`RATest`: An `R` package for Randomization Tests with an application to testing the continuity of the baseline covariates in RDD using Approximate Permutation Tests

Mauricio Olivares-González [†]
Department of Economics
UIUC
lvrsgnz2@illinois.edu

Ignacio Sarmiento-Barbieri
Department of Economics
UIUC
srmntbr2@illinois.edu

July 13, 2021

Abstract

This paper introduces the `RATest` package in `R`, a collection of randomization tests. This package implements the approximate permutation test proposed by Canay and Kamat (2017) for testing the null hypothesis of continuity of the distribution of the baseline covariates at the cutoff in the Regression Discontinuity Design (RDD). We revisit the construction of permutation tests in general, and their properties under the approximate group invariance in particular. We illustrate these ideas and the proposed package in the context of the RDD of Lee (2008).

Keywords: Regression Discontinuity Design, Permutation Test, Induced Order Statistics, `R`.

1 Introduction

The fundamental hypothesis of continuity of the baseline covariates at the cutoff is intrinsically intestable. A stronger condition is proposed by Lee (2008), which leads to testable indentification hypotheses. Specifically, identification of the ATE requires that i) the distribution of the running variable is continuous at the cutoff, and ii) the continuity of the distribution of the baseline covariates at the cutoff.

This paper implements the ideas and methods of Canay and Kamat (2017), who propose a permutation test approach for the null hypothesis of continuity of the distribution of the baseline covariates at the cutoff. Permutation tests have several advantages in the testing problem we are concerned. They can be applied without parametric assumptions of the underlying distribution generating the data. They can control the limiting rejection probability under general assumptions.

The practical relevance of this continuity assumption is everywhere in the regression discontinuity empirical literature. In practice[1], though, the assessment of the validity of the RDD frequently relies on graphical inspection, or checking the continuity of the *conditional means* of the baseline covariates at the cutoff by means of formal test, neither of which in a test for the continuity of the baseline covariates at the cutoff.

This paper is organized as follows. Section 2 introduces the environment and testable identification assumptions in the RDD case. In Section 3 the induced order statistics environment and test statistics will be described. A brief introduction to permutation tests and their validity under certain group invariance assumptions is developed in Section 4. Section 4.3 establishes the asymptotic validity of the permutation test in Canay et al. (2017). Sections 5 presents technical details on the computation and the implementation of the proposed testing procedure in `RATest`. Section 6 concerns the empirical illustration and further examples. Finally, Section 7 concludes. Readers versed in the theoretical dimension of this problem can skip sections 2 to 4.3.

2 Testable Hypothesis

2.1 Potential Outcomes

Consider the simplest model for a randomized experiment with subject i's (continuous) response Y_i to a binary treatment A_i. The treatment assignment in the sharp RDD follows the rule $A_i = \{Z_i \geq \bar{z}\}$, where Z_i is the so called running variable, and $\bar{z}$ is the cutoff at which the discontinuity arises. This threshold is conveniently assumed to be equal to 0.

For every subject i, there are two mutually exclusive potential outcomes - either subject gets treated or not. If subject i receives the treatment ($A_i = 1$), we will say the potential outcome is $Y_i(1)$. Similarly, if subject i belongs to the control group ($A_i = 0$), the potential outcome is $Y_i(0)$. We are interested in the average treatment effect (ATE) at the cutoff, i.e.

$$\mathbb{E}(Y_i(1) - Y_i(0)|Z = 0)$$

The identification assumption is not testable nonetheless as we only get to observe at most one of the potential outcomes[2]. Lee (2008) established a more restrictive but testable sufficient condition for identification - units can control the running variable except around the cutoff[3]. The identifying assumption implies that the baseline covariates are continuously distributed at the cutoff

$$H(w|z) \ z = 0 \quad \text{for all} \ w \in \mathcal{W} \tag{1}$$

where $W \in \mathcal{W}$ denotes the baseline covariates. We can cast condition (1) in terms of a two-sample hypothesis testing problem. Let

$$H^-(w|0) = \lim_{z \uparrow 0} H(w|0) \quad \text{and} \quad H^+(w|0) = \lim_{z \downarrow 0} H(w|0)$$

---

[1]This has been highlighted by Canay and Kamat (2017). See Appendix E for a survey of the topic in leading journals from 2011 to 2015.

