform distribution between random bounds and a random noise ϵ is generated from the standard normal distribution. Finally, b is set to be zero and the Y_i are set to be:

$$Y_i = a + bX_i + \epsilon_i.$$

For each iteration, $b = 100$ Bootstrap observations are generated for each one of the resampling schemes GR, NHR. For every Bootstrap sample, the pvalues are calculated according to the statistics T_n^{*c}, T_n^{*eq} and are appended to a list. Hence we obtain 4 lists containing 1000 pvalues, one pvalue for each one of the Bootstrap samples generated and one list for each one of the combinations among GR,NHR and T_n^{*c}, T_n^{*eq}. The histogram of the pvalues for each combination are displayed in Figure 1.1.

Figure 1.1: Histogram of the Bootstrap p-values when H_0 is true



Blue color histograms represent consistent Bootstrap processes. Red color histograms represent inconsistent Bootstrap.

From this figure, we make an obvious remark. The pvalues histograms colored with blue seem to approximate a uniform distribution, while the red colored ones seem to have some erratic behavior among them and between the red. This observation leads us to the inclination to believe that Corollary 1.5.6 is satisfied for the pvalues displayed in the main diagonal of the table, while the red histograms remain to present unidentified behavior. The asymptotic uniform distribution of the pvalues is a direct consequence of a consistent Bootstrap process. This property renders the decision rule meaningful, such that the statistician can control the Type I Error and falsely rejects a null hypothesis with probability α.

We proceed with providing a definition of the two types of Bootstrap schemes we examine throughout this project. This is an informal version of the Definition provided in Chapter 5 and should not be examined with a strict point of view but rather with an intuitive approach.

Definition 1.6.1 (Informal Version of Definition 5.1.2)**.** *Let $X_1, \dots, X_n$ be a random sample from $\mathbb{H}$ and let $\mathcal{P}_0, \mathcal{P}\backslash\mathcal{P}_0$ be a partition of the class of probability measures $\mathcal{P}$, such that the null hypothesis holds if $\mathbb{H} \in \mathcal{P}_0$. Let $\mathbb{H}_n$ denote its empirical measure. A Bootstrap scheme $\hat{R} = (R_n)_{n=1}^{\infty}$ is called*

- ***General Resampling (GR)*** *scheme if:*

$$\forall \mathbb{H} \in \mathcal{P},\ R_n \approx \mathbb{H}_n,$$

- ***Null Resampling (NHR)*** *scheme if:*

$$\forall \mathbb{H} \in \mathcal{P}_0,\ R_n \approx \mathbb{H}_n, \text{ and } \forall \mathbb{H} \in \mathcal{P},\ R_n \in \mathcal{P}_0.$$

This leads us to the following conjecture.

Conjecture:

- Under GR Bootstrap, R_n, for the Bootstrap statistic T_n^{*c}:

$$\text{for all } \mathbb{H} \in \mathcal{P}_0,\ \lim_{n\to\infty} \pi(\mathbb{H}; T_n, T_n^{*c}) = \alpha,$$

  and

$$\text{for all } \mathbb{H} \in \mathcal{P}\backslash\mathcal{P}_0,\ \lim_{n\to\infty} \pi(\mathbb{H}; T_n, T_n^{*c}) = 1.$$

- Under NHR Bootstrap, R_n, for the Bootstrap statistic T_n^{*}:

$$\text{for all } \mathbb{H} \in \mathcal{P}_0,\ \lim_{n\to\infty} \pi(\mathbb{H}; T_n, T_n^{*eq}) = \alpha,$$

  and

$$\text{for all } \mathbb{H} \in \mathcal{P}\backslash\mathcal{P}_0,\ \lim_{n\to\infty} \pi(\mathbb{H}; T_n, T_n^{*eq}) = 1,$$

Notice that in order to show that a Bootstrap Hypothesis test will not lead to a good power, it suffices to show that the Bootstrap statistic used in not consistent with the true statistic according equivalent condition 5 of Lemma 1.5.5. The goal of this project is to show that given a Bootstrap scheme $\hat{R}$ that belongs to one of the categories GR or NHR, only one of the two bootstrapped test statistics T_n^{*eq} or T_n^{*c} will lead to good power, even given arbitrary sample size n of data from $\mathbb{P}$ and arbitrary computer power B to simulate new replications.

2

Empirical Processes and Weak Convergence

In this chapter, we provide the tools that are used to examine the asymptotic behavior of the true and Bootstrap statistics. Naturally, the statistics that are considered are mostly based on empirical measures constructed from samples and Bootstrap samples. Hence, for a probability measure $\mathbb{P}$, a finite sample $\mathbf{X}_n$ and its empirical measure $\mathbb{P}_n$ as defined in 1.5.3, the quantity $\mathbb{P}_n - \mathbb{P}$ is of extreme interest. The Central Limit Theorem for processes implies that this quantity converges at the rate $O\left(\frac{1}{\sqrt{n}}\right)$ to a Gaussian process, the Brownian Bridge, defined in Section 2.1. Section 2.2 provides the necessary conditions for the Central Limit Theorem for empirical processes to hold. Finally, in Section 2.3, we introduce the equivalent Central Limit Theorem for the empirical process, but instead of estimating a fixed distribution measure $\mathbb{P}$, the random samples are generated from a different probability measure for each n. Assuming we have a finite sample $\mathbf{X}_n^*$ generated by a probability measure R_n, we are interested in characterizing the asymptotic behavior of the empirical measure, as in the standard case. The difference is that the data is not assumed to be generated by a fixed probability measure $\mathbb{P}$, but for every n, a different probability measure R_n is used to simulate. This sequence of measures represents the Bootstrap scheme in each case, where for every n, R_n depends on the original sample $\mathbf{X}_n$.

All theorems and definitions of this chapter are derived from [10].

2.1. The Brownian Bridge

In this thesis, the process:

$$\mathbb{G}_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (\delta(X_i) - \mathbb{P})$$

is of primary interest. Its asymptotic behavior will demonstrate which bootstrap statistics are consistent and under which bootstrap scheme. We shall lay the conditions and framework for which $\mathbb{G}_n$ converges weakly to a suitable tight limit $\mathbb{G}$.

Let T be an arbitrary set. The space $\ell^\infty(T)$ is the set of all bounded, real-valued functions on T, denoted by $z : T \to \mathbb{R}$ such that $\|z\|_T := \sup_{t\in T} |z(t)| < \infty$. It is a metric space with respect to the uniform distance $d(z_1, z_2) = \|z_1 - z_2\|_T$. This space, or a suitable subspace of it, is a natural space for stochastic processes with bounded sample paths.

Definition 2.1.1 (Page 34 in [10]). *A stochastic process is an indexed collection $X(t) : t \in T$ of random variables defined on the same probability space, where each $X(t) : \Omega \to \mathbb{R}$ is a measurable map. If every sample path $t \mapsto X(t, \omega)$ is bounded, then the stochastic process yields a map $X : \Omega \to \ell^\infty(T)$.*

Definition 2.1.2 (Page 16 in [10]). *A Borel probability measure L on a space $\mathbb{D}$ is said to be tight if for every $c > 0$, there exists a compact set $K \subset \mathbb{D}$ such that $L(K) \geq 1 - c$.*

A Borel measurable map $X : \Omega \to \mathbb{D}$ is called tight if its law $\mathcal{L}(X) = \mathbb{P}(X \in A)$ for $A \in \mathbb{B}(\mathbb{D})$ is tight, where $\mathbb{P}$ denotes the probability measure on the space of probability measures on $\mathbb{B}(\mathbb{D})$.

With this setting, the following Lemma is crucial in the characterization of the Brownian Bridge, a pivotal notion in our research:

Lemma 2.1.3 (Lemma 1.5.3 in [10])**.** *Let X and Y be tight Borel measurable maps into $\ell^\infty(T)$. Then X and Y are equal in law if and only if all corresponding marginals of X and Y are equal in law.*

Definition 2.1.4. *Let X be a tight, Borel measurable mapping in $\ell^\infty(T)$, according to Definition 2.1.1. The mapping $(s,t) \mapsto \rho_p(s,t), where:$*

$$\rho_p(s,t) = (\mathbb{E}|X(s) - X(t)|^p)^{\frac{1}{p}} \tag{2.1}$$

is called the ρ_p semimetric.

Definition 2.1.5 (Example 1.5.10 in [10])**.** *A stochastic process X is called Gaussian if for every $n \in \mathbb{N}$ and $t_1, \dots, t_n \geq 0$ the marginals $(X(t_1), \dots, X(t_n))$ are normally distributed on $\mathbb{R}^n$.*

Remark 2.1.6. *Let $\mathcal{F}$ be a class of functions $f : \mathbb{D} \to \mathbb{E}$, where $\mathbb{D}$ is separable metric space. For a probability measure $\mathbb{P}$ and a measurable function f, we denote:*

$$\mathbb{P}f := \int f d\mathbb{P}.$$

Let:

$$\rho_{\mathbb{P}}(f) = (\mathbb{P}(f - \mathbb{P}f)^2)^{1/2},$$

denote a seminorm $\rho_{\mathbb{P}}$ on $\mathcal{F}$ with $p = 2$ from Equation (2.1)*. Throughout this project by $\rho_{\mathbb{P}}$ we refer to this semimetric.*

Let $X : \Omega \to \mathbb{D}$ be a random variable from $(\Omega, \mathcal{A}, \mathbb{P})$ to a separable metric space $\mathbb{D}$ and let $X_1, \dots, X_n$ be an i.i.d. sample from $\mathbb{P}$. Let also $\mathcal{F}$ be a collection of measurable functions $f : \mathbb{D} \to \mathbb{E}$. The empirical measure $\mathbb{P}_n = \frac{1}{n}\sum_{i=1}^n \delta(X_i)$ induces a mapping from $\mathcal{F}$ to $\mathbb{R}$: $f \mapsto \mathbb{P}_n f$. Note that:

$$\mathbb{P}_n f = \frac{1}{n}\sum_{i=1}^n \int f d\delta(X_i) = \frac{1}{n}\sum_{i=1}^n f(X_i).$$

For a function f such that $f(X) \in L^2(\mathbb{P})$, set

$$\mathbb{G}_n f := \sqrt{n}(\mathbb{P}_n - \mathbb{P})f.$$

From the Law of Large Numbers we get:

$$\mathbb{P}_n f = \frac{1}{n}\sum_{i=1}^n f(X_i) \xrightarrow{a.s.} \mathbb{E}(f(X_1)) = \mathbb{P}f.$$

Further, from the Central Limit Theorem:

$$\mathbb{G}_n f \xrightarrow{d} \mathcal{N}(0, \mathbb{P}(f - \mathbb{P}f)^2).$$

Indeed, we can see that

$$\mathbb{E}\left((\mathbb{P}_n - \mathbb{P})f\right) = \mathbb{E}(f(X_1)) - \mathbb{E}(f(X)) = 0$$

and

$$Var\left((\mathbb{P}_n - \mathbb{P})f\right) = Var\left(\mathbb{P}_n f\right) = \mathbb{P}(f - \mathbb{P}f)^2.$$

Thus we get that for a fixed function f, the process $\mathbb{G}_n$ applied to this function converges to a normal distribution. In the statistical setting, in order to obtain convergence for various statistical quantities it is desirable to expand this property to a large class of functions. The following definitions, describe the classes of functions that make the empirical process feasible to examine from a statistics point of view.

Definition 2.1.7 (Page 81 in [10]). *Denote the norm:*

$$\|\mathbb{P}\| = \sup_{f\in\mathcal{F}} |\mathbb{P}f|.$$

A class of functions $\mathcal{F}$ is called Glivenko-Cantelli if:

$$\|\mathbb{P}_n - \mathbb{P}\| = \sup_{f\in\mathcal{F}} |\mathbb{P}_n f - \mathbb{P}f| \xrightarrow{a.s.} 0.$$

Definition 2.1.8 (Page 81 in [10]). *Let $\mathcal{F}$ be a class of functions, such that for every x: $\sup_{f\in\mathcal{F}} |f(x) - \mathbb{P}f| < \infty$. Then $\mathcal{F}$ is called Donsker class (or $\mathbb{P}$-Donsker), if:*

$$\mathbb{G}_n = \sqrt{n}(\mathbb{P}_n - \mathbb{P}) \xrightarrow{d} \mathbb{G}_{\mathbb{P}} \text{ in } l^{\infty}(\mathcal{F}), \tag{2.2}$$

*for $\mathbb{G}_{\mathbb{P}}$ a **tight** Borel measurable element in $l^{\infty}(\mathcal{F})$.*

Definition 2.1.9. *Let $\mathcal{F}$ be a $\mathbb{P}$-Donsker Class. The limit process $\mathbb{G}_{\mathbb{P}}$ from Definition 2.1.8 is a Gaussian process with the following property:*

$$\mathbb{E}(\mathbb{G}_{\mathbb{P}} f_1) = 0,$$

and

$$\mathbb{E}(\mathbb{G}_{\mathbb{P}} f_1 \mathbb{G}_{\mathbb{P}} f_2) = \mathbb{P}(f_1 - \mathbb{P}f_1)(f_2 - \mathbb{P}f_2) = \mathbb{P}f_1 f_2 - \mathbb{P}f_1 \mathbb{P}f_2, \tag{2.3}$$

for arbitrary functions $f_1, f_2 \in \mathcal{F}$. This process is called the Brownian Bridge.

Remark 2.1.10. *By Lemma 2.1.3 and the expression above, tightness and the covariance function completely determine the distribution of the process in the definition. Moreover we get the uniqueness of the Brownian Bridge. For probability measures $\mathbb{P}, \mathbb{Q}$, $\mathbb{P} = \mathbb{Q}$ implies that $\mathbb{G}_{\mathbb{P}} \overset{d}{=} \mathbb{G}_{\mathbb{Q}}$.*

2.2. VC classes and Entropy Conditions

We proceed with presenting the notions that we use throughout this project in order to assert that we work on Donsker classes of functions. The first one is the Uniform Entropy Bound which implies that the underlying class is Donsker. The second is the VC-dimension of a class which, when finite, implies that it satisfies the Uniform Entropy Bound. Let $(\mathcal{F}, \|\cdot\|)$ be a normed space of real functions $f : \mathbb{D} \to \mathbb{R}$.

Definition 2.2.1 (Covering numbers). *The covering number $N(\epsilon, \mathcal{F}, \|\cdot\|)$ is the minimal number of balls $\{g : \|g - f\| < \epsilon\}$ of radius ϵ needed to cover the set $\mathcal{F}$. The entropy is the logarithm of the covering number.*

Condition 2 (Uniform Entropy). *[2.1.7 in [10]] Let F be a square-integrable function. Denote $\|F\|_{\mathbb{P},2}$ the L_2-norm of F with respect to a probability measure $\mathbb{P}$. Then the uniform entropy condition is satisfied if the uniform entropy bound is finite:*

$$\int_0^{\infty} \sup_Q \sqrt{\log N\left(\epsilon\|F\|_{Q,2}, F, L_2(Q)\right)} d\epsilon < \infty$$

The supremum is taken over all finite discrete probability measures Q on $(\Omega, \mathcal{A})$ such that $\|F\|_{Q,2}^2 = \int F^2 dQ > 0$.

Definition 2.2.2 (VC Classes of Sets). *Let $\mathcal{C}$ be a collection of subsets of a set $\mathcal{X}$. An arbitrary subset of n points $\{x_1, \dots, x_n\}$ from $\mathcal{X}$ possesses 2^n subsets. If a subset of $\{x_1, \dots, x_n\}$ can be formed as $C \cap \{x_1, \dots, x_n\}$ for $C \in \mathcal{C}$, we say that $\mathcal{C}$ picks this subset. The family $\mathcal{C}$ shatters $\{x_1, \dots, x_n\}$ if each of its 2^n subsets can be picked. The VC-index $V(\mathcal{C})$ of the class $\mathcal{C}$ is the smallest n for which no set of size n is shattered by $\mathcal{C}$. Formally, let:*

$$\Delta_n(\mathcal{C}, x_1, \dots, x_n) = \#\{C \cap \{x_1, \dots, x_n\} : C \in \mathcal{C}\},$$

then the $V(\mathcal{C})$ index is defined as:

$$V(\mathcal{C}) = \inf\{n : \max_{x_1,\dots,x_n \in \mathcal{X}} \Delta_n(\mathcal{C}, x_1, \dots, x_n) < 2^n\}.$$

A collection of measurable sets $\mathcal{C}$ is called VC-class if $V(\mathcal{C}) < \infty$.

The following theorem displays a direct connection between covering numbers and the VC index of a class of sets. For a class of sets $\mathcal{C}$, a bound for its covering numbers is provided, that holds uniformly over any probability measure defined on this class.

Theorem 2.2.3 (Theorem 2.6.4 in [10])**.** *There exists a universal constant K such that for any VC-class $\mathcal{C}$ of sets, any probability measure $\mathbb{P}$, $r \geq 1$ and $\epsilon \in (0,1)$:*

$$N(\epsilon, \mathcal{C}, L_r(Q)) \leq KV(\mathcal{C})4^{V(\mathcal{C})} \cdot \epsilon^{V(\mathcal{C})-r(V(\mathcal{C})-1)}.$$

Definition 2.2.4. *For a function $f : \mathbb{D} \to \mathbb{R}$, the subgraph of f is a subset of $\mathbb{D} \times \mathbb{R}$ given by $\{(x,t) : t < f(x)\}$. A collection $\mathcal{F}$ of measurable functions on a sample space is called VC-subgraph class if the collection of all subgraphs of the functions in $\mathcal{F}$ forms a VC-class of sets in $\mathbb{D} \times \mathbb{R}$.*

Denote by $V(\mathcal{F})$ the VC-index of the set of subgraphs of $\mathcal{F}$.

Theorem 2.2.5. *For a VC-class of functions with measurable envelope function F and $r \geq 1$, for any probability measure Q with $\|F\|_{Q,r} > 0$, there exists a universal constant K such that for $\epsilon \in (0,1)$ the following holds:*

$$N(\epsilon\|F\|_{Q,r}, \mathcal{F}, L_r(Q)) \leq KV(\mathcal{F})16^{V(\mathcal{F})}\epsilon^{V(\mathcal{F})(1-r)+r}$$

Definition 2.2.6 (Envelope Function)**.** *Let $\mathcal{F}$ be a class of measurable functions from a separable metric space $\mathbb{D}$ to the real numbers. A function $F : \mathbb{D} \to \mathbb{R}$ is called the Envelope Function of $\mathcal{F}$, if for every $f \in \mathcal{F}$ and for every $x \in \mathbb{D}$: $|f(x)| \leq F(x)$. The mapping $x \mapsto \sup_{f\in\mathcal{F}} f(x)$ is called the Minimal Envelope Function.*

In the rest of this text, when we refer to the envelope function of a class, we mean the Minimal Envelope Function and we assume that it is a.s. finite.

Theorem 2.2.7 (Theorem 2.5.2 in [10])**.** *Let $\mathcal{F}$ be a class of measurable functions, with envelope function F, that satisfies the Uniform Entropy Condition 2. Let the classes $\mathcal{F}_{\delta,\mathbb{P}} = \{f - g : f, g \in \mathcal{F}, \|f - g\|_{\mathbb{P},2} < \delta\}$ and $\mathcal{F}_\infty^2 := \{(f - g)^2 : f, g \in \mathcal{F}\}$ are $\mathbb{P}$-measurable for every $\delta > 0$. If $\mathbb{P}F < \infty$ then $\mathcal{F}$ is $\mathbb{P}$-Donsker.*

Remark 2.2.8. *Notice that the right side bound of the Covering numbers for a probability measure Q from Theorem 2.2.5 does not depend to the probability measure itself. Hence, the preceding theorems imply that a VC-class of functions satisfies the Uniform Entropy Condition. Thus if it also satisfies certain measurability conditions the class is P-Donsker for any probability measure P for which the envelope function is square integrable.*

Remark 2.2.9. *The collection of cells $(-\infty, t]$ for $t \in \mathbb{R}$ is a VC-class with $V(\mathcal{C}) = 2$. Indeed, every single point set is shattered but no two point set is. Further, if a collections of sets $\mathcal{C}$ is a VC-class, then the collection of indicators of sets in $\mathcal{C}$ is a VC-subgraph calls of the same index. Hence the collection of indicators of the form $1_{(-\infty,t]}$ is a VC-class and satisfies the Uniform Entropy Condition and is P-Donsker. This property of the class of indicators on the real line will be of great importance in Chapter 3 where the general proofs for the consistency of the statistics of empirical measures will be specified for the equivalent statistics of empirical distribution functions.*

2.3. CLT under Sequences

Consider the triangular array scheme where for each $n \in \mathbb{N}$, $X_1^*, \ldots, X_n^* \overset{i.i.d.}{\sim} R_n$ for sequence of probability measures R_n. Denote the empirical measure:

$$\hat{\mathbb{H}}_n = \sum_{i=1}^{n} \delta(X_i^*).$$

The interest now shifts to the convergence of the process:

$$\mathbb{G}_{n,R_n} = \sqrt{n}(\hat{\mathbb{H}}_n - R_n).$$

The sequence R_n can be imagined as the bootstrap measure, that describes the probability distribution of a bootstrap sample $X_1^*, \dots, X_n^*$ conditional on a finite sample $X_1, \dots, X_n$. Assuming that the sequence R_n conditional on $\mathbf{X}_n$ converges to a measure R we shall display the framework for which:

$$\mathbb{G}_{n,R_n}|\mathbf{X}_n \xrightarrow{d} \mathbb{G}_R \text{ in } \ell^\infty(\mathcal{F}).$$

The following Theorem is the central result in proving that a Bootstrap scheme is meaningful and examining the asymptotic behavior of the empirical distribution for samples generated under a resampling scheme R_n. Although the theorem provides a general convergence result, the convergence that is displayed in this project is conditional on a finite sample $\mathbf{X}_n$, since the resampling schemes are conditional.