[2]To put it in a more compact way, we say individual i's observed outcome, Y_i^* is $Y_i^* = Y_i(1)A_i + Y_i(0)(1-A_i)$, whereas the identification assumption in Hahn et al. (2001) requires that both

$$\mathbb{E}(Y_i(1)|Z = z) \quad \text{and} \quad \mathbb{E}(Y_i(0)|Z = z) \ \text{are continuous in} \ z \ \text{at} \ 0$$

[3]See condition 2b in the aforementioned paper.

Condition (1) is equivalent to $H(w|z)$ being right continuous at $z = 0$ and

$$H^-(w|0) = H^+(w|0) \quad \text{for all} \ \ w \in \mathcal{W} \tag{2}$$

Therefore, testing the null hypothesis of continuity of the baseline covariates at the cutoff $Z = 0$ reduces to testing for condition (2).

Remark 2.1. In the empirical literature, a hypothesis of the form (2) is commonly replaced by a weaker hypothesis

$$\mathbb{E}(W|Z = z) \quad \text{is continuous in} \ \ z \ \ \text{at} \ \ z = 0$$

This poses some limitations, most notably that there might be distributions which conditional means are continuous, yet the some other features of the conditional distribution of W might be discontinuous. See Canay and Kamat (2017), appendix E for a thorough revision of this practice in the literature. ■

3 Induced Order Statistics

Consider a random sample $X^{(n)} = \{(Y_i^*, W_i, Z_i)\}_{i=1}^n$ from a distribution P of (Y^*, W, Z). The order statistics of the sample of the running variable, $Z_{(1)} \leq Z_{(2)} \leq \cdots \leq Z_{(n)}$ will *induce an order* in sample of the baseline covariate, say, $W_{[1]}, W_{[2]}, \ldots, W_{[n]}$ according to the rule: if $Z_{(j)} = Z_k$ then $W_{[j]} = W_k$ for all $k = 1, \ldots, n$. It is worth mentioning that the values of this induced order statistics are not necessarily ordered.

3.1 The test statistic

The test statistic exploits the behavior of the closest units to the left and right of the cutoff $\bar{z} = 0$. More precisely, fix $q \in \mathbb{N}$[4] and take the q closest values of the order statistics of $\{Z_i\}$ to the right, and the q closest values to the left:

$$Z_{(q)}^- \leq Z_{(q-1)}^- \leq \cdots \leq Z_{(1)}^-$$

and

$$Z_{(1)}^+ \leq Z_{(2)}^+ \leq \cdots \leq Z_{(q)}^+$$

respectively. The induced order for the baseline covariates is then

$$W_{[q]}^-, W_{[q-1]}^-, \ldots, W_{[1]}^-$$

and

$$W_{[1]}^+, W_{[2]}^+, \ldots, W_{[q]}^+$$

respectively. The random variabes

$$\{W_{[q]}^-, W_{[q-1]}^-, \ldots, W_{[1]}^-\}$$

---

[4]The number q has to be small, relative to n. More on this in the upcoming sections.