Theorem 2.3.1 (Theorem 2.8.9 in [10])**.** *Let $\mathcal{F}$ a class of measurable functions, with measurable envelope function F and denote:*

$$\mathcal{F}_{\delta,\mathbb{P}} = \{f - g : f, g \in \mathcal{F}, \rho_{\mathbb{P}}(f - g) < \delta\}.$$

Assume that $\mathcal{F}_{\delta,R_n} = \{f - g : f, g \in \mathcal{F}, \|f - g\|_{R_n,2} < \delta\}$ and $\mathcal{F}_\infty^2 := \{(f - g)^2 : f, g \in \mathcal{F}\}$ are R_n-measurable for every $\delta > 0$ and n. Suppose that $\mathcal{F}$ satisfies the Uniform Entropy Condition 2 and $R_n F^2 = O(1)$. Finally assume that the following conditions are satisfied:

1. *ρ_{R_n} seminorms converge uniformly to ρ_{P_0}:*

$$\sup_{f \in \mathcal{F}} |\rho_{R_n}(f) - \rho_R(f)| \xrightarrow{a.s.} 0. \tag{2.4}$$

2. *For every $\epsilon > 0$:*

$$\limsup_{n\to\infty} R_n F^2 1(F \geq \epsilon\sqrt{n}) = 0 \ a.s.. \tag{2.5}$$

Then:

$$\mathbb{G}_{n,R_n} \xrightarrow{d} \mathbb{G}_R \textit{ in } \ell^\infty(\mathcal{F}),$$

where $\mathbb{G}_R$ is a R-Brownian Bridge.

3

Independence Test

In this section we state and prove the special case of our claim in the independence test. We proceed by stating the formal independence test setting, then we provide all the necessary theoretical tools for the proof and we conclude with the complete proof.

Let $\mathcal{P}$ be the set of all bivariate distributions and $\mathcal{P}_0$ be the set of independent bivariate distributions. A bivariate random sample $(X_1, Y_1), \dots, (X_n, Y_n)$ is obtained from the joint distribution $\mathbb{H}$. We denote $\mathbb{P}, \mathbb{Q}$ the marginal distributions of X, Y, namely $X \sim \mathbb{P}, Y \sim \mathbb{Q}$. We are interested in testing whether the random variables X, Y are independent. This translates to the joint distribution of the joint $\mathbb{H}$ being equal to the product of the marginal distributions $\mathbb{P}, \mathbb{Q}$. Thus, in this framework the test is stated:

$$H_0 : \mathbb{H} \in \mathcal{P}_0 \text{ vs } H_1 : \mathbb{H} \in \mathcal{P} \backslash \mathcal{P}_0.$$

For a finite random sample, set the empirical measures:

$$\mathbb{P}_n = \frac{1}{n}\sum_{i=1}^{n} \delta(X_i), \quad \mathbb{Q}_n = \frac{1}{n}\sum_{i=1}^{n} \delta(Y_i), \quad \text{and } \mathbb{H}_n = \frac{1}{n}\sum_{i=1}^{n} \delta(X_i, Y_i). \tag{3.1}$$

We also set the following measure product: $\mathbb{P}_n \otimes \mathbb{Q}_n = \frac{1}{n^2}\sum_{i,j=1}^{n} \delta(X_i, Y_j)$.

The processes defined above will completely define the bootstrap schemes, test statistics and their asymptotic behavior. It is of high importance to highlight the difference between $\mathbb{H}_n$ and $\mathbb{P}_n \otimes \mathbb{Q}_n$. We can see that while the process $\mathbb{H}_n$ is defined so that it estimates the joint distribution of $(X_i, Y_i)_{i=1}^{n}$, the process $\mathbb{P}_n \otimes \mathbb{Q}_n$ differs. When used as resampling schemes, $\mathbb{H}_n$ represents a simulation from the joint distribution, while $\mathbb{P}_n \otimes \mathbb{Q}_n$ represents independent resampling from the marginals and a generation of a joint sample by concatenating the marginal simulations.

Given a resampling scheme a Bootstrap sample $(X_i^*, Y_i^*)_{i=1}^{n}$ is generated and we denote:

$$\hat{\mathbb{P}}_n = \frac{1}{n}\sum_{i=1}^{n} \delta(X_i^*), \quad \hat{\mathbb{Q}}_n = \frac{1}{n}\sum_{i=1}^{n} \delta(Y_i^*), \quad \text{and } \hat{\mathbb{H}}_n = \frac{1}{n}\sum_{i=1}^{n} \delta(X_i^*, Y_i^*).$$

Remark 1.5.4 implies that the empirical distribution can be seen as a special case of the empirical measure, for instance: $\hat{F}(s,t) = \mathbb{H}_n((-\infty, s] \times (-\infty, t]), \hat{F}^*(s,t) = \hat{\mathbb{H}}_n((-\infty, s] \times (-\infty, t])$.

In the next Chapter we provide a concrete characterization of the independence test setting, two options for resampling and describe the true statistic T_n, as well as the Bootstrap statistics T_n^{*eq}, T_n^{*c}.

3.1. Setting

Let X, Y be two random variables defined on measurable spaces $(\mathcal{X}, \mathcal{A}), (\mathcal{Y}, \mathcal{B})$ respectively, such that $X \sim \mathbb{P}$, $Y \sim \mathbb{Q}$, and let the joint random variable (X, Y) defined on the product space $(\mathcal{X} \times \mathcal{Y}, \mathcal{A} \times \mathcal{B})$

such that $(X,Y) \sim \mathbb{H}$. Let $\mathcal{F}, \mathcal{G}$ be classes of functions defined on the spaces $\mathcal{X}, \mathcal{Y}$ respectively and let $\mathcal{F} \times \mathcal{G}$ be the class of functions $f \times g$ such that:

$$\mathcal{F} \times \mathcal{G} = \{f \times g : (f \times g)(x,y) = f(x)g(y)\}.$$

In our setting, testing independence is formalized as the following hypothesis test of equality between probability measures.

$$H_0 : \mathbb{H} = \mathbb{P} \otimes \mathbb{Q} \text{ vs } H_1 : \mathbb{H} \neq \mathbb{P} \otimes \mathbb{Q}.$$

Since there is a one-to-one mapping between probability measures and cumulative distribution functions, for random variables in $\mathbb{R}^d$ the independence test can be equivalently rewritten as the test of:

$$H_0 : F = F_X \times F_Y \text{ vs } H_0 : F \neq F_X \times F_Y,$$

where F, F_X, F_Y are respectively the cumulative distribution function of $\mathbb{H}, \mathbb{P}, \mathbb{Q}$.

Set the true and bootstrap statistics:

$$\begin{aligned} T_n &= \left\|\sqrt{n}(\hat{F} - \hat{F}_X \times \hat{F}_Y)\right\|, \\ T_n^{*eq} &= \left\|\sqrt{n}(\hat{F}^* - \hat{F}_X^* \times \hat{F}_Y^*)\right\|, \\ T_n^{*c} &= \left\|\sqrt{n}(\hat{F}^* - \hat{F} + \hat{F}_X \times \hat{F}_Y - \hat{F}_X^* \times \hat{F}_Y^*)\right\|. \end{aligned}$$

For a finite sample $\mathbf{X}_n, \mathbf{Y}_n$ from $\mathbb{H}$, denote $\mathbf{U}_n = (\mathbf{X}_n, \mathbf{Y}_n)$. The objective is to show that, when the null hypothesis holds, under GR Bootstrap:

$$d_{BL}\left(T_n\,,\, T_n^{*c}|\mathbf{U}_n\right) \xrightarrow{\mathbb{H}} 0,$$

and under NHR Bootstrap:

$$d_{BL}\left(T_n\,,\, T_n^{*eq}|\mathbf{U}_n\right) \xrightarrow{\mathbb{H}} 0,$$

so that the equivalent condition (4) from Lemma 1.5.5 holds and the respective statistics are consistent with the true statistic under each resampling scheme.

3.2. Main results

Following the discussion in the previous Section, the combinations we pick to apply Lemma 1.5.5 are $(T_n^{*eq}, \mathrm{NHR})$ and (T_n^{*c}, GR).

For a Bootstrap statistic T_n^*, in the independence test we define the power of the test as:

$$\mathbb{H} \mapsto \pi(\mathbb{H}; T_n, T_n^*) := \mathbb{H}(T_n^* < T_n|\mathbf{U}_n).$$

Theorem 3.2.1. *Under the null hypothesis $\mathbb{H} \in \mathcal{P}_0$ or equivalently $\mathbb{H} = \mathbb{P} \otimes \mathbb{Q}$, for the expressions of the Bootstrap statistics we have the following:*

1. *For almost every sequence $(X_1,Y_1), \ldots, (X_n,Y_n)$ and $(X_1^*,Y_1^*), \ldots, (X_n^*,Y_n^*)$ distributed according to the probability measure $P_n = \mathbb{H}_n$:*
$$d_{BL}\left(T_n\,,\, T_n^{*c}|\mathbf{U}_n\right) \xrightarrow{\mathbb{H}} 0.$$
2. *For almost every sequence $(X_1,Y_1), \ldots, (X_n,Y_n)$ and $(X_1^*,Y_1^*), \ldots, (X_n^*,Y_n^*)$ distributed according to the probability measure $P_n = \mathbb{P}_n \otimes \mathbb{Q}_n$:*
$$d_{BL}\left(T_n\,,\, T_n^{*eq}|\mathbf{U}_n\right) \xrightarrow{\mathbb{H}} 0.$$

Theorem 3.2.2. *Under the null hypothesis the following hold:*

1. *For a GR Bootstrap Scheme (generate Bootstrap samples from $\mathbb{H}_n$):*
$$pvalue_n = \mathbb{H}\left(T_n^{*c} > T_n|\mathbf{U}_n\right) \xrightarrow{d} U(0,1).$$

2. *For a NHR Bootstrap Scheme (generate Bootstrap samples from $\mathbb{P}_n \otimes \mathbb{Q}_n$):*

$$pvalue_n = \mathbb{H}\left(T_n^{*eq} > T_n|\mathbf{U}_n\right) \xrightarrow{d} U(0,1).$$

Theorem 3.2.3. *Under the alternative hypothesis the following hold:*

1. *For a Bootstrap Scheme of GR (generate Bootstrap samples from $\mathbb{H}_n$):*

$$\mathbb{H}\left(T_n^{*c} < T_n|\mathbf{U}_n\right) \xrightarrow{d} 1.$$

2. *For a Bootstrap Scheme of NHR (generate Bootstrap samples from $\mathbb{P}_n \otimes \mathbb{Q}_n$):*

$$\mathbb{H}\left(T_n^{*eq} < T_n|\mathbf{U}_n\right) \xrightarrow{d} 1.$$

Theorem 3.2.4. *In the independence test of significance α, the following hold.*

- *Under GR Bootstrap, $R_n = \mathbb{H}_n$, for the Bootstrap statistic T_n^{*c}:*

$$\text{for all } \mathbb{H} \in \mathcal{P}_0,\ \lim_{n\to\infty} \pi(\mathbb{H}; T_n, T_n^{*c}) = \alpha,$$

  and

$$\text{for all } \mathbb{H} \in \mathcal{P}\backslash\mathcal{P}_0,\ \lim_{n\to\infty} \pi(\mathbb{H}; T_n, T_n^{*c}) = 1,$$

- *Under NHR Bootstrap, $R_n = \mathbb{P}_n \otimes \mathbb{Q}_n$, for the Bootstrap statistic T_n^{*eq}:*

$$\text{for all } \mathbb{H} \in \mathcal{P}_0,\ \lim_{n\to\infty} \pi(\mathbb{H}; T_n, T_n^{*eq}) = \alpha,$$

  and

$$\text{for all } \mathbb{H} \in \mathcal{P}\backslash\mathcal{P}_0,\ \lim_{n\to\infty} \pi(\mathbb{H}; T_n, T_n^{*eq}) = 1,$$

3.2.1. Inconsistent Combinations

In this section, we show that the combinations $(T_n^{*eq}, \text{GR })$, (T_n^{*c}, NHR), do not lead to hypothesis tests with good power. To achieve that, it suffices to show that under the null hypothesis the respective bootstrap statistics are not consistent with the true statistic T_n, hence it is impossible to control the Type I error.

Theorem 3.2.5. *Then under the null hypothesis the following hold:*

$$d_{BL}\left(T_n\,,\,T_n^{*eq}|\mathbf{U}_n\right) \nrightarrow 0 \text{ under GR Bootstrap}, R_n = \mathbb{H}_n,$$
$$d_{BL}\left(T_n\,,\,T_n^{*c}|\mathbf{U}_n\right) \nrightarrow 0 \text{ under NHR Bootstrap}, R_n = \mathbb{P}_n \otimes \mathbb{Q}_n.$$

3.3. Empirical process point-of-view and proofs of the results

We define the following processes:

$$\mathbb{Z}_n = \sqrt{n}\left[(\mathbb{H}_n - \mathbb{P}_n \otimes \mathbb{Q}_n) - (\mathbb{H} - \mathbb{P} \otimes \mathbb{Q})\right],$$

and conditional on a finite sample $\mathbf{U}_n$:

$$\hat{\mathbb{S}}_n = \sqrt{n}\left[(\hat{\mathbb{H}}_n - \hat{\mathbb{P}}_n \otimes \hat{\mathbb{Q}}_n) - (\mathbb{H}_n - \mathbb{P}_n \otimes \mathbb{Q}_n)\right], \tag{3.2}$$

$$\hat{\mathbb{T}}_n = \sqrt{n}\left(\hat{\mathbb{H}}_n - \hat{\mathbb{P}}_n \otimes \hat{\mathbb{Q}}_n\right). \tag{3.3}$$

Under the null hypothesis $\mathbb{H} = \mathbb{P} \otimes \mathbb{Q}$ and by applying the product indicator $1(X \leq s) \otimes 1(Y \leq t)$ we get the test statistics T_n, T_n^{*c}, T_n^{*eq} respectively, as special cases of the norms of the previous processes.

Theorems 3.3.3 and 3.3.2, imply the consistency of the empirical measure statistics:

$$\hat{\mathbb{S}}_n = \sqrt{n}\left[(\hat{\mathbb{H}}_n - \hat{\mathbb{P}}_n \otimes \hat{\mathbb{Q}}_n) - (\mathbb{H}_n - \mathbb{P}_n \otimes \mathbb{Q}_n)\right] \qquad \text{under GR Bootstrap}, R_n = \mathbb{H}_n,$$
$$\hat{\mathbb{T}}_n = \sqrt{n}\left(\hat{\mathbb{H}}_n - \hat{\mathbb{P}}_n \otimes \hat{\mathbb{Q}}_n\right) \qquad \text{under NHR Bootstrap}, R_n = \mathbb{P}_n \otimes \mathbb{Q}_n,$$

with respect to the true empirical measure statistic:

$$\mathbb{Z}_n = \sqrt{n}\left[(\mathbb{H}_n - \mathbb{P}_n \otimes \mathbb{Q}_n) - (\mathbb{H} - \mathbb{P} \otimes \mathbb{Q})\right].$$

The following theorem from [10] describes the asymptotic behavior of $\mathbb{Z}_n$, and equivalently the distribution of the statistic T_n, as a special case of the former.

Theorem 3.3.1 (Theorem 3.8.1 in [10])**.** *Let $\mathcal{F},\mathcal{G}$ be classes of measurable functions on measurable spaces $(X,\mathcal{A})$ and $(Y,\mathcal{B})$ respectively. If $\mathcal{F}\times\mathcal{G},\mathcal{F}$, and $\mathcal{G}$ are $\mathbb{H}$-Donsker and $\|\mathbb{P}\|_{\mathcal{F}} < \infty$ and $\|\mathbb{Q}\|_{\mathcal{G}} < \infty$, then the sequence of independence processes $\mathbb{Z}_n$ conditionally on a finite sample $\mathbf{U}_n$ converges in distribution in $l^\infty(\mathcal{F}\times\mathcal{G})$ to the Gaussian process:*

$$\mathbb{Z}_{\mathbb{H}}(f\times g) = \mathbb{G}_{\mathbb{H}}((f-\mathbb{P}f)\times(g-\mathbb{Q}g)),$$

for a tight $\mathbb{H}$-Brownian bridge $\mathbb{G}_{\mathbb{H}}$.

3.3.1. NHR

We proceed by presenting the case for the NHR Bootstrap, i.e. resampling from the null hypothesis. Specifically, the Bootstrap new sample is generated from the empirical product $\mathbb{P}_n \otimes \mathbb{Q}_n$. In practice, what this means is that two bootstrap samples $(X_i^*)_{i=1}^n \sim \mathbb{P}_n$, $(Y_i^*)_{i=1}^n \sim \mathbb{Q}_n$ independently and concatenate to results to create the joint sample $(X_i^*, Y_i^*)_{i=1}^n \sim \mathbb{P}_n \otimes \mathbb{Q}_n$.

Theorem 3.3.2 (Theorem 3.8.3 in [10])**.** *Let $\mathcal{F}$ and $\mathcal{G}$ be separable classes of measurable functions on measurable spaces $(X,\mathcal{A})$ and $(Y,\mathcal{B})$ respectively such that $\mathcal{F}\times\mathcal{G}$ satisfies the Uniform Entropy Condition 2 for envelope functions F,G and $F\times G$ are $\mathbb{H}$-square integrable. Given a sequence $\mathbf{U}_n = (\mathbf{X}_n,\mathbf{Y}_n)$, let $(X_1^*,Y_1^*),...,(X_n^*,Y_n^*)$ generated by the probability measures $P_n = \mathbb{P}_n\otimes\mathbb{Q}_n$. Then:*

$$\hat{\mathbb{G}}_n = \sqrt{n}(\hat{\mathbb{H}}_n - \mathbb{P}_n\otimes\mathbb{Q}_n) \xrightarrow{d} \mathbb{G}_{\mathbb{P}\otimes\mathbb{Q}}, \text{ in } l^\infty(\mathcal{F}\times\mathcal{G}), \text{ conditionally on } \mathbf{U}_n,$$

$$\hat{\mathbb{T}}_n = \sqrt{n}[(\hat{\mathbb{H}}_n - \hat{\mathbb{P}}_n\otimes\hat{\mathbb{Q}}_n)] \xrightarrow{d} \mathbb{Z}_{\mathbb{P}\otimes\mathbb{Q}}, \text{ in } l^\infty(\mathcal{F}\times\mathcal{G}), \text{ conditionally on } \mathbf{U}_n,$$

given H^∞-almost every sequence $(X_1,Y_1),(X_2,Y_2)...$, where $\mathbb{G}_{\mathbb{P}\otimes\mathbb{Q}}$ is a $\mathbb{P}\otimes\mathbb{Q}$-Brownian bridge.

Proof. We can rewrite the bootstrap independence process as follows:

$$\begin{aligned}\hat{\mathbb{T}}_n(f\times g) = {}& \sqrt{n}(\hat{\mathbb{H}}_n - \mathbb{P}_n\otimes\mathbb{Q}_n)(f\times g) - \sqrt{n}(\hat{\mathbb{P}}_n - \mathbb{P}_n)f\mathbb{Q}_n g,\\ & - \mathbb{P}_n f\sqrt{n}(\hat{\mathbb{Q}}_n - \mathbb{Q}_n)g - \sqrt{n}(\hat{\mathbb{P}}_n - \mathbb{P}_n)f(\hat{\mathbb{Q}}_n - \mathbb{Q}_n)g.\end{aligned}$$

First, we show that conditionally on every sequence of original observations $\mathbf{U}_n$, the following hold:

$$(\mathbb{P}_n\otimes\mathbb{Q}_n)(F\times G)^2 = O(1), \tag{3.4}$$
$$(\mathbb{P}_n\otimes\mathbb{Q}_n)(F\times G)^2 1_{|F\times G|\geq\epsilon\sqrt{n}} \to 0, \text{ for every } \epsilon > 0, \tag{3.5}$$
$$\sup_{h_1,h_2\in\mathcal{F}\times\mathcal{G}} |\rho_{\mathbb{P}_n\otimes\mathbb{Q}_n}(h_1,h_2) - \rho_{\mathbb{P}\otimes\mathbb{Q}}(h_1,h_2)| \to 0. \tag{3.6}$$

Equation 3.4 holds since the envelope $F\times G$ is $\mathbb{H}$ square integrable. Then equation (3.5) can be written as:

$$\frac{1}{n^2}\sum_{i,j=1}^n F^2(X_i)G^2(Y_j)1(|f\times G| \geq \epsilon\sqrt{n}),$$

and for $M^2 \leq \epsilon\sqrt{n}$ this quantity is bounded by above by:

$$\mathbb{P}_n F^2 1_{F\geq M}\mathbb{Q}_n G^2 + \mathbb{P}_n F^2\mathbb{Q}_n G^2 1_{G\geq M}, \tag{3.7}$$

which converges almost surely to a fixed value, which can be made arbitrarily small by choosing M large. In fact

$$\mathbb{P}_n F^2 1_{F\geq M}\mathbb{Q}_n G^2 + \mathbb{P}_n F^2 \mathbb{Q}_n G^2 1_{G\geq M} = \frac{1}{n^2}\sum_{i,j=1}^{n} F(X_i)G(Y_j)1_{[F(X_i)\geq M]\cup[G(X_j)\geq M]},$$

and since $\{F(X_i)G(Y_j) \geq \epsilon\sqrt{n}\} \subset \{F(X_i) \geq M\} \cup \{G(X_j) \geq M\}$ for $M^2 \leq \epsilon\sqrt{n}$, the upper bound follows. This argument shows the validity of Condition 2.5 in this setting.

Finally:

$$\begin{aligned}\rho_{\mathbb{P}_n\otimes\mathbb{Q}_n} &= (\mathbb{P}_n \otimes \mathbb{Q}_n)(f \times g - (\mathbb{P}_n \otimes \mathbb{Q}_n)(f \times g))^2 \\ &= \mathbb{P}_n f^2 \cdot \mathbb{Q}_n g^2 - (\mathbb{P}_n f)^2(\mathbb{Q}_n g)^2,\end{aligned}$$

and:

$$\rho_{\mathbb{P}\otimes\mathbb{Q}} = \mathbb{P}f^2 \cdot \mathbb{Q}g^2 - (\mathbb{P}f)^2(\mathbb{Q}g)^2.$$

By substracting the previous quantities and taking absolute values and supremum over $f \times g \in \mathcal{F} \times \mathcal{G}$ we show that Equation 3.6 is satisfied, which implies that Condition 2.4 holds.

Hence we can apply Theorem 2.3.1, for $P_n = \mathbb{P}_n \otimes \mathbb{Q}_n$ and we obtain the convergence of $\hat{\mathbb{G}}_n$.

We have:

$$\begin{aligned}(\hat{\mathbb{H}}_n - \mathbb{P}_n \otimes \mathbb{Q}_n)(f \times g) &= \int f \times g d(\hat{\mathbb{H}}_n - \mathbb{P}_n \otimes \mathbb{Q}_n) \\ &= \int f \times g d\hat{\mathbb{H}}_n - \int f \times g d\mathbb{P}_n \times \mathbb{Q}_n.\end{aligned}$$

But:

$$\begin{aligned}\int f \times g d\hat{\mathbb{H}}_n &= \frac{1}{n}\sum_{i=1}^{n}\int f \times g d\delta(X_i^*, Y_i^*) \\ &= \frac{1}{n}\sum_{i=1}^{n} f(X_i^*)g(Y_i^*),\end{aligned}$$

and:

$$\int f \times g d\mathbb{P}_n \otimes \mathbb{Q}_n = \frac{1}{n^2}\sum_{i,j=1}^{n} f(X_i)g(Y_j).$$

By setting $g = 1$, in the previous we get:

$$\begin{aligned}\int f \times 1 d\hat{\mathbb{H}}_n &= \frac{1}{n}\sum_{i=1}^{n} f(X_i^*) = \hat{\mathbb{P}}_n f, \\ \int f \times 1 d\mathbb{P}_n \otimes \mathbb{Q}_n &= \frac{1}{n}\sum_{i=1}^{n} f(X_i) = \mathbb{P}_n f, \\ \sqrt{n}(\hat{\mathbb{H}}_n - \mathbb{P}_n \otimes \mathbb{Q}_n)(f \times 1) &= \sqrt{n}(\hat{\mathbb{P}}_n - \mathbb{P}_n) \xrightarrow{d} \mathbb{G}_{\mathbb{P}\otimes\mathbb{Q}}(f \times 1),\end{aligned}$$

and by setting $f = 1$ we get:

$$\int 1 \times g d\hat{\mathbb{H}}_n = \frac{1}{n}\sum_{i=1}^{n} g(Y_i^*) = \hat{\mathbb{Q}}_n g,$$

$$\int 1 \times g d\mathbb{P}_n \otimes \mathbb{Q}_n = \frac{1}{n}\sum_{i=1}^{n} g(Y_i) = \mathbb{Q}_n g,$$

$$\sqrt{n}(\hat{\mathbb{H}}_n - \mathbb{P}_n \otimes \mathbb{Q}_n)(1 \times g) = \sqrt{n}(\hat{\mathbb{Q}}_n - \mathbb{Q}_n) \xrightarrow{d} \mathbb{G}_{\mathbb{P}\otimes\mathbb{Q}}(1 \times g),$$

and by making use of the Continuous Mapping Theorem and Slutsky's Lemma, we get that the independence process:

$$\hat{\mathbb{T}}_n(f \times g) \xrightarrow{d} \mathbb{G}_{\mathbb{P}\otimes\mathbb{Q}}(f \times g) - \mathbb{G}_{\mathbb{P}\otimes\mathbb{Q}}(f \times 1)\mathbb{Q}g - \mathbb{P}f\mathbb{G}_{\mathbb{P}\otimes\mathbb{Q}}(1 \times g),$$

which is a Gaussian process with the same covariance as $\mathbb{Z}_{\mathbb{P}\otimes\mathbb{Q}}$.

□

3.3.2. GR

Theorem 3.3.3. *Let $\mathcal{F}$ and $\mathcal{G}$ be separable classes of measurable functions on measurable spaces $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ respectively such that $\mathcal{F} \times \mathcal{G}$ satisfies the Uniform Entropy Condition 2 for envelope functions F, G and $F \times G$, that are $\mathbb{H}$-square integrable. Given a sequence $\mathbf{U}_n = (\mathbf{X}_n, \mathbf{Y}_n)$, let $(X_1^*, Y_1^*), ..., (X_n^*, Y_n^*)$ distributed according to the probability measure $R_n = \mathbb{H}_n$. Then:*

$$\hat{\mathbb{G}}_n = \sqrt{n}(\hat{\mathbb{H}}_n - \mathbb{H}_n) \xrightarrow{d} \mathbb{G}_{\mathbb{H}}, \text{ in } l^\infty(\mathcal{F} \times \mathcal{G}), \text{ conditionally on } \mathbf{U}_n,$$

$$\hat{\mathbb{S}}_n = \sqrt{n}[(\hat{\mathbb{H}}_n - \hat{\mathbb{P}}_n \otimes \hat{\mathbb{Q}}_n) - (\mathbb{H}_n - \mathbb{P}_n \otimes \mathbb{Q}_n)] \xrightarrow{d} \mathbb{Z}_{\mathbb{H}}, \text{ in } l^\infty(\mathcal{F} \times \mathcal{G}), \text{ conditionally on } \mathbf{U}_n,$$

given H^∞-almost every sequence $(X_1, Y_1), (X_2, Y_2)...$, where $\mathbb{G}_{\mathbb{H}}$ is an $\mathbb{H}$-Brownian bridge.

Proof. We can rewrite the bootstrap independence process as follows:

$$\begin{aligned}\hat{\mathbb{S}}_n(f \times g) = \sqrt{n}(\hat{\mathbb{H}}_n - \mathbb{H}_n)(f \times g) - \sqrt{n}(\hat{\mathbb{P}}_n - \mathbb{P}_n)f\mathbb{Q}_n g \\ - \mathbb{P}_n f\sqrt{n}(\hat{\mathbb{Q}}_n - \mathbb{Q}_n)g - \sqrt{n}(\hat{\mathbb{P}}_n - \mathbb{P}_n)f(\hat{\mathbb{Q}}_n - \mathbb{Q}_n)g.\end{aligned}$$

First, we show that conditionally on every sequence of original observations $\mathbf{U}_n$, the following hold:

$$(\mathbb{H}_n)(F \times G)^2 = O(1), \tag{3.8}$$

$$\mathbb{H}_n(F \times G)^2 1_{|F\times G| \geq \epsilon\sqrt{n}} \to 0, \text{ for every } \epsilon > 0, \tag{3.9}$$

$$\sup_{h_1, h_2 \in \mathcal{F}\times\mathcal{G}} |\rho_{\mathbb{H}_n}(h_1, h_2) - \rho_{\mathbb{H}}(h_1, h_2)| \to 0. \tag{3.10}$$

Equation (3.8) holds since the enbelope function is $\mathbb{H}$ squre integrable. Then Equation (3.9) can be written as:

$$\frac{1}{n^2}\sum_{i=1}^{n} F^2(X_i)G^2(Y_i)1_{|F\times G| \geq \epsilon\sqrt{n}},$$

which is again bounded by the quantity in (3.7). This argument shows the validity of Condition 2.5 in this setting.

Finally:

$$\begin{aligned}\rho_{\mathbb{H}_n} &= (\mathbb{H}_n)(f \times g - \mathbb{H}_n(f \times g))^2 \\ &= \mathbb{H}_n(f \times g)^2 - (\mathbb{H}_n(f \times g))^2,\end{aligned}$$

and:

$$\rho_{\mathbb{H}} = \mathbb{H}(f \times g)^2 - (\mathbb{H}(f \times g))^2.$$

By substracting the previous quantities and taking absolute values and supremum over $f \times g \in \mathcal{F} \times \mathcal{G}$ we show that Equation 3.10 is satisfied, which implies that Condition 2.4 holds.

Then, we apply Theorem 2.3.1, for $R_n = \mathbb{H}_n$ and we obtain the convergence of $\hat{\mathbb{G}}_n$.

We have:

$$\begin{aligned}(\hat{\mathbb{H}}_n - \mathbb{H}_n)(f \times g) &= \int f \times g d(\hat{\mathbb{H}}_n - \mathbb{H}_n)\\ &= \int f \times g d\hat{\mathbb{H}}_n - \int f \times g d\mathbb{H}_n.\end{aligned}$$

But:

$$\begin{aligned}\int f \times g d\hat{\mathbb{H}}_n &= \frac{1}{n}\sum_{i=1}^{n}\int f \times g d\delta(X_i^*, Y_i^*)\\ &= \frac{1}{n}\sum_{i=1}^{n} f(X_i^*)g(Y_i^*),\end{aligned}$$

and:

$$\int f \times g d\mathbb{H}_n = \frac{1}{n^2}\sum_{i=1}^{n} f(X_i)g(Y_i).$$

By setting $g = 1$, in the previous we get:

$$\int f \times 1 d\hat{\mathbb{H}}_n = \frac{1}{n}\sum_{i=1}^{n} f(X_i^*) = \hat{\mathbb{P}}_n f,$$

$$\int f \times 1 d\mathbb{P}_n \otimes \mathbb{Q}_n = \frac{1}{n}\sum_{i=1}^{n} f(X_i) = \mathbb{P}_n f,$$

$$\sqrt{n}(\hat{\mathbb{H}}_n - \mathbb{H}_n)(f \times 1) = \sqrt{n}(\hat{\mathbb{P}}_n - \mathbb{P}_n) \xrightarrow{d} \mathbb{G}_{\mathbb{H}}(f \times 1),$$

and by setting $f = 1$ we get:

$$\int 1 \times g d\hat{\mathbb{H}}_n = \frac{1}{n}\sum_{i=1}^{n} g(Y_i^*) = \hat{\mathbb{Q}}_n g,$$

$$\int 1 \times g d\mathbb{P}_n \otimes \mathbb{Q}_n = \frac{1}{n}\sum_{i=1}^{n} g(Y_i) = \mathbb{Q}_n g,$$

$$\sqrt{n}(\hat{\mathbb{H}}_n - \mathbb{H}_n)(1 \times g) = \sqrt{n}(\hat{\mathbb{Q}}_n - \mathbb{Q}_n) \xrightarrow{d} \mathbb{G}_{\mathbb{H}}(1 \times g),$$

and by making use of the Continuous Mapping Theorem and Slutsky's Lemma, we get that the independence process:

$$\mathbb{S}_n(f \times g) \xrightarrow{d} \mathbb{G}_{\mathbb{H}}(f \times g) - \mathbb{G}_{\mathbb{H}}(f \times 1)\mathbb{Q}g - \mathbb{P}f\mathbb{G}_{\mathbb{H}}(1 \times g), \tag{3.11}$$

which is a Gaussian process with the same covariance as $\mathbb{Z}_H$.

According to 2.1.3 and the discussion resulting in the covariance of a $\mathbb{P}$-Brownian Bridge in (2.3), a Brownian Bridge is completely determined by tightness and by:

$$\mathbb{E}_{\mathbb{P}}(\mathbb{G}f_1\mathbb{G}f_2) = \mathbb{P}(f_1 f_2) - \mathbb{P}(f_1)\mathbb{P}(f_2).$$

By making use of this expression for the covariance and the Fubini Theorem for probability measures we verify the claim that the limit in (3.11) has the same covariance as $\mathbb{Z}_{\mathbb{H}}$ and the limit process has indeed the same distribution with the Brownian Bridge. □

3.3.3. General Proofs for Independence Test

Theorem 3.3.4. *Assume the null hypothesis holds.*

Then, for almost every sequence $\mathbf{U}_n = (\mathbf{X}_n, \mathbf{Y}_n)$ and $(X_1^*, Y_1^*), ..., (X_n^*, Y_n^*)$ distributed according to the probability measure $P_n = \mathbb{H}_n$:

$$d_{BL}(\mathbb{Z}_n, \hat{\mathbb{S}}_n|\mathbf{U}_n) \xrightarrow{\mathbb{H}} 0.$$

Similarly, for almost every sequence $(X_1, Y_1), ..., (X_n, Y_n)$ and $(X_1^*, Y_1^*), ..., (X_n^*, Y_n^*)$ generated by the probability measures $P_n = \mathbb{P}_n \otimes \mathbb{Q}_n$:

$$d_{BL}(\mathbb{Z}_n, \hat{\mathbb{T}}_n|\mathbf{U}_n) \xrightarrow{\mathbb{H}} 0.$$

The preceding theorem is a direct consequence of Theorem 3.3.3 and Theorem 3.3.2.

We now proceed with the consistency of the Bootstrap statistics under the null hypothesis and the desired result for the asymptotic p-values.

Proof of Theorem 3.2.1. First, as mentioned earlier, by applying the tensor product $(1(X \leq s) \times 1(Y \leq t))$ to the process $\mathbb{Z}_n$, under the null hypothesis the true statistic T_n is obtained.

For Assertion 1 of the Theorem we apply the same tensor product to the process $\hat{\mathbb{S}}_n$ and by making use of the Theorem 3.3.4 and the Continuous Mapping Theorem we obtain the weak convergence:

$$d_{BL}\,(T_n\,,\,T_n^{*c}|\mathbf{U}_n) \xrightarrow{\mathbb{H}} 0.$$

For Assertion 2 of the Theorem we apply the same tensor product to the process $\hat{\mathbb{T}}_n$ and by making use of the Theorem 3.3.4 and the Continuous Mapping Theorem we obtain the weak convergence:

$$d_{BL}\left(T_n\,,\,T_n^{*eq}|\mathbf{U}_n\right) \xrightarrow{\mathbb{H}} 0.$$

□

We now present the proof of Theorems 3.2.2 and 3.2.3, for the p-values under the null and alternative hypothesis, respectively.

Proof of Theorem 3.2.2. Combine Theorem 3.2.1 with Corollary 1.6 and the assertion follows.

□

Proof of Theorem 3.2.3. Let $\hat{\mathbb{R}}_n = \mathbb{Z}_n + \sqrt{n}(\mathbb{H} - \mathbb{P} \otimes \mathbb{Q})$, denote the process that produces the true statistic T by applying the tensor product $f \times g = 1(-\infty, s) \times 1(-\infty, t) = 1_A$ and the norm, where:

$$A = (-\infty, s) \times (-\infty, t). \tag{3.12}$$

Furthermore, denote $h : \mathcal{P} \to \mathbb{R}$ the continuous mapping that applies $f \times g$ and the norm to the independence processes, i.e. :

$$T_n(X_1, \ldots, X_n) = h(\hat{\mathbb{R}}_n),$$
$$T_n^{*c}(X_1, \ldots, X_n) = h(\hat{\mathbb{S}}_n),$$
$$T_n^{*eq}(X_1, \ldots, X_n) = h(\hat{\mathbb{T}}_n).$$

Specifically, we get:

$$\|\hat{\mathbb{R}}_n(f \times g)\| = \|\mathbb{Z}_n(f \times g) + \sqrt{n}(\mathbb{H}(1(-\infty, s) \times 1(-\infty, t)))\|.$$

Now, recall from Theorem 3.3.1 that:

$$\mathbb{Z}_n \xrightarrow{d} \mathbb{Z}_{\mathbb{H}}.$$

Then by applying ϕ to $\hat{\mathbb{R}}_n$, under H_1, we get:

$$\lim_{n\to\infty} T(X_1, \ldots, X_n) = \lim_{n\to\infty} \|\mathbb{Z}_n(f \times g) + \sqrt{n}(\mathbb{H}(1(-\infty, s) \times 1(-\infty, t)))\| \tag{3.13}$$

$$\geq \lim_{n\to\infty} \|\sqrt{n}(\mathbb{H}(1(-\infty, s) \times 1(-\infty, t)))\| - \|\mathbb{Z}_{\mathbb{H}}\| = +\infty, \tag{3.14}$$

where we used the triangle inequality and the Continuous Mapping Theorem for h.

For Assertion 1 of the Theorem we recall from Theorem 3.3.3 that:

$$\hat{\mathbb{S}}_n | \mathbf{U}_n \xrightarrow{d} \mathbb{Z}_{\mathbb{H}},$$

which is a Gaussian process and:

$$\mathbb{H}(\mathbb{Z}_{\mathbb{H}} < \infty) = 1.$$

Hence by applying h to $\hat{\mathbb{S}}_n$:

$$\lim_{n\to\infty} \mathbb{H}(T_n^{*c} < T_n | \mathbf{U}_n) = \lim_{n\to\infty} \mathbb{H}(h(\hat{\mathbb{S}}_n) < h(\hat{\mathbb{R}}_n) | \mathbf{U}_n)$$
$$= \mathbb{H}(h(\mathbb{Z}_{\mathbb{H}}) < \infty) = 1.$$

Similarly for Assertion 2, from Theorem 3.3.2:

$$\hat{\mathbb{R}}_n | \mathbf{U}_n \xrightarrow{d} \mathbb{Z}_{\mathbb{P}\otimes\mathbb{Q}},$$

which is a Gaussian process and:

$$\mathbb{H}(\mathbb{Z}_{\mathbb{P}\otimes\mathbb{Q}} < \infty) = 1.$$

Hence by applying h to $\hat{\mathbb{R}}_n$:

$$\lim_{n\to\infty} \mathbb{H}(T_n^{*eq} < T_n | \mathbf{U}_n) = \lim_{n\to\infty} \mathbb{H}(h(\hat{\mathbb{R}}_n) < h(\hat{\mathbb{R}}_n) | \mathbf{U}_n)$$
$$= \mathbb{H}(h(\mathbb{Z}_{\mathbb{P}\otimes\mathbb{Q}}) < \infty) = 1.$$

and the claim has been proved. □

Finally, we show that the combinations (T_n^{*eq}, GR), (T_n^{*c}, NHR) do not result in hypothesis tests with good power. .