can be viewed as an independent sample of W conditional on Z being close to the cutoff from the left. Analogously,

$$\{W_{[1]}^+, W_{[2]}^+, \ldots, W_{[q]}^+\}$$

can be thought of an independent sample of W conditional on Z being close to the cutoff from the right. Let $H_n^-(w)$ and $H_n^+(w)$ be the empirical CDFs of the two samples of size q, respectively,

$$H_n^-(w) = \frac{1}{q} \sum_{i=1}^q I\{W_{[i]}^- \leq w\}$$

and

$$H_n^+(w) = \frac{1}{q} \sum_{i=1}^q I\{W_{[i]}^+ \leq w\}$$

Stack all the $2q$ observations of the baseline covariates into

$$S_n = (S_{n,1}, \ldots, S_{n,2q}) = (W_{[1]}^-,$$

$$\ldots, W_{[q]}^-, W_{[1]}^+, \ldots, W_{[q]}^+)$$

The test statistic is a Cramér-von Mises type test:

$$T(S_n) = \frac{1}{2q} \sum_{i=1}^{2q} \left(H_n^-(S_{n,i}) - H_n^+(S_{n,i}) \right)^2 \tag{3}$$

4   Permutation Test

The general theory of permutation tests is presented, following section 15 in Lehmann and Romano (2006). Specifically, how they control type I error if the randomization hypothesis holds.

4.1   Hueristic Introduction to Permutation Tests

The rationale and construction of the permutation tests dates back to Fisher (1934). In a nutshell, these tests arise from recomputing the test statistic over permutations of the data. Consider the Cramér-von Mises test statistic (3). Given S_n, the observed value of the test statistic is given by $T(S_n)$. Define $\mathbf{G}_{2q}$ as the group of permutations of $\{1, \ldots, 2q\}$ onto itself. Compute $T(\cdot)$ for all permutations π, i.e. $T(S_{n,\pi(1)}, \ldots, S_{n,\pi(2q)})$ for all $\pi \in \mathbf{G}_{2q}$, and order these values

$$T^{(1)}(S_n) \leq T^{(2)}(S_n) \leq \cdots \leq T^{(N)}(S_n)$$

where $N = (2q)!$ is the cardinality of $\mathbf{G}_{2q}$. Fix a nominal level $\alpha \in (0,1)$, and define $k = N - \lfloor N\alpha \rfloor$ where $\lfloor \nu \rfloor$ is the largest integer less than or equal to ν. Let $M^+(S_n)$ and $M^0(S_n)$ be the number of values $T^{(j)}(S_n)$, $j = 1, \ldots, N$, which are greater than $T^{(k)}(S_n)$ and equal to $T^{(k)}(S_n)$ respectively. Set

$$a(S_n) = \frac{\alpha N - M^+(S_n)}{M^0(S_n)}$$

Define the randomization test function $\varphi(S_n)$ as

$$\varphi(S_n) = \begin{cases} 1 & T(S_n) > T^{(k)}(S_n) \\ a(S_n) & T(S_n) = T^{(k)}(S_n) \\ 0 & T(S_n) < T^{(k)}(S_n) \end{cases}$$

Moreover, define the randomization distribution based on $T(S_n)$ as

$$\hat{R}_N(t) = \frac{1}{N} \sum_{\pi \in \mathbf{G}_{2q}} I\{T(S_{n,\pi(1)}, \ldots, S_{n,\pi(2q)}) \leq t\} \tag{4}$$

Hence, the permutation test rejects the null hypothesis (2) if $T(S_n)$ is bigger than the $1 - \alpha$ quantile of the randimization distribution (4).

4.2 Why does this construction work?

Permutation tests have favorable finite sample properties, i.e. their construction yields an exact level α test for a fixed sample size, provided the fundamental *randomization hypothesis* holds[5]. In order to understand the scope of this hypothesis, let $\mathbf{P}_0$ be the family of distributions $P \in \mathbf{P}$ satisfying the null hypothesis (2):

$$\mathbf{P}_0 = \{P \in \mathbf{P} : H^-(w|0) = H^+(w|0) \ \text{ for all } \ w \in \mathcal{W}\}$$

then, the randomization hypothesis says that $\mathbf{P}_0$ remains invariant under $\pi \in \mathbf{G}_{2q}$. For the sake of exposition, suppose that $(S_1, \ldots, S_{2q})$ were i.i.d. with CDF $H(w|0)$. Then the randomization hypothesis tell us that $(S_{\pi(1)}, \ldots, S_{\pi(2q)}) \overset{d}{=} (S_1, \ldots, S_{2q})$ for all permutations $\pi \in \mathbf{G}_{2q}$.