Proof of Theorem 3.2.5. First, from Theorem 3.3.1 we have that:

$$\mathbb{Z}_n \xrightarrow{d} \mathbb{Z}_{\mathbb{H}}.$$

For the first part, from Theorem 3.3.3, we have that $\hat{\mathbb{S}}_n | \mathbf{U}_n \xrightarrow{d} \mathbb{Z}_{\mathbb{H}}$. We can rewrite $\hat{\mathbb{T}}_n$ and obtain the convergence under H_0:

$$\hat{\mathbb{T}}_n = \hat{\mathbb{S}}_n + \mathbb{Z}_n \xrightarrow{d} 2\mathbb{Z}_{\mathbb{H}} \text{ conditional on } \mathbf{U}_n.$$

From the Continuous Mapping Theorem:

$$\lim_{n\to\infty} d_{BL}\left(T_n, T_n^{*eq} | \mathbf{U}_n\right) = \lim_{n\to\infty} d_{BL}(h(\mathbb{Z}_n), h(\hat{\mathbb{T}}_n)) = d_{BL}(h(\mathbb{Z}_{\mathbb{H}}), h(2\mathbb{Z}_{\mathbb{H}})) \neq 0.$$

For the second part, from Theorem 3.3.2 we have $\hat{\mathbb{T}}_n \xrightarrow{d} \mathbb{Z}_{\mathbb{P}\otimes\mathbb{Q}}$. Similarly under H_0:

$$\hat{\mathbb{S}}_n = \hat{\mathbb{T}}_n - \mathbb{Z}_n \xrightarrow{d} \mathbb{Z}_{\mathbb{P}\otimes\mathbb{Q}} - \mathbb{Z}_{\mathbb{H}} \stackrel{d}{=} 0,$$

where the convergence holds conditionally on $\mathbf{U}_n$. By applying the Continuous Mapping Theorem:

$$\lim_{n\to\infty} d_{BL}\left(T_n, T_n^{*c} | \mathbf{U}_n\right) = \lim_{n\to\infty} d_{BL}(h(\mathbb{Z}_n), h(\hat{\mathbb{S}}_n)) = d_{BL}(h(\mathbb{Z}_{\mathbb{H}}), h(0)) \neq 0.$$

□

4

Regression Slope Test

4.1. Setting

Consider the Regression setting as defined in Section 1.6.1. In this setting we observe a joint sample $(X_i, Y_i)_{i=1}^n$ from a joint distribution $(X, Y) \sim \mathbb{H}$. Then the regression model admits the following assumptions:

$$Y_i = a + bX_i + \epsilon_i, \tag{4.1}$$

$$\mathbb{E}(\epsilon_i) = 0, \tag{4.2}$$

$$Var(\epsilon_i) = \sigma_\epsilon^2. \tag{4.3}$$

We denote $\mathbb{P}, \mathbb{Q}$ the marginal distributions of X, Y respectively, i.e. $X \sim \mathbb{P}, Y \sim \mathbb{Q}$. Similarly, as in the independence test, we assume that the random variables X, Y are defined on measurable spaces $(\mathcal{X}, \mathcal{A}), (\mathcal{Y}, \mathcal{B})$. Next, denote $\phi : \mathcal{D}[0, 1] \to \mathbb{R}$ the Hadamard differentiable mapping:

$$\phi(\mathbb{H}) = \frac{\mathbb{E}_{\mathbb{H}}(XY) - \mathbb{E}_{\mathbb{H}}(X)\mathbb{E}_{\mathbb{H}}(Y)}{\mathbb{E}_{\mathbb{H}}(X^2) - \mathbb{E}_{\mathbb{H}}(X)^2}.$$

It is known that $b = \phi(\mathbb{H})$. We want to test the following hypothesis:

$$H_0 : b = 0 \text{ } vs \text{ } H_1 : b \neq 0.$$

or the test can be framed equivalently:

$$H_0 : \phi(\mathbb{H}) = 0 \text{ } vs \text{ } H_1 : \phi(\mathbb{H}) \neq 0.$$

Let $\mathbb{H}_n = \sum_{i=1}^n \delta(X_i, Y_i)$ the empirical process of the joint sample. Then for the OLS estimator $\hat{b}$ the following holds:

$$\begin{aligned} \hat{b} &= \frac{\mathbb{E}_{\mathbb{H}_n}(XY) - \mathbb{E}_{\mathbb{H}_n}(X)\mathbb{E}_{\mathbb{H}_n}(Y)}{\mathbb{E}_{\mathbb{H}_n}(X^2) - \mathbb{E}_{\mathbb{H}_n}(X)^2} \\ &= \phi(\mathbb{H}_n). \end{aligned}$$

Given an original sample $(X_i, Y_i)_{i=1}^n$ and a Bootstrap sample $(X_i^*, Y_i^*)_{i=1}^n$, we remind the empirical processes:

$$\mathbb{H}_n = \frac{1}{n}\sum_{i=1}^n \delta(X_i, Y_i),$$

$$\hat{\mathbb{H}}_n = \frac{1}{n}\sum_{i=1}^n \delta(X_i^*, Y_i^*).$$

Based on the empirical processes, the following processes are defined:

$$\mathbb{Z}_n = \sqrt{n}(\phi(\mathbb{H}_n) - \phi(\mathbb{H})), \tag{4.4}$$

$$\mathbb{S}_n = \sqrt{n}(\phi(\hat{\mathbb{H}}_n) - \phi(\mathbb{H}_n)), \tag{4.5}$$

$$\mathbb{T}_n = \sqrt{n}\phi(\hat{\mathbb{H}}_n), \tag{4.6}$$

and we set the true and Bootstrap statistics as follows:

$$T_n = \sqrt{n}|\hat{b}|,$$
$$T_n^{*c} = |\mathbb{S}_n| = \sqrt{n}|\hat{b}^* - \hat{b}|,$$
$$T_n^{*eq} = |\mathbb{T}_n| = \sqrt{n}|\hat{b}^*|.$$

The objective is to show that, when the null hypothesis holds, under GR Bootstrap:

$$d_{BL}\,(T_n\,,\,T_n^{*c}|\mathbf{U}_n) \xrightarrow{\mathbb{H}} 0,$$

while under NHR Bootstrap:

$$d_{BL}\left(T_n\,,\,T_n^{*eq}|\mathbf{U}_n\right) \xrightarrow{\mathbb{H}} 0.$$

Under the null hypothesis we obtain the true statistic is obtained by $T_n = \|\mathbb{Z}_n\|$. Theorems 4.3.3, 4.3.4 imply the consistency with respect to the true statistic $\mathbb{Z}_n$:

$$\mathbb{S}_n \text{ in GR Bootstrap, } R_n = \mathbb{H}_n,$$
$$\mathbb{T}_n \text{ in NHR Bootstrap, } R_n = \mathbb{P}_n \otimes \mathbb{Q}_n.$$

First, we display the asymptotic behaviour of the original process $\mathbb{Z}_n$, the limit of which, under the null hypothesis, is the same as the limit of the true statistic T_n, normed. In order to examine the asymptotic behaviour of these processes we make use of the Delta Method.

Theorem 4.1.1 (Delta Method)**.** *[Theorem 3.9.4 in[10]] Let $\mathbb{D}, \mathbb{E}$ be metrizable topological spaces. Let $\phi : \mathbb{D}_\phi \mapsto \mathbb{E}$ be Hadamard differentiable at θ tangentially to $\mathbb{D}_0$. Let $X_n : \Omega_n \mapsto \mathbb{D}_\phi$ be maps with $r_n(X_n - \theta) \xrightarrow{d} X$, for a sequence of constants $r_n \to \infty$, where X is separable and takes values in $\mathbb{D}_0$. Then:*

$$r_n(\phi(X_n) - \phi(\theta)) \xrightarrow{d} \phi'_\theta(X).$$

In this setting we are interested in the asymptotic behavior of a Hadamard transformation of $\mathbb{H}_n$, hence we replace X_n by $\mathbb{H}_n$, where the randomness of $\mathbb{H}_n$ is derived from the random sample $\mathbf{X}_n$. The Donsker theorem gives a convergence rate of $r_n = \sqrt{n}$, $\theta = \mathbb{H}$, and $\mathbb{G}_\mathbb{H}$ is the asymptotic limit of the empirical process, equivalent to the limit X, in the preceding theorem. Following this discussion, we apply the Delta method to obtain the limit of the process $\mathbb{Z}_n$.

Theorem 4.1.2. *Let X, Y random variables defined on on measurable spaces $(\mathcal{X}, \mathcal{A}), (\mathcal{Y}, \mathcal{B})$ respectively and $\mathcal{F}, \mathcal{G}$ classes of functions the equivalent measure spaces. Denote $\mathbb{H}$ the joint distribution of X, Y on the measurable space $(\mathcal{X} \times \mathcal{Y}, \mathcal{A} \times \mathcal{B})$. If the class $\mathcal{F} \times \mathcal{G}$ is $\mathbb{H}$-Donsker then the sequence $\mathbb{Z}_n$ converges in distribution in $\ell^\infty(\mathcal{F} \times \mathcal{G})$:*

$$\mathbb{Z}_n \xrightarrow{d} \phi'_\mathbb{H}(\mathbb{G}_\mathbb{H}),$$

where $\mathbb{G}_\mathbb{H}$ is an $\mathbb{H}-$Brownian Bridge and ϕ' is the Hadamard derivative of ϕ.

Proof. The class of functions $\mathcal{F} \times \mathcal{G}$ is Donsker thus, for the empirical measure $\mathbb{H}_n$, it holds that:

$$\sqrt{n}(\mathbb{H}_n - \mathbb{H}) \xrightarrow{d} \mathbb{G}_\mathbb{H},$$

where $\mathbb{G}_\mathbb{H}$ is an $\mathbb{H}-$Brownian Bridge. Further, ϕ is a Hadamard differentiable function, thus the Delta method implies:

$$\sqrt{n}(\phi(\mathbb{H}_n) - \phi(\mathbb{H})) \xrightarrow{d} \phi'_\mathbb{H}(\mathbb{G}_\mathbb{H}),$$

which concludes the proof. □

Remark 4.1.3. *Note that the Hadamard derivative is a linear transformation of its inputs. Hence, ϕ' calculated at a Brownian Bridge $\mathbb{G}$, returns a Gaussian process with mean zero.*

4.2. Main Results

Theorem 4.2.1. *Under the null hypothesis $b = \phi(\mathbb{H}) = 0$, for the expressions of the true and Bootstrap statistics, as in* (4.4)*, the following hold:*

1. *For almost every sequence $(X_1, Y_1), ..., (X_n, Y_n)$ and $(X_1^*, Y_1^*), ..., (X_n^*, Y_n^*)$ distributed according to the probability measure $R_n = \mathbb{H}_n$:*

$$d_{BL}\,(T_n\,,\,T_n^{*c}|\mathbf{U}_n) \xrightarrow{\mathbb{H}} 0.$$

2. *For almost every sequence $(X_1, Y_1), ..., (X_n, Y_n)$ and $(X_1^*, Y_1^*), ..., (X_n^*, Y_n^*)$ distributed according to the probability measure $R_n = \mathbb{P}_n \otimes \mathbb{Q}_n$:*

$$d_{BL}\left(T_n\,,\,T_n^{*eq}|\mathbf{U}_n\right) \xrightarrow{\mathbb{H}} 0.$$

Theorem 4.2.2. *Under the null hypothesis the following hold:*

1. *For a Bootstrap Scheme of GR (generate Bootstrap samples from $\mathbb{H}_n$):*

$$pvalue_n = \mathbb{H}\left(T_n^{*c} > T_n|\mathbf{U}_n\right) \xrightarrow{d} U(0,1).$$

2. *For a Bootstrap Scheme of NHR (generate Bootstrap samples from $\mathbb{P}_n \otimes \mathbb{Q}_n$):*

$$pvalue_n = \mathbb{H}\left(T_n^{*eq} > T_n|\mathbf{U}_n\right) \xrightarrow{d} U(0,1).$$

Theorem 4.2.3. *Under the alternative hypothesis the following hold:*

1. *Under GR Bootstrap, $R_n = \mathbb{H}_n$:*

$$\mathbb{H}\left(T_n^{*c} < T_n|\mathbf{U}_n\right) \xrightarrow{d} 1.$$

2. *Under GR Bootstrap, $R_n = \mathbb{P}_n \otimes \mathbb{Q}_n$:*

$$\mathbb{H}\left(T_n^{*eq} < T_n|\mathbf{U}_n\right) \xrightarrow{d} 1.$$

Theorem 4.2.4. *Then under the null hypothesis the following hold:*

$$d_{BL}\,(T_n\,,\,T_n^{*c}|\mathbf{U}_n) \nrightarrow 0 \textit{ under GR Bootstrap}, R_n = \mathbb{H}_n,$$
$$d_{BL}\left(T_n\,,\,T_n^{*eq}|\mathbf{U}_n\right) \nrightarrow 0 \textit{ under NHR Bootstrap}, R_n = \mathbb{P}_n \otimes \mathbb{Q}_n.$$

For a Bootstrap statistic T_n^*, in the independence test we define the power of the test as:

$$\mathbb{H} \mapsto \pi(\mathbb{H}; T_n, T_n^*) := \mathbb{H}(T_n^* < T_n|\mathbf{X}_n).$$

Theorem 4.2.5. *In the Regression Slope test of significance α, the following hold.*

- *Under GR Bootstrap, $R_n = \mathbb{H}_n$, for the Bootstrap statistic T_n^{*c}:*

  $$\text{for all } \mathbb{H} \in \mathcal{P}_0,\ \lim_{n\to\infty} \pi(\mathbb{H}; T_n, T_n^{*c}) = \alpha,$$

  and

  $$\text{for all } \mathbb{H} \in \mathcal{P}\backslash\mathcal{P}_0,\ \lim_{n\to\infty} \pi(\mathbb{H}; T_n, T_n^{*c}) = 1.$$

- *Under NHR Bootstrap, $R_n = \mathbb{P}_n \otimes \mathbb{Q}_n$, for the Bootstrap statistic T_n^*:*

  $$\text{for all } \mathbb{H} \in \mathcal{P}_0,\ \lim_{n\to\infty} \pi(\mathbb{H}; T_n, T_n^{*eq}) = \alpha,$$

  and

  $$\text{for all } \mathbb{H} \in \mathcal{P}\backslash\mathcal{P}_0,\ \lim_{n\to\infty} \pi(\mathbb{H}; T_n, T_n^{*eq}) = 1,$$

4.3. Hadamard Derivative and Proof of Results

Let $(X_i, Y_i) \sim \mathbb{H}$ a random sample and set the empirical processes:

$$\mathbb{P}_n = \frac{1}{n}\sum_{i=1}^{n} \delta(X_i),$$

$$\mathbb{Q}_n = \frac{1}{n}\sum_{i=1}^{n} \delta(Y_i).$$

As mentioned in the independence test, we can generate different Bootstrap samples by simulating from $\mathbb{H}_n$ and $\mathbb{P}_n \otimes \mathbb{Q}_n$ and they belong to GR and NHR Bootstrap schemes respectively.

The Donsker Theorem implies that:

$$\sqrt{n}(\mathbb{H}_n - \mathbb{H}) \xrightarrow{d} \mathbb{G}_{\mathbb{H}}.$$

The mapping ϕ satisfies the assumptions for the Delta method, hence:

$$\sqrt{n}(\phi(\mathbb{H}_n) - \phi(\mathbb{H})) \xrightarrow{d} \phi'_{\mathbb{H}}(\mathbb{G}_{\mathbb{H}}).$$

Under the null hypothesis it holds that:

$$\sqrt{n}\phi(\mathbb{H}_n) \xrightarrow{d} \phi'_{\mathbb{H}}(\mathbb{G}_{\mathbb{H}}).$$

In order to show consistency we would require that the asymptotic limit of the Bootstrap statistic would be equal to $\phi'_{\mathbb{H}}(\mathbb{G}_{\mathbb{H}})$, hence the motivation to calculate the Hadamard Derivative of ϕ and for probability measure $\mathbb{H}$ and random variable ϵ:

$$\phi'_H(\epsilon) = \frac{\mathbb{E}_\epsilon(XY) - \mathbb{E}_H(X)\mathbb{E}_\epsilon(Y) - \mathbb{E}_\epsilon(X)\mathbb{E}_H(Y) - \phi(H)\mathbb{E}_H(Y)\mathbb{E}_\epsilon(X^2) + 2\phi(H)\mathbb{E}_H(X)}{\mathbb{E}_H(X^2) - \mathbb{E}_H(X)^2}.$$

The proof of this calculation is given on Lemma A.0.1 of the Appendix A.

In a natural manner, it is of high interest to extend the Delta Method for sequences according to the following theorem.

Definition 4.3.1. *A mapping $\phi : \mathbb{D}_\phi \mapsto \mathbb{E}$ is called Uniformly Hadamard differentiable if:*

$$r_n(\phi(\theta_n + r_n^{-1}h_n) - \phi(\theta_n)) \xrightarrow{d} \phi'_\theta(h), \tag{4.7}$$

for a sequence of constants $r_n \to \infty$, for every converging sequence h_n with $\theta_n + r_n^{-1}h_n \in \mathbb{D}_\phi$ for all n and $h_n \to h \in \mathbb{D}_0$ and some arbitrary map ϕ'_θ on $\mathbb{D}_0$.

Theorem 4.3.2 (Delta Method). *[Theorem 3.9.5 in[10]] Let $\mathbb{D}, \mathbb{E}$ be metrizable topological variable spaces and let r_n sequence of constants such that $r_n \to \infty$. Let $\phi : \mathbb{D}_\phi \mapsto \mathbb{E}$ Uniformly Hadamard differentiable according to Definition 4.3.1.*

Let $X_n : \Omega_n \mapsto \mathbb{D}_\phi$ be maps with $r_n(X_n - \theta_n) \xrightarrow{d} X$, for a sequence of constants $r_n \to \infty$, where X is separable and takes values in $\mathbb{D}_0$. Then:

$$r_n(\phi(X_n) - \phi(\theta_n)) \xrightarrow{d} \phi'_\theta(X).$$

The condition in Definition (4.3.1) is the sufficient condition for ϕ, to extend the the Delta Method for the Central Limit Theorem under sequences.

Theorem 4.3.3. *[GR Bootstrap] Let $\mathcal{F}, \mathcal{G}$ separable classes of measurable functions on measurable spaces $(\mathcal{X}, \mathcal{A}), (\mathcal{Y}, \mathcal{B})$ such that $\mathcal{F} \times \mathcal{G}$ satisfies the Uniform Entropy Condition for envelope functions*

$F, G, F \times G$ that are $\mathbb{H}$-square integrable. Given a sample $\mathbf{U}_n = (\mathbf{X}_n, \mathbf{Y}_n)$ from $\mathbb{H}_n$, let $(X_i^*, Y_i^*)_{i=1}^n$ a Bootstrap sample generated by $R_n = \mathbb{H}_n$. Conditionally on $\mathbf{U}_n$ denote:

$$\hat{\mathbb{S}}_n = \sqrt{n}(\phi(\hat{\mathbb{H}}_n) - \phi(\mathbb{H}_n)).$$

Then:

$$\hat{\mathbb{S}}_n | \mathbf{U}_n \xrightarrow{d} \phi'_{\mathbb{H}}(\mathbb{G}_{\mathbb{H}}).$$

for $\mathbb{H}^\infty$-almost every sequence $(X_1, Y_1), (X_2, Y_2), \ldots$, where $\mathbb{G}_{\mathbb{H}}$ is a $\mathbb{H}$ Brownian Bridge.

Proof. As mentioned in the previous chapter, the sequence $R_n = \mathbb{H}_n$, satisfies the conditions for the Theorem 2.3.1 with limit $\mathbb{H}$, hence:

$$\sqrt{n}(\hat{\mathbb{H}}_n - \mathbb{H}_n) \xrightarrow{d} \mathbb{G}_{\mathbb{H}}.$$

Further, ϕ satisfies the conditions for Theorem 4.3.2, hence:

$$\sqrt{n}(\phi(\hat{\mathbb{H}}_n) - \phi(\mathbb{H}_n)) \xrightarrow{d} \phi'_{\mathbb{H}}(\mathbb{G}_{\mathbb{H}}),$$

which concludes the proof. □

Theorem 4.3.4. *[NHR Bootstrap] Let $\mathcal{F}, \mathcal{G}$ separable classes of measurable functions on measurable spaces $(\mathcal{X}, \mathcal{A}), (\mathcal{Y}, \mathcal{B})$ such that $\mathcal{F} \times \mathcal{G}$ satisfies the Uniform Entropy Condition for envelope functions $F, G, f \times g$ that are $\mathbb{H}$-square integrable. Given a sample $\mathbf{U}_n = (\mathbf{X}_n, \mathbf{Y}_n)$ from $\mathbb{H}$, let $(X_i^*, Y_i^*)_{i=1}^n$ a Bootstrap sample generated by $R_n = \mathbb{P}_n \otimes \mathbb{Q}_n$. Conditionally on $\mathbf{U}_n$ denote:*

$$\hat{\mathbb{T}}_n = \sqrt{n}\phi(\hat{\mathbb{H}}_n).$$

Then:

$$\hat{\mathbb{T}}_n | \mathbf{U}_n \xrightarrow{d} \phi'_{\mathbb{P}\otimes\mathbb{Q}}(\mathbb{G}_{\mathbb{P}\otimes\mathbb{Q}}),$$

for $\mathbb{H}^\infty$-almost every sequence $(X_1, Y_1), (X_2, Y_2), \ldots$, where $\mathbb{G}_{\mathbb{P}\otimes\mathbb{Q}}$ is a $\mathbb{P} \otimes \mathbb{Q}$ Brownian Bridge.

Proof. As mentioned in the previous chapter, the sequence $R_n = \mathbb{P}_n \otimes \mathbb{Q}_n$, satisfies the conditions for the Theorem 2.3.1 with limit $\mathbb{P} \otimes \mathbb{Q}$, hence:

$$\sqrt{n}(\hat{\mathbb{H}}_n - \mathbb{P}_n \otimes \mathbb{Q}_n) \xrightarrow{d} \mathbb{G}_{\mathbb{P}\otimes\mathbb{Q}}.$$

Further, ϕ satisfies the conditions for Theorem 4.3.2, hence:

$$\sqrt{n}(\phi(\hat{\mathbb{H}}_n) - \phi(\mathbb{P}_n \otimes \mathbb{Q}_n)) \xrightarrow{d} \phi'_{\mathbb{P}\otimes\mathbb{Q}}(\mathbb{G}_{\mathbb{P}\otimes\mathbb{Q}}).$$

However, the way ϕ is defined, it holds that $\phi(\mathbb{P}_n \otimes \mathbb{Q}_n) = 0$, which concludes the proof. □

4.3.1. General Proofs for the Regression Test

Given the Theorems of the previous subsections we proceed with the proofs for the consistency.

Proof of Theorem 4.2.1. First, Theorem 4.1.2 implies the convergence of $\mathbb{Z}_n$ to $\phi'_{\mathbb{H}}(\mathbb{G}_{\mathbb{H}})$. For $T_n^{*c} = \|\mathbb{S}_n\|$ and $R_n = \mathbb{H}_n$, the proof is straightforward since Theorem 4.3.3 implies the convergence of $\mathbb{S}_n$ to $\phi'_{\mathbb{H}}(\mathbb{G}_{\mathbb{H}})$. The result follows by applying the Continuous Mapping Theorem to these processes, for the continuous mapping of the norm.

For the combination $T_n^{*eq} = \|\mathbb{T}_n\|$ and $R_n = \mathbb{P}_n \otimes \mathbb{Q}_n$, Theorem 4.3.4 implies the convergence of $\mathbb{T}_n$ to $\phi'_{\mathbb{P}\otimes\mathbb{Q}}(\mathbb{G}_{\mathbb{P}\otimes\mathbb{Q}})$. Hence, the objective is to show that under the null hypothesis:

$$\phi'_{\mathbb{H}}(\mathbb{G}_{\mathbb{H}}) \stackrel{d}{=} \phi'_{\mathbb{P}\otimes\mathbb{Q}}(\mathbb{G}_{\mathbb{P}\otimes\mathbb{Q}}).$$

This claim is fully proved in Lemma A.0.2, therefore we provide a sketch of the proof. As noted in Remark 4.1.3, we know that since ϕ' is a linear transformation of its argument and the Brownian Bridge is a zero mean Gaussian process, then both $\phi'_{\mathbb{H}}(\mathbb{G}_{\mathbb{H}})$ and $\phi'_{\mathbb{P}\otimes\mathbb{Q}}(\mathbb{G}_{\mathbb{P}\otimes\mathbb{Q}})$ are zero mean Gaussian processes. Thus, in order to prove that they are equal in distribution, it suffices to show that their variances are equal.

Under the null hypothesis, for $\phi'_{\mathbb{H}}(\mathbb{G}_{\mathbb{H}})$, the following holds:

$$\begin{aligned}\phi'_{\mathbb{H}}(\mathbb{G}_{\mathbb{H}}) &= \frac{\mathbb{E}_{\mathbb{G}_{\mathbb{H}}}(XY) - \mathbb{E}_{\mathbb{H}}(X)\mathbb{E}_{\mathbb{G}_{\mathbb{H}}}(Y) - \mathbb{E}_{\mathbb{G}_{\mathbb{H}}}(X)\mathbb{E}_{\mathbb{H}}(Y)}{\mathbb{E}_{\mathbb{H}}(X^2) - \mathbb{E}_{\mathbb{H}}(X)^2} \\ &= \frac{\int XYd(\mathbb{G}_{\mathbb{H}} - \mathbb{H} \otimes \mathbb{G}_{\mathbb{H}} - \mathbb{G}_{\mathbb{H}} \otimes \mathbb{H})}{\mathbb{E}_{\mathbb{H}}(X^2) - \mathbb{E}_{\mathbb{H}}(X)^2} \\ &= \frac{\int XYd(\mathbb{G}_{\mathbb{H}} - \mathbb{P} \otimes \mathbb{G}_{\mathbb{Q}} - \mathbb{G}_{\mathbb{P}} \otimes \mathbb{Q})}{\mathbb{E}_{\mathbb{H}}(X^2) - \mathbb{E}_{\mathbb{H}}(X)^2},\end{aligned}$$

where we used the fact that $\int Xd\mathbb{G}_{\mathbb{H}} = \int Xd\mathbb{G}_{\mathbb{P}}$ and $\int Yd\mathbb{G}_{\mathbb{H}} = \int Yd\mathbb{G}_{\mathbb{Q}}$. The proof of these claims is given in the Appendix. Similarly, we calculate:

$$\phi'_{\mathbb{P}\otimes\mathbb{Q}}(\mathbb{G}_{\mathbb{P}\otimes\mathbb{Q}}) = \frac{\int XYd(\mathbb{G}_{\mathbb{P}\otimes\mathbb{Q}} - \mathbb{P} \otimes \mathbb{G}_{\mathbb{Q}} - \mathbb{G}_{\mathbb{P}} \otimes \mathbb{Q})}{\mathbb{E}_{\mathbb{H}}(X^2) - \mathbb{E}_{\mathbb{H}}(X)^2}.$$

It holds that $\int XYd\mathbb{G}_{\mathbb{H}} \overset{d}{=} \int XYd\mathbb{G}_{\mathbb{P}\otimes\mathbb{Q}}$. The result is a direct consequence of the Continuous Mapping Theorem for the norm, and the proof is concluded. □

Proof of Theorem 4.2.1. First, Theorem 4.1.2 implies that:

$$\mathbb{Z}_n \xrightarrow{d} \phi'_{\mathbb{H}}(\mathbb{G}_{\mathbb{H}}).$$

Under the null hypothesis and by applying the norm to process $\mathbb{Z}_n$, the true statistic T_n is obtained. The continuous mapping theorem then, implies that:

$$T_n = |\mathbb{Z}_n| \xrightarrow{d} |\phi'_{\mathbb{H}}(\mathbb{G}_{\mathbb{H}})|.$$

For the first part of the Theorem, it is known from Theorem 4.3.3 that:

$$\sqrt{n}(\phi(\hat{\mathbb{H}}_n) - \phi(\mathbb{H}_n)) \xrightarrow{d} \phi'_{\mathbb{H}}(\mathbb{G}_{\mathbb{H}}).$$

By applying the norm to the previous process, we obtain the Bootstrap statistic T_n^{*c}. The Continuous Mapping Theorem implies that $T_n^{*c} \xrightarrow{d} |\phi'_{\mathbb{H}}(\mathbb{G}_{\mathbb{H}})|$, hence:

$$d_{BL}\,(T_n\,, T_n^{*c}|\mathbf{U}_n) \xrightarrow{\mathbb{H}} 0,$$

and the proof of the first part is concluded.

For the second part, Theorem 4.3.4 implies that:

$$\sqrt{n}\phi(\hat{\mathbb{H}}_n) \xrightarrow{d} \phi'_{\mathbb{P}\otimes\mathbb{Q}}(\mathbb{G}_{\mathbb{P}\otimes\mathbb{Q}}).$$

By applying the norm to the previous process, the Bootstrap statistic T_n^{*eq} is obtained. The Continuous Mapping Theorem implies that $T_n^{*eq} \xrightarrow{d} |\phi'_{\mathbb{P}\otimes\mathbb{Q}}(\mathbb{G}_{\mathbb{P}\otimes\mathbb{Q}})|$. Lemma A.0.2 implies that under H_0:

$$\phi'_{\mathbb{P}\otimes\mathbb{Q}}(\mathbb{G}_{\mathbb{P}\otimes\mathbb{Q}}) \overset{d}{=} \phi'_{\mathbb{H}}(\mathbb{G}_{\mathbb{H}}).$$

Hence:

$$d_{BL}\left(T_n\,, T_n^{*eq}|\mathbf{U}_n\right) \xrightarrow{\mathbb{H}} 0,$$

which concludes the second part. □

Proof of Theorem 4.2.2. Combine Theorem 4.2.1 with Corollary 1.6 and the assertion follows. □

Denote $\mathbb{R}_n = \mathbb{Z}_n + \sqrt{n}\phi(\mathbb{H})$. Under the null hypothesis it holds that $\mathbb{R}_n = \mathbb{Z}_n$. Recall from Theorem 4.1.2 that $\mathbb{Z}_n \xrightarrow{d} \phi'_{\mathbb{H}}(\mathbb{G}_{\mathbb{H}})$.

Proof of Theorem 4.2.3. Under the alternative hypothesis $\phi(\mathbb{H}) \neq 0$. The triangle inequality implies that:

$$|\mathbb{R}_n| \geq \sqrt{n}|\phi(\mathbb{H})| + |\mathbb{Z}_n| \xrightarrow{d} \infty,$$

since $\phi(\mathbb{H}) \neq 0$ and $\mathbb{Z}_n \xrightarrow{d} \phi'_{\mathbb{H}}(\mathbb{G}_{\mathbb{H}})$. For the first part:

$$\mathbb{H}\left(T_n^{*c} < T_n \middle| \mathbf{U}_n\right) = \mathbb{H}\left(|\hat{\mathbb{S}}_n| < |\mathbb{R}_n| \middle| \mathbf{U}_n\right) \to \mathbb{H}\left(|\phi'_{\mathbb{H}}(\mathbb{G}_{\mathbb{H}})| < \infty\right) = 1$$

For the second part:

$$\mathbb{H}\left(T_n^{*eq} < T_n \middle| \mathbf{U}_n\right) = \mathbb{H}\left(|\hat{\mathbb{T}}_n| < |\mathbb{R}_n| \middle| \mathbf{U}_n\right) \to \mathbb{H}\left(|\phi'_{\mathbb{P}\otimes\mathbb{Q}}(\mathbb{G}_{\mathbb{P}\otimes\mathbb{Q}})| < \infty\right) = 1$$

□

Proof of Theorem 4.2.4. For the first part, assume resampling scheme of $R_n = \mathbb{H}_n$ and observe that conditionally on $\mathbf{U}_n$:

$$\hat{\mathbb{T}}_n = \hat{\mathbb{S}}_n + \mathbb{R}_n.$$

Theorem 4.3.3 implies that $\hat{\mathbb{S}}_n|\mathbf{U}_n \xrightarrow{d} \phi'_{\mathbb{H}}(\mathbb{G}_{\mathbb{H}})$. The Continuous Mapping Theorem implies that $\hat{\mathbb{T}}_n|\mathbf{U}_n \xrightarrow{d} 2\phi'_{\mathbb{H}}(\mathbb{G}_{\mathbb{H}})$, which concludes the first part.

For the second part, assume resampling scheme of $R_n = \mathbb{P}_n \otimes \mathbb{Q}_n$ and observe that conditionally on $\mathbf{U}_n$:

$$\hat{\mathbb{S}}_n = \hat{\mathbb{T}}_n - \mathbb{R}_n.$$

Theorem 4.3.4 implies that $\hat{\mathbb{T}}_n|\mathbf{U}_n \xrightarrow{d} \phi'_{\mathbb{P}\otimes\mathbb{Q}}(\mathbb{G}_{\mathbb{P}\otimes\mathbb{Q}})$. The Continuous Mapping Theorem implies that $\hat{\mathbb{T}}_n|\mathbf{U}_n \xrightarrow{d} 0$, which concludes the second part.

□

5

General Case

5.1. Main Result

In the previous Chapter, we saw that for a resampling scheme R_n, with limit R, the condition $\phi_R(\mathbb{G}_R) \stackrel{d}{=} \phi_{\mathbb{H}}(\mathbb{G}_{\mathbb{H}})$, was the necessary condition for the Bootstrap statistic to be consistent with the true statistic. In the Regression setting, the test was reduced to the test $\phi(\mathbb{H}) = 0$, where ϕ was a Hadamard differentiable function. Naturally, it makes sense to search for a similar function that describes the Independence Test with the same reduction.

Let $\mathcal{P}$ denote the set of bivariate probability measures. Apparently, for random variables $X \sim \mathbb{P}, Y \sim \mathbb{Q}$ and $(X, Y) \sim \mathbb{H}$, the test can be posed:

$$H_0 : \phi(\mathbb{H}) = 0 \text{ vs } H_1 : \phi(\mathbb{H}) \neq 0,$$

where, $\phi : \mathcal{P} \to \mathbb{E}$, with $\phi(\mathbb{H}) = \mathbb{H} - \mathbb{P} \otimes \mathbb{Q}$. The partition of the measure class is then $\mathcal{P} = \mathcal{P}_0 \cup \mathcal{P} \backslash \mathcal{P}_0$, where

$$\mathcal{P}_0 = \{\mathbb{H} \in \mathcal{P} : \phi(\mathbb{H}) = 0\}.$$

The functional ϕ, maps an argument to the difference of the identity and tensor product of its projections on each dimension. In the Appendix B, it is shown that in this case ϕ is indeed Uniformly Hadamard differentiable and we can work similarly as in the Regression Chapter, to derive consistent Bootstrap statistics.

Hence, we are inclined to examine whether both the regression slope test and the independence test belong to a more general class of tests that could generalize the common assumptions and properties between them. We present a general theorem for Bootstrap Hypothesis Testing which is verified to be a general case for these two settings and could potentially include a larger class of testing settings.

Theorem 5.1.1. *Let $\mathbb{D}$ and $\mathbb{E}$ be two normed vector spaces. Let $\mathcal{P}$ be a set of probability measures on $\mathbb{D}$. Let $\mathcal{F}$ be a separable class on $\mathbb{D}$. Let $X_1, \ldots, X_n$ be an i.i.d. sample from a distribution $\mathbb{H} \in \mathcal{P}$. Consider the test:*

$$H_0 : \phi(\mathbb{H}) = 0 \ vs \ H_1 : \phi(\mathbb{H}) \neq 0,$$

for a uniformly Hadamard differentiable functional $\phi : \mathcal{P} \to \mathbb{E}$, where $\mathcal{P}$ is seen as a subspace of $\ell^\infty(\mathcal{F})$, equipped with the norm $\|\mathbb{H}\| := \sup_{f \in \mathcal{F}} |\mathbb{H}f|$. Let $(R_n)_{n=1}^\infty$ be a resampling scheme and assume that there exists a limit probability measure $R = \lim_{n\to\infty} R_n$, in the sense of the CLT under sequences:

$$\sqrt{n}(\hat{\mathbb{H}}_n - R_n) \xrightarrow{d} \mathbb{G}_R \text{ in } \ell^\infty(\mathcal{F}) \tag{5.1}$$

with the limit R satisfying

$$\phi'_R(\mathbb{G}_R) = \phi'_{\mathbb{H}}(\mathbb{G}_{\mathbb{H}}), \tag{5.2}$$

whenever $\phi(\mathbb{H}) = 0$. Let $T_n = \sqrt{n}\|\phi(\mathbb{H}_n)\|$ be the true test statistic. Then, using the bootstrap test statistic

$$T_n^* := \|\sqrt{n}(\phi(\hat{\mathbb{H}}_n) - \phi(R_n))\|,$$

yields a test such that

$$\text{for all } \mathbb{H} \in \mathcal{P}_0,\ \lim_{n\to\infty} \pi(\mathbb{H}; T_n, T_n^*) = \alpha,$$

and

$$\text{for all } \mathbb{H} \in \mathcal{P}\backslash\mathcal{P}_0,\ \lim_{n\to\infty} \pi(\mathbb{H}; T_n, T_n^*) = 1,$$

where $\mathcal{P}_0 := \{\mathbb{H} \in \mathcal{P} : \phi(\mathbb{H}) = 0\}$.

Proof. The empirical measure $\mathbb{H}_n$ satisfies:

$$\sqrt{n}(\mathbb{H}_n - \mathbb{H}) \xrightarrow{d} \mathbb{G}_{\mathbb{H}}.$$

The Delta method implies:

$$\mathbb{Z}_n = \sqrt{n}(\phi(\mathbb{H}_n) - \phi(\mathbb{H})) \xrightarrow{d} \phi'_{\mathbb{H}}(\mathbb{G}_{\mathbb{H}}),$$

for $\mathbb{G}_{\mathbb{H}}$ an $\mathbb{H}$-Brownian Bridge.

The function ϕ is Uniformly differentiable hence, the extension of the Delta method for sequences and the assumption (5.1) implies:

$$\sqrt{n}(\phi(\hat{\mathbb{H}}_n) - \phi(R_n)) \xrightarrow{d} \phi'_R(\mathbb{G}_R),$$

Under the null hypothesis the continuous mapping theorem implies that:

$$\begin{aligned} d_{BL}\,(T_n\,,\,T_n^*|\mathbf{X}_n) &= d_{BL}\Big(\sqrt{n}\|\phi(\mathbb{H}_n)\|\,,\,\sqrt{n}\|\phi(\hat{\mathbb{H}}_n) - \phi(R_n)\|\big|\mathbf{X}_n\Big) \\ &\xrightarrow{\mathbb{H}} d_{BL}(\|\phi'_{\mathbb{H}}(\mathbb{G}_{\mathbb{H}})\|, \|\phi'_R(\mathbb{G}_R)\|) = 0, \end{aligned}$$

where in the last equality we used (5.2). Further, it holds that:

$$\begin{aligned} T_n &= \|\mathbb{Z}_n + \sqrt{n}\phi(\mathbb{H})\| \\ &\geq \sqrt{n}\|\phi(\mathbb{H})\| - \|\mathbb{Z}_n\| \xrightarrow{d} \infty, \end{aligned}$$

under the alternative hypothesis $\phi(\mathbb{H}) \neq 0$ and by using the fact that $\lim_{n\to\infty}\|\mathbb{Z}_n\| = \|\phi'_{\mathbb{H}}(\mathbb{G}_{\mathbb{H}})\|$. Hence:

$$\mathbb{H}_{H_1}(T_n^* < T_n|\mathbf{X}_n) \xrightarrow{d} \mathbb{H}_{H_1}(\|\phi'_R(\mathbb{G}_R)\| < \infty) = 1,$$

where we used the fact that ϕ' is a linear functional of Gaussian processes and thus its norm is finite with probability 1. □

A Hypothesis Test in this framework is framed by the root of any Hadamard differentiable mapping ϕ of the measure $\mathbb{H}$. The mapping $\phi(\mathbb{H}_n)$ naturally yields the true statistic, since it directly inherits the distributional properties of the original sample, when the classic Delta Method 4.1.1 is applicable. For the examination Bootstrap processes asymptotic behaviour, the “slightly” stronger assumption of uniform differentiability 4.3.1 is required to extend the Delta method for the Bootstrap. The preceding theorem provides a concrete answer on to which Bootstrap statistic is the consistent with the true, depending directly on the resampling scheme R_n that is picked. Condition 5.2 states that a resampling scheme R_n should be generating new samples that preserve some similar structure of the null hypothesis. This equality of Hadamard derivatives summarizes this information via ϕ' by the resampling scheme R, its Brownian Bridge $\mathbb{G}_R$ and, of course, the structure of the setting described in ϕ itself. Hence, when possible to extend the Delta Method, and since the Delta Method yields a unique representation of the Bootstrap statistic ,this theorem implies that the Bootstrap consistency depends only on a proper choice of resampling scheme R_n.