Hence, if the randomization hypothesis holds, the permutation test described in section 4.1 based on the Cramér-von Mises test statistic (3) is such that[6]

$$\mathbb{E}_P(\varphi(S_n)) = \alpha \quad \text{for all} \quad P \in \mathbf{P}_0$$

This hypothesis, however, is hard to sustain in the present context. The null hypothesis (2) does not guarantee that $(S_{\pi(1)}, \ldots, S_{\pi(2q)}) \overset{d}{=} (S_1, \ldots, S_{2q})$ for all permutations $\pi \in \mathbf{G}_{2q}$, because S_n is not i.i.d. from $H(w|0)$. See Remark 4.1 in Canay and Kamat (2017). That is, the group invariance property fails, which leads us to the approximate invariance described in section 4.3.

4.3 Asymptotic Validity in the (sharp) RD case

In the absence of the group invariance assumption, Canay and Kamat (2017) developed a framework to explore the validity of permutation tests for testing the hypothesis (2) under an *approximate* invariance assumption[7]. Rather than assuming $(S_{\pi(1)}, \ldots, S_{\pi(2q)}) \overset{d}{=} (S_1, \ldots, S_{2q})$ for all permutations $\pi \in \mathbf{G}_{2q}$, we know only require $S = (S_1, \ldots, S_{2q})$ to be invariant to $\pi \in \mathbf{G}_{2q}$, whereas S_n might not be. As a result, the permutation test will control the type I error asymptotically.

The following conditions suffice to establish asymptotic validity of the permutation test based on the structure of the rank test statistics[8]:

---

[5]See Lehmann and Romano (2006), definition 15.2.1.

[6]See theorem 15.2.1 in Lehmann and Romano (2006).

[7]This approach has was first proposed by Canay et al. (2017), where the finite group of transformations $\mathbf{G}$ consisted of sign changes. This asymptotic framework deviates from the one developed by Hoeffding (1952), and later extended by Romano (1990), and Chung and Romano (2013). See Canay and Kamat (2017) Remark 4.4.

[8]See Assumption 4.4 in Canay and Kamat (2017).

Assumption 4.1. If $P \in \mathbf{P}_0$, then

(i) $S_n = S_n(X^{(n)}) \overset{\mathrm{d}}{\to} S$ under P.

(ii) $(S_{\pi(1)}, \ldots, S_{\pi(2q)}) \overset{d}{=} (S_1, \ldots, S_{2q})$ for all $\pi \in \mathbf{G}_{2q}$.

(iii) S is a continuous random variable taking values in $\mathcal{S} \subset \mathbb{R}^{2q}$.

(iv) $T : \mathcal{S} \to \mathbf{R}$ is invariant to rank

■

Two comments are worth mentioning. First, Assumption 4.1 is strengthened in Canay and Kamat (2017) Assumption 4.1 in a way that is easier to interpret. This condition imposes certain restrictions in the context of the model as well. Specifically, it requires the baseline covariate W to be a scalar random variable that is continuously distributed conditional on $Z = 0$[9]. However, the multidimensional case is also taken into account. More of this in section 6. Second, Assumption 4.1 requires S to be continuously distributed, an assumption that is relaxed in Canay and Kamat (2017) assumption 4.5.

In spite the assumption stated here emphasizes the continuous case, the asymptotic validity of the permutation test follows regardless of whether the running variable, or the baseline covariate is continuous or discrete, scalar or vector. In other words, the permutation test based on the Cramér-von Mises test statistic (3) is asymptotically valid:

$$\mathbb{E}_P(\varphi(S_n)) \to \alpha \,, \text{ as } \ n \to \infty \ \text{ as long as } \ P \in \mathbf{P}_0$$

where $\varphi(\cdot)$ is constructed as in section 4.1. See Canay and Kamat (2017) theorem 4.2.

5 Implementation

5.1 Computing the p-values

We argued that the permutation test rejects the null hypothesis (2) if $T(S_n)$ is bigger than the $1 - \alpha$ quantile of the randimization distribution (4). Alternatively, we can define the p-value of a permutation test, $\hat{p}$, as

$$\hat{p} = \frac{1}{N} \sum_{\pi \in \mathbf{G}_{2q}} I\{T(S_{n,\pi(1)}, \ldots, S_{n,\pi(2q)}) \geq T(S_n)\} \tag{5}$$

where $T(S_n) = T(S_{n,1}, \ldots, S_{n,2q})$ is the observed sample, and N is the cardinality of $\mathbf{G}_N$. It can be shown[10]

$$P(\hat{p} \leq u) \leq u \quad \text{for all}$$
$$0 \leq u \leq 1,$$
$$P \in \mathbf{P}_0$$

therefore, the test that rejects when $\hat{p} \leq \alpha$ is level α.

---

[9]The discrete case is also addressed. See assumption 4.2, *ibid.*

[10]This section applied to randomization tests in general, not only to permutation tests. See Lehmann and Romano (2006), section 15.2, page 636.

5.2 Stochastic approximation

When computing the permutation distribution in (4), we often encounter the situation that the cardinality of $\mathbf{G}_{2q}$ might be large such that it becomes computationally prohibitive. In this situation, it is possible to approximate the p-values the following way. Randomly sample permutations π from $\mathbf{G}_{2q}$ with or without replacement. Suppose the sampling is with replacement, then $\pi_1, \ldots, \pi_N$ are i.i.d. and uniformly distributed on $\mathbf{G}_{2q}$. Then

$$\tilde{p} = \frac{1}{B}\left(1 + \sum_{i=1}^{B-1} I\{T(S_{n,\pi_i(1)}, \ldots, S_{n,\pi_i(2q)}) \geq T(S_n)\}\right) \tag{6}$$

is such that

$$P(\tilde{p} \leq u) \leq u \quad \text{for all} \quad 0 \leq u \leq 1, \ P \in \mathbf{P}_0 \tag{7}$$

where this P takes into account the randomness of $T(\cdot)$ and the sampling of the π_i. Like in the case developed in Section 5.1, the test that rejects when $\tilde{p} \leq \alpha$ is level α.

It is worth noticing that the approximation $\tilde{p}$ satisfies (7) regardless of B, although a bigger B will improve the approximation. As a matter of fact, $\tilde{p} - \hat{p} = o_p(1)$ as $B \to \infty$. The `RATest` package uses $B = 499$ by default.

5.3 Tuning parameter q

The implementation of the test statistic heavily relies on q, the number of closest values of the running variable to the left and right of the cutoff. This quantity is small relative to the sample size n, and remains fixed as $n \to \infty$. Canay and Kamat (2017) recomend the rule of thumb

$$q = \left\lceil f(0)\sigma_Z\sqrt{10 * (1 - \rho^2)}\frac{n^{3/4}}{\log n} \right\rceil \tag{8}$$

where $\lceil \nu \rceil$ is the smallest integer greater or equal to ν, $f(0)$ is the density if Z at zero, ρ is the coefficient of correlation W and Z, and σ_Z^2 is the variance of Z. Additionally, the authors have considered the following alternative rule of thumb

$$q^a = \left\lceil f(0)\sigma_Z\sqrt{1 - \rho^2}\frac{n^{0.9}}{\log n} \right\rceil \tag{9}$$

finding similar results. The main difference in these rules is that the latter grows more rapidly. These quantities are only rules of thumb and have to be seen under this light. Whether or not it's an optimal rule is something for further research. See Canay and Kamat (2017) section 3.1 for additional motivation.