It is natural to wonder whether the consistent statistic is indeed unique in some sense. Apparently, for a finite Bootstrap sample generated by R_n the quantity $\phi(R_n)$ is of high importance. The NHR scheme should by definition satisfy the null hypothesis, which in our setting means that $\phi(R_n) = 0$. This reduces the consistent Bootstrap process to $\sqrt{n}\phi(\hat{\mathbb{H}}_n)$. In contrast, the case of GR, where $R_n = \mathbb{H}_n$ yields a consistent process of the form $\sqrt{n}(\phi(\hat{\mathbb{H}}_n) - \phi(\mathbb{H}_n))$. It would be naive to state that these two Bootstrap processes differ only in the term $\phi(\mathbb{H}_n)$. The reason is that the Bootstrap samples are generated by different probability measures, hence the Bootstrap empirical measure $\hat{\mathbb{H}}_n$ should also reflect this difference. This discussion leads us to the equivalent formal version of Definition 1.6.1 that summarizes the information of Conjecture 1.6.1.

Definition 5.1.2 (Formal version of Definition 1.6.1)**.** *Let $X_1, \dots, X_n$ be a random sample from $\mathbb{H}$ and let $\mathcal{P}_0, \mathcal{P}\backslash\mathcal{P}_0$ be a partition of the class of probability measures $\mathcal{P}$, where:*

$$\mathcal{P}_0 = \{\mathbb{H} \in \mathcal{P} : \phi(\mathbb{H}) = 0\}.$$

Let $\mathbb{H}_n$ denote the empirical measure of the finite sample $\mathbf{X}_n$. A Bootstrap scheme $\hat{R} = (R_n)_{n=1}^{\infty}$ is called

- ***General Resampling (GR)** scheme if:*

$$\phi(R_n) \overset{d}{=} \phi(\mathbb{H}_n), \forall n \in \mathbb{N}.$$

- ***Null Resampling (NHR)** scheme if:*

$$\phi(R_n) = 0, \forall n \in \mathbb{N}.$$

This definition provides clarity regarding the distinct types of resampling methods investigated in this project. Firstly, we observe that a General Resampling (GR) scheme encompasses any sequence of measures that emulates the structure of the original sample. As a result, the original empirical measure, denoted as $\mathbb{H}_n$, serves as our point of reference. In contrast, a Null Resampling (NR) scheme constitutes any Bootstrap method that generates samples consistent with the null hypothesis. This observation, in conjunction with Theorem 5.1.1 and the consistent Bootstrap formula, leads us to a second observation and establishes a connection with the initial Conjecture 1.6.1. Specifically, we note that in the case of General Resampling, the consistent Bootstrap process is essentially ”centered” around the true statistic, as evidenced by $\phi(R_n) = \phi(\mathbb{H}_n)$. Conversely, in Null Resampling, the Bootstrap process is simplified to the term $\phi(\hat{\mathbb{H}}_n)$. However, it is crucial to avoid jumping to the conclusion that centering is the sole distinction. While the calculations leading to $\phi(\hat{\mathbb{H}}_n)$ remain consistent regardless of the choice of R_n, the difference in the simulation process is also manifest in $\hat{\mathbb{H}}_n$, given the structural variations in the distribution of the generated samples.

5.2. Application: Independence test revisited

In this section, we present the proof of Conjecture 1.6.1 by leveraging Theorem 5.1.1. We maintain our focus on the Independence Test setting, as outlined in Section 3.1. The argument for the Regression Slope test is straightforward, and the foundation for the assumptions in Theorem 5.1.1 was established within this context. In Chapter 4, we demonstrated that this setting constitutes a specific instance of the general theorem. Consequently, our current objective is to establish the same relationship for the Independence test.

We remind that for X, Y random variables defined on measurable spaces $(\mathcal{X}, \mathcal{A}), (\mathcal{Y}, \mathcal{B})$ respectively, such that $X \sim \mathbb{P}$, $Y \sim \mathbb{Q}$, and the joint $(X, Y) \sim \mathbb{H}$, on the product space $(\mathcal{X} \times \mathcal{Y}, \mathcal{A} \times \mathcal{B})$, the test is:

$$H_0 : \mathbb{H} = \mathbb{P} \otimes \mathbb{Q} \text{ vs } H_1 : \mathbb{H} \neq \mathbb{P} \otimes \mathbb{Q},$$

where $\mathbb{H} \in \mathcal{P}$ and $\mathcal{P}$ is the set of all bivariate probability measures. Let $\phi : \mathcal{P} \mapsto \mathbb{E}$, with $\phi(\mathbb{H}) = \mathbb{H} - \mathbb{H}(\cdot \times \mathbb{R}) \otimes \mathbb{H}(\mathbb{R} \times \cdot)$. We can see that for this setting the projections of the joint are exactly the marginal distributions $\mathbb{P}, \mathbb{Q}$ of X, Y. We can restate the test:

$$H_0 : \phi(\mathbb{H}) = 0 \text{ vs } H_1 : \phi(\mathbb{H}) \neq 0,$$

since $\phi(\mathbb{H}) = \mathbb{H} - \mathbb{P} \otimes \mathbb{Q}$. In Lemma B.0.1 of the Appendix, we show that ϕ is indeed Uniformly Hadamard differentiable and obtain the Hadamard derivative of ϕ calculated as:

$$\phi'_H(\epsilon) = \phi'_H(\epsilon) = \epsilon - H(\cdot \times \mathbb{R}) \otimes \epsilon(\mathbb{R} \times \cdot) - \epsilon(\cdot \times \mathbb{R}) \otimes H(\mathbb{R} \times \cdot).$$

Set $T_n = \sqrt{n}\left\|\phi(\mathbb{H}_n)\right\| = \sqrt{n}\|\mathbb{H}_n - \mathbb{P}_n \otimes \mathbb{Q}_n\|$. Under the null hypothesis the expression for the asymptotic empirical limit describing T_n as:

$$\begin{aligned}\phi'_{\mathbb{H}}(\mathbb{G}_{\mathbb{H}}) &= \mathbb{G}_{\mathbb{H}} - \mathbb{H}(\cdot \times \mathbb{R}) \otimes \mathbb{G}_{\mathbb{H}}(\mathbb{R} \times \cdot) - \mathbb{G}_{\mathbb{H}}(\cdot \times \mathbb{R}) \otimes \mathbb{H}(\mathbb{R} \times \cdot)\\ &= \mathbb{G}_{\mathbb{H}} - \mathbb{P}(\cdot \times \mathbb{R}) \otimes \mathbb{G}_{\mathbb{H}}(\mathbb{R} \times \cdot) - \mathbb{G}_{\mathbb{H}}(\cdot \times \mathbb{R}) \otimes \mathbb{Q}(\mathbb{R} \times \cdot)\\ &= \mathbb{G}_{\mathbb{H}} - \mathbb{P}(\cdot \times \mathbb{R}) \otimes \mathbb{G}_{\mathbb{Q}}(\mathbb{R} \times \cdot) - \mathbb{G}_{\mathbb{P}}(\cdot \times \mathbb{R}) \otimes \mathbb{Q}(\mathbb{R} \times \cdot).\end{aligned}$$

where in the third equality we used Lemma B.0.2 from Appendix B and the fact that under the null hypothesis $\mathbb{H} = \mathbb{P} \otimes \mathbb{Q}$. We examine the two resampling schemes proposed earlier $\mathbb{H}_n, \mathbb{P}_n \otimes \mathbb{Q}_n$.

Lemma 5.2.1. *The resampling schemes $R_n = \mathbb{H}_n$ and $R_n = \mathbb{P}_n \otimes \mathbb{Q}_n$ satisfy the conditions of Theorem 5.1.1.*

Proof. Both of these schemes have been shown in Chapter 3 to satisfy the conditions for CLT Under Sequences Theorem 2.3.1 with limit measures $\mathbb{H}, \mathbb{P}\otimes\mathbb{Q}$ respectively. It remains to derive the expression for T_n^* and the asymptotics of $\phi'_R(\mathbb{G}_R)$ for each one of those schemes. Denote $\hat{\mathbb{H}}_n$ the empirical measure of a joint Bootstrap sample (X, Y) distributed according to R_n. Denote $\hat{\mathbb{P}}_n, \hat{\mathbb{Q}}_n$ the empirical measures of the Bootstrap samples X and Y respectively.

Let $R_n = \mathbb{H}_n$. Then $\phi(\mathbb{H}_n) = \mathbb{H}_n - \mathbb{P}_n \otimes \mathbb{Q}_n$ according to the definition of ϕ and obviously $\phi(R_n) \overset{d}{=} \phi(\mathbb{H}_n)$, hence this is a GR scheme according to Definition 5.1.2. Since $R_n \to \mathbb{H}$, we get that $\phi_R(\mathbb{G}_R) = \phi'_{\mathbb{H}}(\mathbb{G}_{\mathbb{H}})$, which implies that Condition (5.2) is trivially satisfied. Hence the assumptions of Theorem 5.1.1 are satisfied for $R_n = \mathbb{H}_n$ and we obtain the consistent Bootstrap statistic:

$$\begin{aligned}T_n^* &= \sqrt{n}\left\|\phi(\hat{\mathbb{H}}_n) - \phi(R_n)\right\|\\ &= \sqrt{n}\left\|\phi(\hat{\mathbb{H}}_n) - \phi(\mathbb{H}_n)\right\|\\ &= \sqrt{n}\left\|\hat{\mathbb{H}}_n - \hat{\mathbb{P}}_n \otimes \hat{\mathbb{Q}}_n - \mathbb{H}_n + \mathbb{P}_n \otimes \mathbb{Q}_n\right\|\\ &= T_n^{*c}.\end{aligned}$$

Let $R_n = \mathbb{P}_n \otimes \mathbb{Q}_n$. Then $\phi(\mathbb{H}_n) = 0$ according to the definition of ϕ and obviously $\phi(R_n) \overset{d}{=} \phi(\mathbb{H}_n)$, hence we see that this is a NHR scheme according to Definition 5.1.2. It holds that $R_n = \mathbb{P}_n \otimes \mathbb{Q}_n \to \mathbb{P} \otimes \mathbb{Q}$. Lemma B.0.3 implies that under the null hypothesis $\phi'_{\mathbb{H}}(\mathbb{G}_{\mathbb{H}}) \overset{d}{=} \phi'_{\mathbb{P}\otimes\mathbb{Q}}(\mathbb{G}_{\mathbb{P}\otimes\mathbb{Q}})$. Hence the assumptions of Theorem 5.1.1 are satisfied for $R_n = \mathbb{P}_n \otimes \mathbb{Q}_n$ and we obtain the consistent Bootstrap statistic:

$$\begin{aligned}T_n^* &= \sqrt{n}\left\|\phi(\hat{\mathbb{H}}_n) - \phi(R_n)\right\|\\ &= \sqrt{n}\left\|\phi(\hat{\mathbb{H}}_n) - \phi(\mathbb{P}_n \otimes \mathbb{Q}_n)\right\|\\ &= \sqrt{n}\left\|\hat{\mathbb{H}}_n - \hat{\mathbb{P}}_n \otimes \hat{\mathbb{Q}}_n\right\|\\ &= T_n^{*eq},\end{aligned}$$

which concludes the proof. □

6

Simulation study

In this Chapter we present the simulation research that was conducted and led to the main Conjecture 1.6.1 and finally to the motivation to prove it. The experiments were performed for the settings of Regression Slope test and Independence test. The experiments are divided in two main parts.

The first part of the simulation includes generation of a random samples of size n that either satisfies the null hypothesis or does not. For this sample the true statistic T_n is calculated. Two resampling schemes are picked one from each of the two types (GR, NHR). B Bootstrap samples of size n are generated and for $r = 1, \ldots, B$ a Bootstrap statistic $T_n^{*(r)}$ is obtained, one from each of the two types T_n^{*eq}, T_n^{*c}. Finally, the histogram for each combination of Bootstrap schemes and Bootstrap statistic type is plotted.

The second part is related to the p-values and includes a part of the process from the previous part. Instead of generating one sample, many samples of size n are generated. Say s is the size of the collection of samples generated. For each sample, we generate B Bootstrap samples of size n and calculate the $\mathrm{pvalue}_{n,B}$. Hence s values for $\mathrm{pvalue}_{n,B}$ are calculated and stored. The histograms of the p-values are plotted. Naturally four histograms are obtained, one for each combination of Bootstrap scheme and Bootstrap statistic type.

The histograms display the four combinations in a way to distinguish between the consistent and inconsistent ones. Histograms with red color display the consistent Bootstrap processes while blue represent the inconsistent. In the Bootstrap statistics histograms the black line represents the value of the true statistic T_n, while in the $\mathrm{pvalue}_{n,B}$ histograms the black line represents the rejection level α which is set to 0.05.

6.1. Regression Test

In this section we present the simulation results for the Regression Slope test.

For the first part of the experiment we set the following parameters:

- $B = 100$,
- $n = 100$.

For the H_0 case, we generate a random intercept a between 0 and 100 and a random sample $\mathbf{X}_n$ from the uniform distribution. We also generate the random noise ϵ from the standard Gaussian distribution. Finally we set $b = 0$ and generate the sample $\mathbf{Y}_n$ as follows:

$$Y_i = a + b \cdot X_i + \epsilon. \tag{6.1}$$

The following quantities are calculated using the OLS:

- $\hat{b}$,
- $\hat{a}$,
- $\hat{\epsilon}_i$ for $i = 1, \ldots, n$.

The true statistic is set to be $T_n = |\hat{b}|$.

For the GR scheme case we perform the following procedure. B Bootstrap samples of size n are generated using the Non-Parametric Bootstrap from $X_i, \hat{\epsilon}_i$, and the OLS $\hat{b}$. More specifically we generate $X_i^{*(r)}, \hat{\epsilon}_i^{*(r)}$ and for every $i = 1, \ldots, n$, we set:

$$Y_i^{*(r)} = \hat{a} + \hat{b}X_i^{*(r)} + \hat{\epsilon}_i^{*(r)}. \tag{6.2}$$

Then the OLS On the Bootstrap sample returns $\hat{b}^*$. Finally we calculate for every r the Bootstrap statistics:

- $T_n^{*c} = |\hat{b}^* - \hat{b}|$,
- $T_n^{*eq} = |\hat{b}^*|$.

For the NHR scheme we perform the following procedure instead. B Bootstrap samples of size n are generated using the Non-Parametric Bootstrap from $X_i, \hat{\epsilon}_i$, and the a slope of zero. More specifically we generate $X_i^{*(r)}, \hat{\epsilon}_i^{*(r)}$ and for every $i = 1, \ldots, n$, we set:

$$Y_i^{*(r)} = \hat{a} + 0 \cdot X_i^{*(r)} + \hat{\epsilon}_i^{*(r)}. \tag{6.3}$$

Then the OLS On the Bootstrap sample returns $\hat{b}^*$. Finally we calculate for every r the Bootstrap statistics:

- $T_n^{*c} = |\hat{b}^* - \hat{b}|$,
- $T_n^{*eq} = |\hat{b}^*|$.

The histograms are displayed in Figure 6.1.

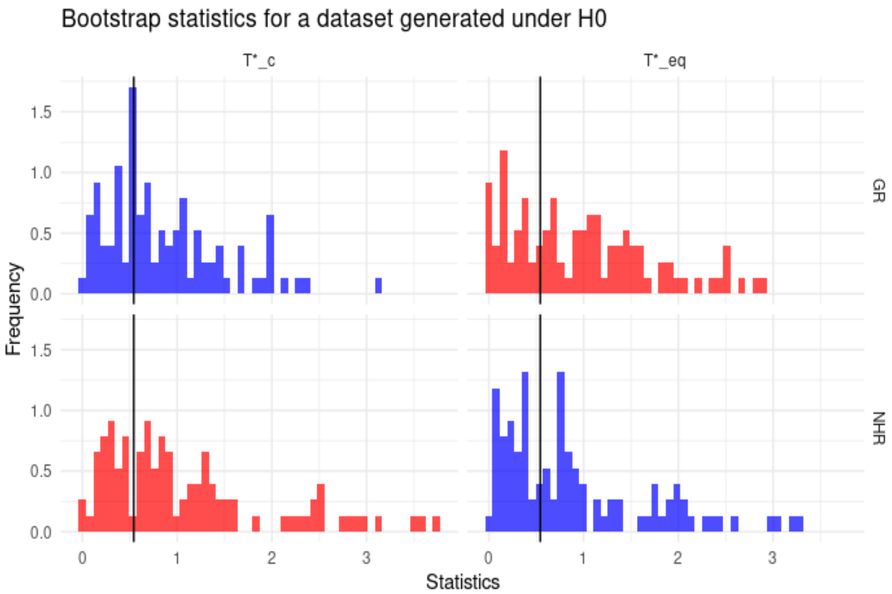


Figure 6.1: Regression Slope Test - Histogram of the Bootstrap statistics when H_0 is true .Blue color histograms represent consistent Bootstrap processes. Red color histograms represent inconsistent Bootstrap.

For the H_1 case, we generate a random intercept a between 0 and 100 and a random sample $\mathbf{X}_n$ from the uniform distribution. We also generate the random noise ϵ from the standard Gaussian distribution. Finally we set $b = 10$ and generate the sample $\mathbf{Y}_n$ as follows:

$$Y_i = a + b \cdot X_i + \epsilon. \tag{6.4}$$

The same statistical quantities $\hat{b}, \hat{a}, \hat{\epsilon}_i$are calculated using the OLS. The true statistic is set to be $T_n = |\hat{b}|$. We then pick the same Boostrap schemes to perform the simulations and obtain the results displayed in Figure 6.2

We now proceed to the second part of the experiment and set the following parameters:

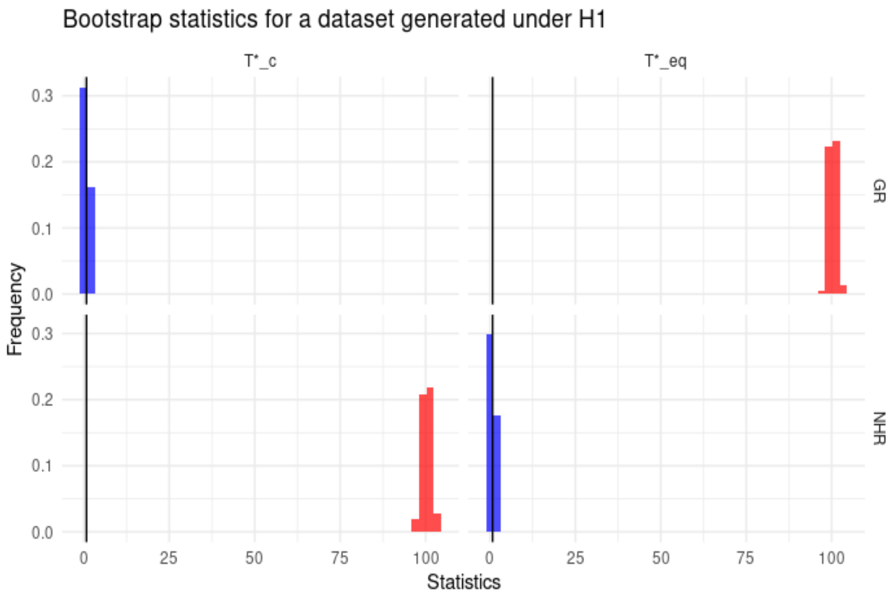


Figure 6.2: Regression Slope Test - Histogram of the Bootstrap statistics when H_0 is not true. Blue color histograms represent consistent Bootstrap processes. Red color histograms represent inconsistent Bootstrap.

- $s = 1000$,
- $B = 100$,
- $n = 100$.

In the H_0 case we generate B Bootstrap samples of size n based on both the GR and NHR schemes and calculated the T_n^{*c}, T_n for each case. Finally, for each combination we calculate the $\text{pvalue}_{n,B}$. This simulation is performed for s iterations and the $\text{pvalue}_{n,B}$ (for each one of the combinations of Bootstrap scheme and type of statistic) are appended to a list. The histograms are displayed in Figure 6.3.

We observe the blue colored p-values tending to resemble to a uniform distribution. These combinations represent the consistent Bootstrap processes.

We follow the same procedure for the H_1 case. The histograms are displayed in Figure 6.4. We observe that in this case following the blue colored p-values we would have correctly rejected the null hypothesis every single time, while we would have falsely accepted the null hypothesis every time in the rest two of the cases.

6.2. Independence Test

In this section we present the simulation results for the Independence test.

For the first part of the experiment we set the following parameters:

- $B = 100$,
- $n = 100$.