5.3.1 Scalar Case

Equations (8)-(9) can be estimated from sample. Consider Equation (8) first. The feasible tuning parameter is

$$\hat{q} = \left\lceil \max\left\{ \min\left\{ \hat{f}_n(0)\hat{\sigma}_{n,z}\sqrt{10 * (1 - \hat{\rho}_n^2)}\frac{n^{3/4}}{\log n}, q_{UB}\right\}, q_{LB}\right\} \right\rceil \tag{10}$$

where $q_{LB} = 10$, and $q_{UB} = n^{0.9}/\log n$. The lower bound, q_{LB} represents situations in which the randomized and non-randomized versions of the permutation test differ, whereas the upper bound, q_{UB} guarantees the rate of convergence does not violate the formal results in Canay and Kamat (2017), theorem 4.1. The same reasoning applies if we replace q with q^a.

The density function $\hat{f}_n(\cdot)$ was estimated employing the univariate adaptative kernel density estimation *à la Silverman* (e.g. Portnoy and Koenker, 1989; Koenker and Xiao, 2002; Silverman, 1986), and the results were obtained directly from the **R** package **quantreg** (Koenker (2016)). Finally, ρ and σ_Z were estimated by their sample counterparts.

5.3.2 Vector Case

The rules of thumb in (8)-(9) are not quite suitable when W is a K-dimensional vector, since the variances and correlations are not scalars. Motivated by Canay and Kamat (2017), we will consider two cases. First, we are interested in testing (2) individually, i.e. testing for continuity of the baseline covariates one by one. In this case, the following algorithm applies. We estimate $\hat{q}$ (or $\hat{q}^a$) for each of the K baseline covariates as in section 5.3.1. When testing (2) for the j-th covariate, we will use $\hat{q}_j$ to determine the $\hat{q}_j$ closest values of the order statistics of $\{Z_i\}$ to the right and to the left of the cutoff:

$$Z^-_{(\hat{q}_j)} \leq \cdots \leq Z^-_{(1)} \quad \text{and} \quad Z^+_{(1)} \leq \cdots \leq Z^+_{(\hat{q}_j)}$$

Then, the induced order statistics for the j-th baseline covariate is

$$W^-_{j,[\hat{q}_j]}, \ldots, W^-_{j,[1]} \quad \text{and} \quad W^+_{j,[1]}, \ldots, W^+_{j,[\hat{q}_j]}$$

Second, we may want to test whether or not the *joint* distribution of the baseline covariates is continuos at the cutoff. In this case, we will estimate the tuning parameter q for each of the K baseline covariates as we just described, but we choose $\hat{q} = \min\{\hat{q}_1, \ldots, \hat{q}_K\}$ and calculate the order statistics

$$Z^-_{(\hat{q})} \leq \cdots \leq Z^-_{(1)} \quad \text{and} \quad Z^+_{(1)} \leq \cdots \leq Z^+_{(\hat{q})}$$

whereas the induced order statistics of the baseline covariate W is

$$W^-_{[\hat{q}]}, \ldots, W^-_{[1]} \quad \text{and} \quad W^+_{[1]}, \ldots, W^+_{[\hat{q}]}$$

5.4 Multidimensional Case

5.4.1 The max statistic

Testing the null hypothesis (2) is equivalent to testing

$$\mathbb{P}(c'W \leq w | Z = z) \text{ is continuous in } z \text{ at } 0 \text{ for all } w \in \mathbb{R} \text{ and all } c \in \mathbf{C} \qquad (11)$$

where $\mathbf{C} \equiv \{a \in \mathbb{R}^k : ||a|| = 1\}$. Let $\hat{\mathbf{C}} \subset \mathbf{C}$, then the max test statistic is

$$M(S_n) = \max_{c \in \hat{\mathbf{C}}} T(c'S_n) \qquad (12)$$

where the test statistic $T(\cdot)$ is the Cramér-von Mises test defined in (3). Following the empirical application in Canay and Kamat (2017), $\hat{\mathbf{C}}$ consists of a random sample of $100 - K$ elements from $\mathbf{C}$, plus the K canonical vectors.

6 Empirical Illustration

The empirical illustration is based on Lee's (2008) of the effect of party incumbency advantage in electoral outcomes. For comparative purposes we follow the same empirical study chosen by Canay and Kamat (2017).
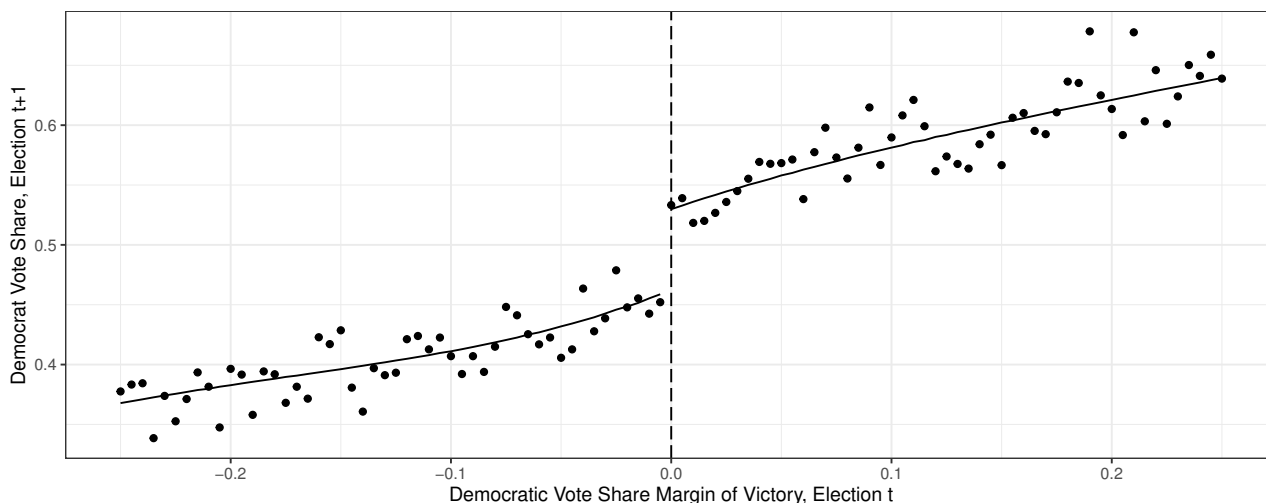
Figure 1: Candidate's probability of winning election $t + 1$, by margin of victory in election t: local averages and logit polynomial fit

The objective of Lee's (2008) is to assess whether a Democratic candidate of the US. House of Representative has an edge over his competitors if his party won the previous election. The causal effect of party incumbency is captured by exploiting the fact that an election winner is determined by $D = 1(Z \geq 0)$ where Z, the running variable, is the vote shares between Democrats and Republicans.

Figure 1 shows Lee (2008) sharp RD stgrategy. The figure illustrates the sharp change in probability of a Democrat winning against the difference in vote share in the preivous election. The data used here and contained in the package have six covariates and 6558 observations with information on the Democrat runner and the opposition. The data set is named `lee2008`, and it is a subset of the publicly available data set in the Mostly Harmless Econometrics Data Archive (http://economics.mit.edu/faculty/angrist/data1/mhe)

One check used by practioners to assess the credibility of the RD designs relies on graphical depiction of the conditional mean of the baseline covariates [11]. Figure 2 plots this for the Democrat vote share in $t-1$. A simple visual inspection would lead the researcher to conclude that there are no discontinuities at the cutoff for these baseline covariates.