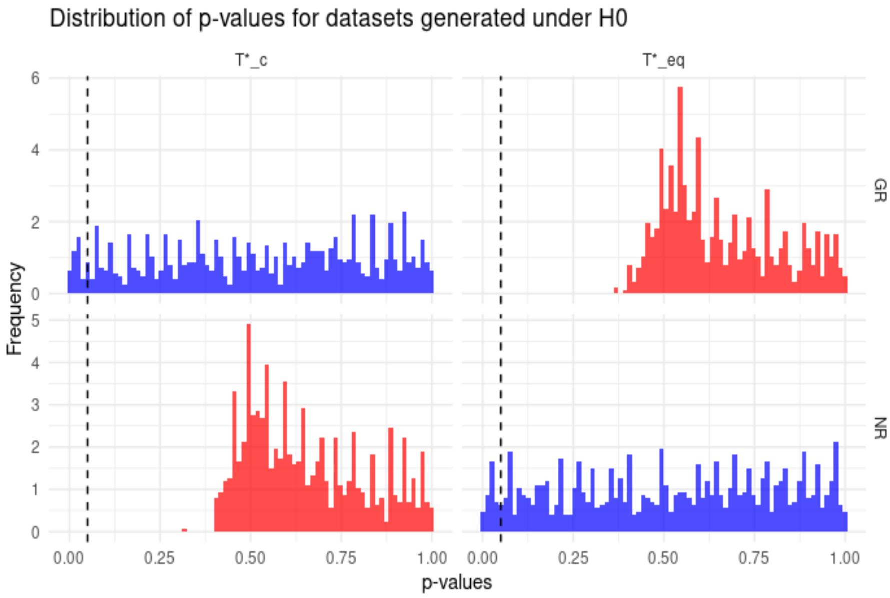
For the H_0 case, we generate two independent random samples $\mathbf{X}_n, \mathbf{Y}_n$ from the normal distribution and calculate the true statistic $T_n = |\hat{F}_{X,Y} - \hat{F}_X\hat{F}_Y|$, where $\hat{F}$ is the empirical distribution.

For the GR scheme we perform the Non-Parametric Bootstrap on the joint sample, namely we simulate from $\mathbb{H}_n$, and generate B Bootstrap samples of size n. For the NHR scheme we simulate from $\mathbb{P}_n \otimes \mathbb{Q}_n$ instead and calculate the Bootstrap statistics. The two types of statistics T_n^{*c}, T_n^{*eq} are calculated for $r = 1, \ldots, B$ for each one of the resampling schemes and the results are displayed in Figure 6.5.

For the H_1 case, we generate a random sample $\mathbf{X}_n$ from the uniform distribution and a Gaussian noise sample denoted by ϵ. Then we create the $\mathbf{Y}_n$ sample the following way:

$$Y = X + \epsilon. \tag{6.5}$$

Figure 6.3: Regression Slope Test - Histogram of the Bootstrap p-values when H_0 is true. Blue color represent consistent Bootstrap processes. Red color histograms represent inconsistent Bootstrap.



The true statistic T_n is calculated. Then, the same Bootstrap schemes are used to generate Bootstrap samples. The histograms for this simulation are displayed in Figure 6.6.

For the second part of the experiment we set $s = 1000, B = 100, n = 100$. We perform the simulations for both the null and alternative hypothesis accordingly to the previous settings describe. The results for the H_0 case are displayed in Figure 6.7, while the results for the alternative are displayed in Figure 6.8.

Figure 6.4: Regression Slope Test - Histogram of the Bootstrap p-values when H_0 is not true. Blue color represent consistent Bootstrap processes. Red color histograms represent inconsistent Bootstrap.

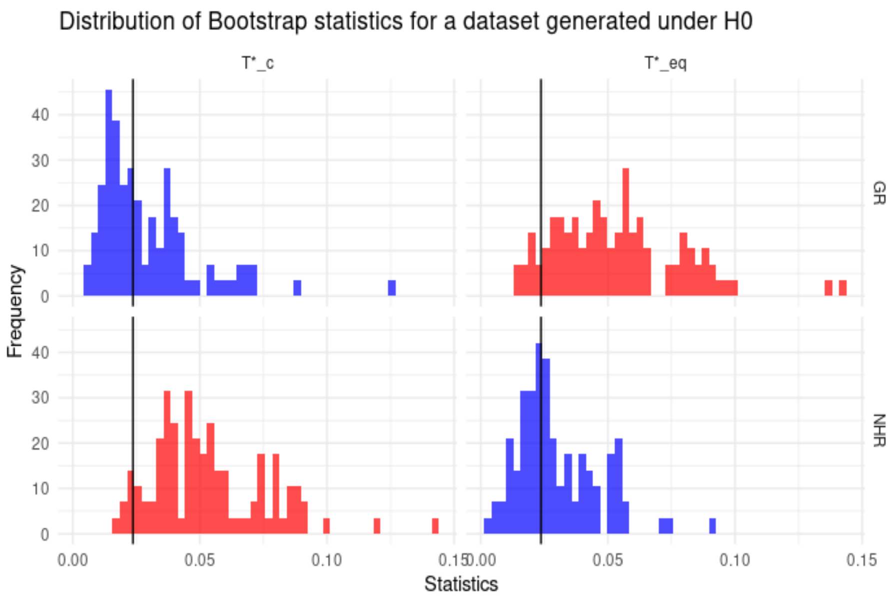


Figure 6.5: Independence Test - Histogram of the Bootstrap statistics when H_0 is true. Blue color histograms represent consistent Bootstrap processes. Red color histograms represent inconsistent Bootstrap.

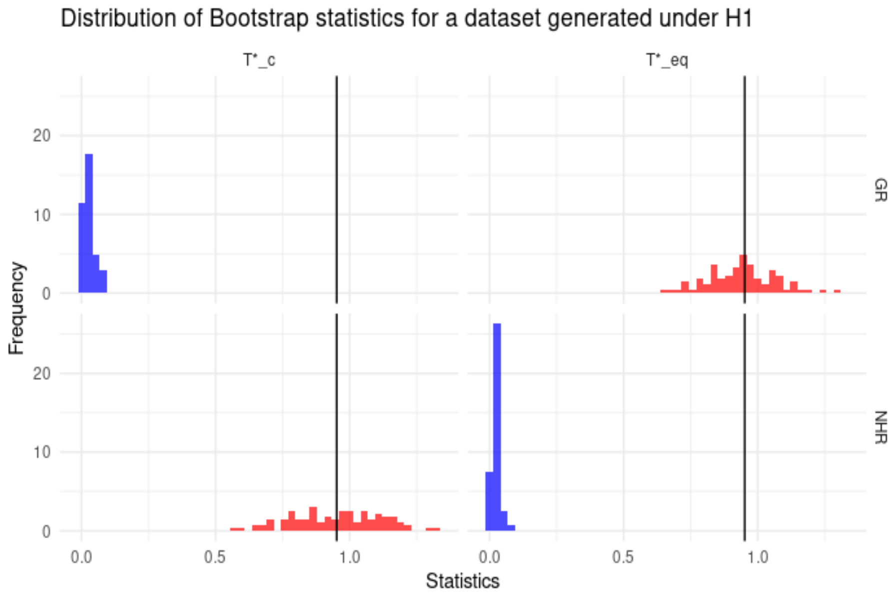
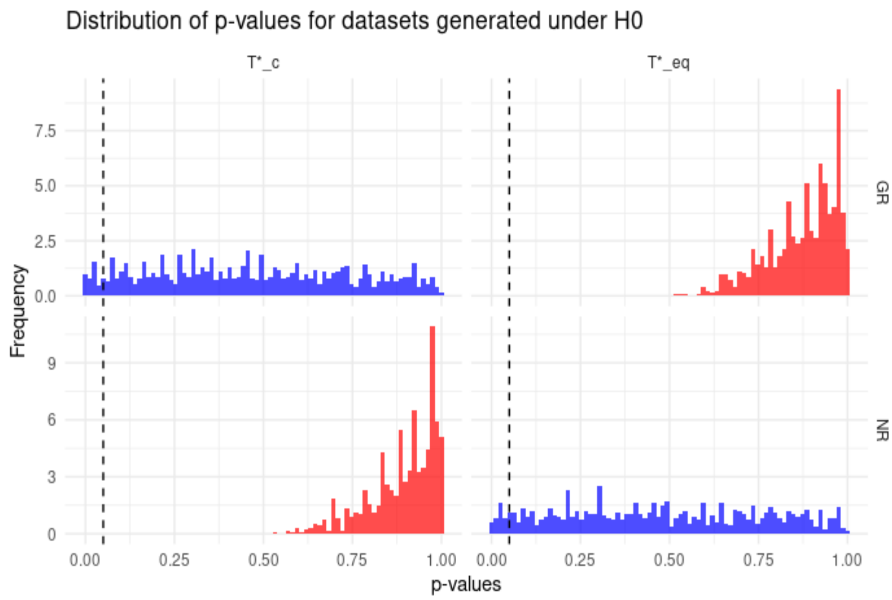


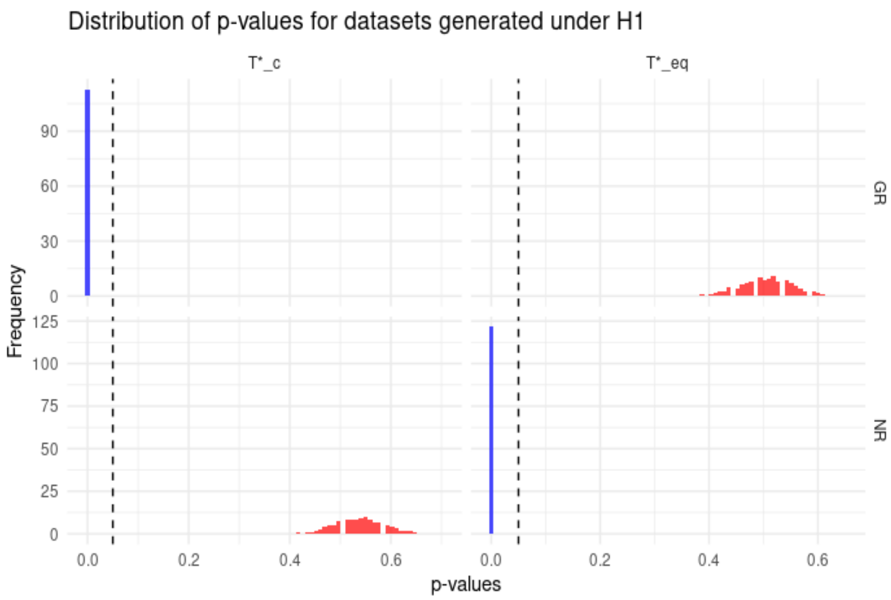
Figure 6.6: Independence Test - Histogram of the Bootstrap statistics when H_0 is true. Blue color histograms represent consistent Bootstrap processes. Red color histograms represent inconsistent Bootstrap.

Figure 6.7: Independence Test - Histogram of the Bootstrap p-values when H_0 is true



Blue color represent consistent Bootstrap processes. Red color histograms represent inconsistent Bootstrap.

Figure 6.8: Independence Test - Histogram of the Bootstrap p-values when H_0 is true



Blue color represent consistent Bootstrap processes. Red color histograms represent inconsistent Bootstrap.

7 Conclusion and Discussion

Within this project, we have delved into fundamental properties of consistent Bootstrap processes in the context of hypothesis testing. Persistent gaps in theoretical understanding have remained in the literature concerning resampling and testing. Our endeavor in this project was aimed at shedding light on the factors that contribute to the construction of a robust hypothesis test with both accurate statistical significance and power, utilizing the Bootstrap. Our initial motivation, sparked by Conjecture 1.6.1, was to gain insights into how the choice of a resampling scheme interacts with Bootstrap statistics and how it influences the significance level of the hypothesis test. Prior research has already touched upon the implications of this conjecture. For instance, in [7], the centered statistic T_n^{*c} is introduced as a practical tool for implementing Bootstrap in the Goodness of Fit test. Similarly, in [10], the T_n^{*eq} statistic is demonstrated to be consistent under a simulated measure of Null Resampling in the context of the Independence test. Additionally, in [6], the T_n^{*eq} statistic is shown to be consistent under a Null Resampling Bootstrap scheme. While these observations and results are well-documented, the absence of universality between the choice of T_n^{*c} and T_n^{*eq} may raise questions within the field of statistics.

We successfully distinguished the varying behaviors of these statistics, both in terms of their effectiveness and failure in maintaining the desired test level. We demonstrated that, in the specific test settings we have examined, each Bootstrap statistic performs well only when paired with its corresponding resampling scheme. This idea initially surfaces in Van Der Vaart's proof of consistency for the statistic T_n^{*eq} in the Independence test, as presented in [10]. By following the steps of the author's proof, we were able to establish an equivalent assertion to Conjecture 1.6.1 for T_n^{*c}. However, we acknowledge that we have yet to fully comprehend the underlying structural factors that lead to this outcome. The convergence of the independence processes $\hat{\mathbb{S}}_n$ and $\hat{\mathbb{T}}_n$ is demonstrated through algebraic calculations applied to the Central Limit Theorem under sequences for the Bootstrap empirical process.

In contrast, the Regression setting led our research in a different approach. The definition of the slope, denoted as b, and the definitions of the statistics naturally pointed toward the Delta Method as the most suitable direction. In this setting, we not only managed to validate Conjecture 1.6.1, but we also leveraged the Delta Method to directly illustrate the distinct behavior of the Bootstrap schemes. The representation of b through the mapping ϕ of the probability measure underlying the data offers a direct reflection of various measures through the image of ϕ. This representation serves not only as a natural plug-in estimator for b but also, through experimentation with the resampling schemes, allows us to develop a more comprehensive understanding of the NHR and GR schemes. Finally, in the case of NHR schemes, we managed to uncover precisely why the T_n^{*eq} statistic emerges as the consistent Bootstrap statistic, thanks to the condition $\phi(R_n) = 0$ and the extension of the Delta Method for sequences.

The approach in the Regression test has provided enlightening insights. In summary, by combining the Donsker Theorem and the Delta Method, we have successfully derived the asymptotic limit of the true statistic, represented by the Hadamard derivative of ϕ calculated within the original measure $\mathbb{H}$, along with its associated Brownian Bridge $\mathbb{G}_{\mathbb{H}}$. Furthermore, we have ensured that the selected resampling scheme aligns with the necessary assumptions for Theorem 2.3.1, including the Lindeberg

Condition. It's important to note that within the existing literature, the sequences $\mathbb{H}_n$ and $\mathbb{P}_n \otimes \mathbb{Q}_n$ (as defined in this project) have been demonstrated to satisfy these conditions and can be effectively employed. When considering a specific resampling scheme R_n, we have effectively utilized the Central Limit Theorem under sequences and the Delta Method for sequences to establish the asymptotic limit of the bootstrap empirical process, which is transformed through the application of ϕ. The equality of Hadamard derivatives under the null hypothesis serves as the condition that guarantees the consistency of the Bootstrap process generated by R_n. Furthermore, we have framed the hypothesis test as an evaluation of whether $\mathbb{H}$ functions as a root of ϕ. This approach has streamlined the assumptions and statements of the original test, effectively incorporating them into the framework of the Hadamard differentiable mapping ϕ.

Naturally, we are left wondering whether it is feasible to achieve a similar compression of the independence test through a comparable mapping. Evidently, as demonstrated in Section 5.2, this is indeed possible. In addition to the quest for a differentiable mapping that transforms the setting into a root test, we also require this mapping to exhibit uniform differentiability, a characteristic observed in the regression setting. Fortunately, this requirement is not overly demanding. Our observation that the independence test shares precisely the same structural framework as the regression test has led to the formulation of Theorem 5.1.1.

Theorem 5.1.1 represents a generalization of hypothesis test setups to a more inclusive framework. We have demonstrated that the two primary settings under scrutiny in this project can be simplified into this comprehensive framework, serving as particular instances of this theorem. Additionally, we have offered insights into the construction of a reliable Bootstrap statistic. In essence, the consistent Bootstrap statistic is a direct equivalent of the empirical process resulting from the resampling scheme R_n, mapped through the test mapping of interest, ϕ.

It is particularly intriguing to observe that the tests we have investigated can all be distilled into a root test of this Hadamard differentiable mapping. In [6], within the specific context of NHR, T_n^{*eq} is demonstrated to maintain consistency in the Goodness of Fit test. This outcome aligns with the initial Conjecture 1.6.1 of this project. Consequently, this motivates further examination of the validity of the assumptions presented in Theorem 5.1.1 within this specific setting. Furthermore, exploring various other scenarios, such as the randomization model, holds significant interest. In summary, the existence of a smooth mapping ϕ that potentially characterizes a hypothesis test as a root test may position this test as a distinct case within the broader framework of Theorem 5.1.1, providing justification for the distinct behavior observed in various Bootstrap statistics under diverse resampling schemes.

From a theoretical standpoint, there is a keen interest in determining whether the assumptions of Theorem 5.1.1 represent the minimal requirements. It can be argued that some level of smoothness assumptions for the mapping ϕ are inherently necessary. The mappings that have been investigated thus far exhibit certain smoothness properties, which are essential for the application of both the classical Delta Method and the extended Delta Method for sequences. Given our predominant focus on Donsker classes of functions, where empirical processes converge to a Brownian Bridge, properties of ϕ facilitating the use of the extended Delta Method for sequences would likely be sufficient to extend the same analytical approach to other scenarios as well.

From an applied perspective, a key takeaway is that the consistent Bootstrap statistic invariably revolves around the computation of ϕ within the chosen resampling scheme R_n. Interestingly, when employing NHR schemes, this centering is precisely at zero. With this understanding in hand, statisticians are now equipped to calculate the consistent Bootstrap statistic in accordance with their selected resampling scheme. This insight paves the way for future research to determine which resampling scheme represents the optimal choice and what criteria should guide this selection. A power analysis, as demonstrated in [4], could potentially offer valuable insights into how rapidly the power function converges to 1 when the data does not align with the null hypothesis.

In conclusion, this project has provided valuable insights into the intricate relationship between resampling schemes and Bootstrap statistics within the context of hypothesis testing. By carefully examining the nuances of various test settings, we have uncovered the crucial importance of aligning the appropriate Bootstrap statistic with its corresponding resampling scheme. Theorems such as 5.1.1 have allowed us to abstract these insights into a broader framework, raising questions about the universality of these findings. The existence of a smooth mapping, such as ϕ, capable of framing hypothesis tests as root tests, has emerged as a central theme, offering a unifying perspective on Bootstrap statistics under different resampling schemes. From both theoretical and practical perspectives, this research paves the way for exploring the minimal assumptions required for such mappings and optimizing resampling scheme selection. As we move forward, these findings hold promise for advancing our understanding of hypothesis testing and enhancing the precision of statistical analyses in various contexts.

A

Regression Hadamard Derivative

Lemma A.0.1. *Let $\phi : \mathbb{D} \mapsto \mathbb{R}^d$, such that:*

$$\phi(H) = \frac{\mathbb{E}_H(XY) - \mathbb{E}_H(X)\mathbb{E}_H(Y)}{\mathbb{E}_H(X^2) - \mathbb{E}_H(X)^2}$$

Then ϕ is Uniformly Hadamard differentiable with:

$$\phi'_H(\epsilon) := \frac{\mathbb{E}_\epsilon(XY) - \mathbb{E}_H(X)\mathbb{E}_\epsilon(Y) - \mathbb{E}_\epsilon(X)\mathbb{E}_H(Y) - \phi(H)\mathbb{E}_\epsilon(X^2) + 2\phi(H)\mathbb{E}_H(X)\mathbb{E}_\epsilon(X)}{\mathbb{E}_H(X^2) - \mathbb{E}_H(X)^2}.$$

Proof.

$$\begin{aligned}
\phi(H+\epsilon) &= \frac{\mathbb{E}_{H+\epsilon}(XY) - \mathbb{E}_{H+\epsilon}(X)\mathbb{E}_{H+\epsilon}(Y)}{\mathbb{E}_{H+\epsilon}(X^2) - \mathbb{E}_{H+\epsilon}(X)^2} \\
&= \frac{\mathbb{E}_H(XY) + \mathbb{E}_\epsilon(XY) - (\mathbb{E}_H(X) + \mathbb{E}_\epsilon(X))\,(\mathbb{E}_H(Y) + \mathbb{E}_\epsilon(Y))}{\mathbb{E}_H(X^2) + \mathbb{E}_\epsilon(X^2) - (\mathbb{E}_H(X) + \mathbb{E}_\epsilon(X))^2} \\
&= \frac{\mathbb{E}_H(XY) + \mathbb{E}_\epsilon(XY) - \mathbb{E}_H(X)\mathbb{E}_H(Y) - \mathbb{E}_H(X)\mathbb{E}_\epsilon(Y) - \mathbb{E}_\epsilon(X)\mathbb{E}_H(Y) + o(\epsilon)}{\mathbb{E}_H(X^2) + \mathbb{E}_\epsilon(X^2) - \mathbb{E}_H(X)^2 - 2\mathbb{E}_H(X)\mathbb{E}_\epsilon(X) + o(\epsilon)} \\
&= \frac{\mathbb{E}_H(XY) + \mathbb{E}_\epsilon(XY) - \mathbb{E}_H(X)\mathbb{E}_H(Y) - \mathbb{E}_H(X)\mathbb{E}_\epsilon(Y) - \mathbb{E}_\epsilon(X)\mathbb{E}_H(Y) + o(\epsilon)}{\left(\mathbb{E}_H(X^2) - \mathbb{E}_H(X)^2\right)\left(1 + \frac{\mathbb{E}_\epsilon(X^2) - 2\mathbb{E}_H(X)\mathbb{E}_\epsilon(X)}{\mathbb{E}_H(X^2) - \mathbb{E}_H(X)^2} + o(\epsilon)\right)} \\
&= \frac{\mathbb{E}_H(XY) + \mathbb{E}_\epsilon(XY) - \mathbb{E}_H(X)\mathbb{E}_H(Y) - \mathbb{E}_H(X)\mathbb{E}_\epsilon(Y) - \mathbb{E}_\epsilon(X)\mathbb{E}_H(Y) + o(\epsilon)}{\mathbb{E}_H(X^2) - \mathbb{E}_H(X)^2} \\
&\times \left(1 - \frac{\mathbb{E}_\epsilon(X^2) - 2\mathbb{E}_H(X)\mathbb{E}_\epsilon(X)}{\mathbb{E}_H(X^2) - \mathbb{E}_H(X)^2} + o(\epsilon)\right) \\
&= \phi(H) + \phi'_H(\epsilon) + \Xi_H(\epsilon) + o(\epsilon^2),
\end{aligned}$$

where

$$\phi'_H(\epsilon) := \frac{\mathbb{E}_\epsilon(XY) - \mathbb{E}_H(X)\mathbb{E}_\epsilon(Y) - \mathbb{E}_\epsilon(X)\mathbb{E}_H(Y) - \phi(H)\mathbb{E}_\epsilon(X^2) + 2\phi(H)\mathbb{E}_H(X)\mathbb{E}_\epsilon(X)}{\mathbb{E}_H(X^2) - \mathbb{E}_H(X)^2}$$

and:

$$\Xi_H(\epsilon) = -\frac{(\mathbb{E}_\epsilon(XY) - \mathbb{E}_H(X)\mathbb{E}_\epsilon(Y) - \mathbb{E}_\epsilon(X)\mathbb{E}_H(Y)) \cdot (\mathbb{E}_\epsilon(X^2) - 2\mathbb{E}_H(X)\mathbb{E}_\epsilon(X))}{(\mathbb{E}_H(X^2) - \mathbb{E}_H(X)^2)^2}$$

We see that ϕ'_H is linear with respect to ϵ, hence the Hadamard derivative is obtained. The quantity $\Xi(\epsilon)$ implies that ϕ is also Uniformly Hadamard differentiable. Indeed:

$$\sqrt{n}(\phi(R_n + \frac{1}{\sqrt{n}}h_n) - \phi(R_n)) = \sqrt{n}\phi_{R_n}\left(\frac{1}{\sqrt{n}}h_n\right) + \sqrt{n}\Xi_{\mathbb{H}_n}\left(\frac{1}{\sqrt{n}}h_n\right) + \sqrt{n}o\left(\frac{1}{n}h_n^2\right)$$

It easy to show that for $\epsilon = \frac{1}{\sqrt{n}} h_n$ and $H = \mathbb{H}_n$, $\sqrt{n}\Xi_{\mathbb{H}_n}\left(\frac{1}{\sqrt{n}} h_n\right) \to 0$, hence ϕ is Uniformly differentiable.

□

Lemma A.0.2. *Under the null hypothesis $\phi(\mathbb{H}) = 0$:*

$$\phi'_{\mathbb{H}}(\mathbb{G}_{\mathbb{H}}) \stackrel{d}{=} \phi'_{\mathbb{P}\otimes\mathbb{Q}}(\mathbb{G}_{\mathbb{P}\otimes\mathbb{Q}}).$$

Proof. Under the null hypothesis $\phi(\mathbb{H}) = 0$:

$$\begin{aligned}\phi'_{\mathbb{H}}(\mathbb{G}_{\mathbb{H}}) &:= \frac{\mathbb{E}_{\mathbb{G}_{\mathbb{H}}}(XY) - \mathbb{E}_{\mathbb{H}}(X)\mathbb{E}_{\mathbb{G}_{\mathbb{H}}}(Y) - \mathbb{E}_{\mathbb{G}_{\mathbb{H}}}(X)\mathbb{E}_{\mathbb{H}}(Y)}{\mathbb{E}_{\mathbb{H}}(X^2) - \mathbb{E}_{\mathbb{H}}(X)^2} \\ &= \frac{\int XYd(\mathbb{G}_{\mathbb{H}} - \mathbb{H} \otimes \mathbb{G}_{\mathbb{H}} - \mathbb{G}_{\mathbb{H}} \otimes \mathbb{H})}{\mathbb{E}_{\mathbb{H}}(X^2) - \mathbb{E}_{\mathbb{H}}(X)^2} \\ &= \frac{\int XYd(\mathbb{G}_{\mathbb{H}} - \mathbb{P} \otimes \mathbb{G}_{\mathbb{Q}} - \mathbb{G}_{\mathbb{P}\otimes\mathbb{Q}})}{\mathbb{E}_{\mathbb{P}}(X^2) - \mathbb{E}_{\mathbb{P}}(X)^2}\end{aligned}$$

Claim:$\int Xd\mathbb{G}_{\mathbb{H}} \stackrel{d}{=} \int Xd\mathbb{G}_{\mathbb{P}}$ and $\int Yd\mathbb{G}_{\mathbb{H}} \stackrel{d}{=} \int Yd\mathbb{G}_{\mathbb{Q}}$. Indeed:

$$\begin{aligned}Var\left(\int Xd\mathbb{G}_{\mathbb{H}}\right) &= \mathbb{E}(\mathbb{G}_{\mathbb{H}}X\mathbb{G}_{\mathbb{H}}X) \\ &= \mathbb{H}(X^2) - \mathbb{H}(X)^2 \\ &= Var_{\mathbb{H}}(X) \\ &= Var_{\mathbb{P}}(X) \\ &= Var\left(\int Xd\mathbb{G}_{\mathbb{P}}\right),\end{aligned}$$

where we used the properties of the Brownian Bridge:

$$\begin{aligned}\mathbb{E}(\mathbb{G}_{\mathbb{P}}f_1\mathbb{G}_{\mathbb{P}}f_2) &= \mathbb{P}(f_1 - \mathbb{P}f_1)(f_2 - \mathbb{P}f_2) = \mathbb{P}f_1f_2 - \mathbb{P}f_1\mathbb{P}f_2, \\ \mathbb{E}(\mathbb{G}_{\mathbb{P}}f_1) &= 0.\end{aligned}$$

The assertion for Y is proven equivalently.

Similarly:

$$\phi'_{\mathbb{P}\otimes\mathbb{Q}}(\mathbb{G}_{\mathbb{P}\otimes\mathbb{Q}}) = \frac{\int XYd(\mathbb{G}_{\mathbb{P}\otimes\mathbb{Q}} - \mathbb{P} \otimes \mathbb{G}_{\mathbb{Q}} - \mathbb{G}_{\mathbb{P}} \otimes \mathbb{Q})}{\mathbb{E}_{\mathbb{P}}(X^2) - \mathbb{E}_{\mathbb{P}}(X)^2}$$

It suffices to show that $\int XYd\mathbb{G}_{\mathbb{H}} \stackrel{d}{=} \int XYd\mathbb{G}_{\mathbb{P}\otimes\mathbb{Q}}$. We use again the property:

$$E(\mathbb{G}_Pf_1\mathbb{G}_Pf_2) = Pf_1f_2 - Pf_1Pf_2.$$

Set $f_1 = f_2 = XY$. Let $\mu = E_{\mathbb{P}}(X), \sigma^2 = Var(X), \sigma_\epsilon^2 = E(\epsilon^2)$. Under the null hypothesis $\phi(\mathbb{H}) = 0$ and $Y = a + \epsilon$.

$$\begin{aligned}Var\left[\int XYd\mathbb{G}_{\mathbb{H}}\right] &= Var\left(\mathbb{G}_H(XY)\right) \\ &= \mathbb{E}_{\mathbb{H}}\left(\mathbb{G}_{\mathbb{H}}(XY)\mathbb{G}_H(XY)\right) - \mathbb{E}_{\mathbb{H}}(\mathbb{G}_{\mathbb{H}}(XY))^2 \\ &= \mathbb{H}(X^2Y^2) - \mathbb{H}(XY)^2 \\ &= E_{\mathbb{H}}(X^2Y^2) - (E_{\mathbb{H}}(XY))^2 \\ &= E_{\mathbb{H}}(X^2E_{\mathbb{H}}(Y^2|X)) - (E_{\mathbb{H}}(XE_{\mathbb{H}}(Y|X)))^2 \\ &= E_{\mathbb{H}}(X^2(a^2 + \sigma_\epsilon^2))) - (E_{\mathbb{H}}(Xa))^2 \\ &= (a^2 + \sigma_\epsilon^2)(\mu^2 + \sigma^2) - a^2\mu^2,\end{aligned}$$

and

$$
\begin{aligned}
Var(\int XYd\mathbb{G}_{\mathbb{P}\otimes\mathbb{Q}}) &= Var(\mathbb{G}_{\mathbb{P}\otimes\mathbb{Q}}(XY)) \\
&= \mathbb{E}(\mathbb{G}_{\mathbb{P}\otimes\mathbb{Q}}(XY)\mathbb{G}_{\mathbb{P}\otimes\mathbb{Q}}(XY)) - \mathbb{E}(\mathbb{G}_{\mathbb{P}\otimes\mathbb{Q}}(XY))^2 \\
&= (\mathbb{P}\otimes\mathbb{Q})(X^2Y^2) - (\mathbb{P}\otimes\mathbb{Q})(XY)^2 \\
&= \mathbb{E}_{\mathbb{P}\otimes\mathbb{Q}}(X^2Y^2) - (\mathbb{E}_{\mathbb{P}\otimes\mathbb{Q}}(XY))^2 \\
&= \mathbb{E}_{\mathbb{P}}(X^2)\mathbb{E}_{\mathbb{Q}}(Y^2) - (\mathbb{E}_{\mathbb{P}}(X)\mathbb{E}_{\mathbb{Q}}(Y))^2 \\
&= (a^2 + \sigma_\epsilon^2)(\mu^2 + \sigma^2) - a^2\mu^2,
\end{aligned}
$$

which concludes the proof.

□

B

Independence Hadamard Derivative

Lemma B.0.1. *Let $\phi : \mathbb{D} \to \mathbb{E}$, such that:*

$$\phi(\mathbb{H}) = \mathbb{H} - \mathbb{H}(\cdot \times \mathbb{R}) \otimes \mathbb{H}(\mathbb{R} \times \cdot).$$

Then ϕ is Uniformly Hadamard differentiable with:

$$\phi'_H(\epsilon) = \epsilon - H(\cdot \times \mathbb{R}) \otimes \epsilon(\mathbb{R} \times \cdot) - \epsilon(\cdot \times \mathbb{R}) \otimes H(\mathbb{R} \times \cdot)$$

Proof.

$$\phi(H) = H - P \otimes Q = H - H(\cdot \times \mathbb{R}) \otimes H(\mathbb{R} \times \cdot)$$

$$\begin{aligned}
\phi(H + \epsilon) &= H - H(\cdot \times \mathbb{R}) \otimes H(\mathbb{R} \times \cdot) \\
&= H + \epsilon - (H + \epsilon)(\cdot \times \mathbb{R}) \otimes (H + \epsilon)(\mathbb{R} \times \cdot) \\
&= \phi(H) + \epsilon - H(\cdot \times \mathbb{R}) \otimes \epsilon(\mathbb{R} \times \cdot) - \epsilon(\cdot \times \mathbb{R}) \otimes H(\mathbb{R} \times \cdot) + \Xi(\epsilon) + o(\epsilon^2) \\
&= \phi(H) + \phi'_H(\epsilon) + \Xi_H(\epsilon) + o(\epsilon^2),
\end{aligned}$$

where:

$$\phi'_H(\epsilon) = \epsilon - H(\cdot \times \mathbb{R}) \otimes \epsilon(\mathbb{R} \times \cdot) - \epsilon(\cdot \times \mathbb{R}) \otimes H(\mathbb{R} \times \cdot)$$

and:

$$\Xi_H(\epsilon) = \epsilon(\cdot \times \mathbb{R}) \otimes \epsilon(\mathbb{R} \times \cdot)$$

We see that ϕ'_H is linear with respect to ϵ, hence the Hadamard derivative is obtained. The quantity $\Xi(\epsilon)$ implies that ϕ is also Uniformly Hadamard differentiable. Indeed:

$$\sqrt{n}(\phi(R_n + \frac{1}{\sqrt{n}} h_n) - \phi(R_n)) = \sqrt{n}\phi_{R_n}\left(\frac{1}{\sqrt{n}} h_n\right) + \sqrt{n}\Xi_{\mathbb{H}_n}\left(\frac{1}{\sqrt{n}} h_n\right) + \sqrt{n} o\left(\frac{1}{n} h_n^2\right)$$

It easy to show that for $\epsilon = \frac{1}{\sqrt{n}} h_n$ and $H = \mathbb{H}_n$, $\sqrt{n}\Xi_{\mathbb{H}_n}\left(\frac{1}{\sqrt{n}} h_n\right) \to 0$, hence ϕ is Uniformly differentiable. □

Lemma B.0.2. *Denote $\mathbb{G}_{\mathbb{H}}, \mathbb{G}_{\mathbb{P}}, \mathbb{G}_{\mathbb{Q}}$ Brownian Bridges where $\mathbb{G}_{\mathbb{H}}$ is a two dimensional process while $\mathbb{G}_{\mathbb{P}}, \mathbb{G}_{\mathbb{Q}}$ are one dimensional processes. Denote $\mathbb{H}^1 = \mathbb{H}(\cdot \times \mathbb{R})$, $\mathbb{H}^2 = \mathbb{H}(\mathbb{R} \times \cdot)$ and $\mathbb{G}^1 = \mathbb{G}(\cdot \times \mathbb{R})$, $\mathbb{G}^2 = \mathbb{G}(\mathbb{R} \times \cdot)$. Under the null hypothesis $\mathbb{H} = \mathbb{P} \otimes \mathbb{Q}$, the following hold:*

$$\mathbb{G}_{\mathbb{H}}(\cdot \times \mathbb{R}) \overset{d}{=} \mathbb{G}_{\mathbb{P}}(\cdot),$$

$$\mathbb{G}_{\mathbb{H}}(\mathbb{R} \times \cdot) \overset{d}{=} \mathbb{G}_{\mathbb{Q}}(\cdot).$$

Proof. We show the claim for the first dimension and the other is proved identically. Since all of the processes are zero-mean Gaussian it suffices to show that the variances are equal. Let $f \in \mathcal{F}$.

$$\begin{aligned} Var(\mathbb{G}_{\mathbb{H}}^1 f) &= \mathbb{E}((\mathbb{G}_{\mathbb{H}}^1 f)^2) \\ &= \mathbb{H}^1(f^2) - \mathbb{H}^1(f)^2 \\ &= \mathbb{P}(f^2) - \mathbb{P}(f)^2 \\ &= Var(\mathbb{G}_{\mathbb{P}} f), \end{aligned}$$

where in the third equality we used that $\mathbb{H} = \mathbb{H}^1 \otimes \mathbb{H}^2 = \mathbb{P} \otimes \mathbb{Q}$. The proof is concluded. □

Lemma B.0.3. *Under the null hypothesis* $\phi(\mathbb{H}) = 0$*:*

$$\phi'_{\mathbb{H}}(\mathbb{G}_{\mathbb{H}}) \stackrel{d}{=} \phi'_{\mathbb{P}\otimes\mathbb{Q}}(\mathbb{G}_{\mathbb{P}\otimes\mathbb{Q}}).$$

Proof.

$$\begin{aligned} \phi'_{\mathbb{H}}(\mathbb{G}_{\mathbb{H}}) &= \mathbb{G}_{\mathbb{H}} - \mathbb{H}(\cdot \times \mathbb{R}) \otimes \mathbb{G}_{\mathbb{H}}(\mathbb{R} \times \cdot) - \mathbb{G}_{\mathbb{H}}(\cdot \times \mathbb{R}) \otimes \mathbb{H}(\mathbb{R} \times \cdot) \\ &= \mathbb{G}_{\mathbb{H}} - \mathbb{P}(\cdot \times \mathbb{R}) \otimes \mathbb{G}_{\mathbb{H}}(\mathbb{R} \times \cdot) - \mathbb{G}_{\mathbb{H}}(\cdot \times \mathbb{R}) \otimes \mathbb{Q}(\mathbb{R} \times \cdot). \end{aligned}$$

Lemma B.0.2 implies that under the null hypothesis $\mathbb{G}_{\mathbb{P}}(\cdot \times \mathbb{R}) \stackrel{d}{=} \mathbb{G}_{\mathbb{P}}(\cdot)$ and $\mathbb{G}_{\mathbb{H}}(\mathbb{R} \times \cdot) \stackrel{d}{=} \mathbb{G}_{\mathbb{Q}}(\cdot)$, thus:

$$\phi'_{\mathbb{H}}(\mathbb{G}_{\mathbb{H}}) = \mathbb{G}_{\mathbb{H}} - \mathbb{P}(\cdot) \otimes \mathbb{G}_{\mathbb{Q}}(\cdot) - \mathbb{G}_{\mathbb{P}}(\cdot) \otimes \mathbb{Q}(\cdot).$$

Similary we calculate:

$$\phi'_{\mathbb{P}\otimes\mathbb{Q}}(\mathbb{G}_{\mathbb{P}\otimes\mathbb{Q}}) = \mathbb{G}_{\mathbb{H}} - \mathbb{P}(\cdot) \otimes \mathbb{G}_{\mathbb{Q}}(\cdot) - \mathbb{G}_{\mathbb{P}}(\cdot) \otimes \mathbb{Q}(\cdot),$$

which concludes the proof. □

Bibliography

[1] Axel Bücher and Ivan Kojadinovic. “A note on conditional versus joint unconditional weak convergence in bootstrap consistency results”. In: *Journal of Theoretical Probability* 32.3 (2019), pp. 1145–1165.

[2] Victor Chernozhukov, Denis Chetverikov, and Kengo Kato. “Gaussian approximations and multiplier bootstrap for maxima of sums of high-dimensional random vectors”. In: *The Annals of Statistics* 41.6 (Dec. 2013). DOI: 10.1214/13-aos1161. URL: https://doi.org/10.1214%2F13-aos1161.

[3] Russell Davidson and Emmanuel Flachaire. “The wild bootstrap, tamed at last”. In: *Journal of Econometrics* 146.1 (2008), pp. 162–169. ISSN: 0304-4076. DOI: https://doi.org/10.1016/j.jeconom.2008.08.003. URL: https://www.sciencedirect.com/science/article/pii/S0304407608000833.

[4] Alexis Derumigny and Jean-David Fermanian. In: *Dependence Modeling* 5.1 (2017), pp. 154–197. DOI: doi:10.1515/demo-2017-0011. URL: https://doi.org/10.1515/demo-2017-0011.

[5] Rick Durrett. *Probability: theory and examples*. Vol. 49. Cambridge university press, 2019.

[6] Christian Genest and Bruno Rémillard. “Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models”. In: *Annales de l’Institut Henri Poincaré, Probabilités et Statistiques* 44.6 (2008), pp. 1096–1127. DOI: 10.1214/07-AIHP148. URL: https://doi.org/10.1214/07-AIHP148.

[7] Peter Hall and Susan R. Wilson. “Two guidelines for bootstrap hypothesis testing”. In: *Biometrics* 47 (1991), pp. 757–762.

[8] Michael A. Martin. “Bootstrap hypothesis testing for some common statistical problems: A critical evaluation of size and power properties”. In: *Computational Statistics and Data Analysis* 51.12 (2007), pp. 6321–6342. ISSN: 0167-9473. DOI: https://doi.org/10.1016/j.csda.2007.01.020. URL: https://www.sciencedirect.com/science/article/pii/S0167947307000230.

[9] Mark Trede. “Bootstrapping inequality measures under the null hypothesis: Is it worth the effort?” In: *Journal of Economics* 77 (2002), pp. 261–282.

[10] Aad W Van Der Vaart et al. *Weak convergence*. Springer, 1996.

[11] J F Walsh and M Reznikoff. “Bootstrapping: a tool for clinical research”. en. In: *J Clin Psychol* 46.6 (Nov. 1990), pp. 928–930.

[12] Stephen J Walters and Michael J Campbell. “The use of bootstrap methods for analysing Health-Related Quality of Life outcomes (particularly the SF-36)”. en. In: *Health Qual Life Outcomes* 2 (Dec. 2004), p. 70.