This package however, implements Canay and Kamat (2017) in the function `RDperm`. The following `R` code performs the test for the continuity of the *Democrat Vote Share, Election t-1* named `demshareprev` in the data set at the threshold. The function requires the name of the baseline covariate to be tested, the running variable `z`, the data set name. We also specify a natural number that will define the q closest values of the order statistics of the running variable (`z`) to the right and to the left of the cutoff. As default, the function uses the Cramér-von Mises test 'CvM'. The function `summary` is available for a concise summary of the result.

```
R> # Lee2008
R> set.seed(101)
R> permtest<-RDperm(W="demshareprev", z="difdemshare",data=lee2008,q_type=51)
R> summary(permtest)



*********************************************************
**       RD Distribution Test using permutations       **
*********************************************************
```

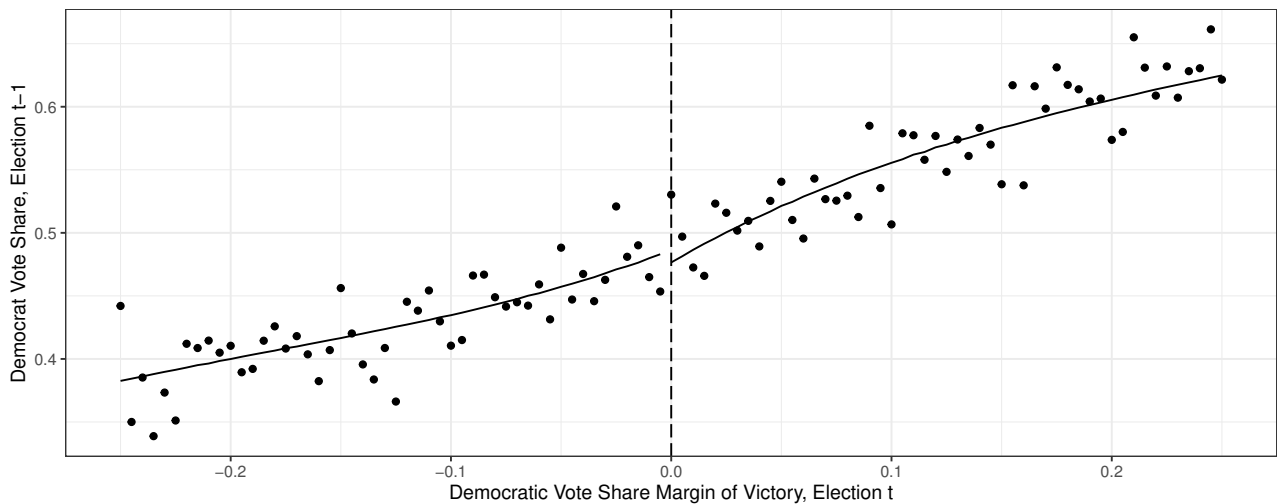[11]For a list in papers using this strategy see Canay and Kamat (2017) Section E: Surveyd papers on RDD

Figure 2: Candidate's probability of winning election $t + 1$, by margin of victory in election t: local averages and logit polynomial fit

```
Running Variable: difdemshare
Cutoff: 0
q: Defined by User
Test Statistic: CvM
Number of Permutations: 499
Number of Obs: 6558


**********************************************************
H0: 'Continuity of the baseline covariates at the cutoff'
**********************************************************


Estimates:
              T(Sn)  Pr(>|z|)  q
demshareprev  0.03   0         51   ***
---
Signif. codes:   0.01 '***' 0.05 '**' 0.1 '*'
```

The `summary` function reports the vaule of the test statistics ($T(Sn)$), the p-value and the number of q closest values used. This is particularly relevant when the user chooses any of the 'rule of thumb' methods for `q`. The function allows for multiple baseline covariates as well, in which case it will return the join test. The following `R` code shows how to do this, using the rule of thumb in Eq (8):

```
R> permtest_rot<-RDperm(W=c("demshareprev","demwinprev", "demofficeexp"),
+                        z="difdemshare",data=lee2008,q_type='rot', n.perm=600)
R> summary(permtest_rot)
```

```
**********************************************************
**        RD Distribution Test using permutations        **
**********************************************************
Running Variable: difdemshare
Cutoff: 0
q: Rule of Thumb
```
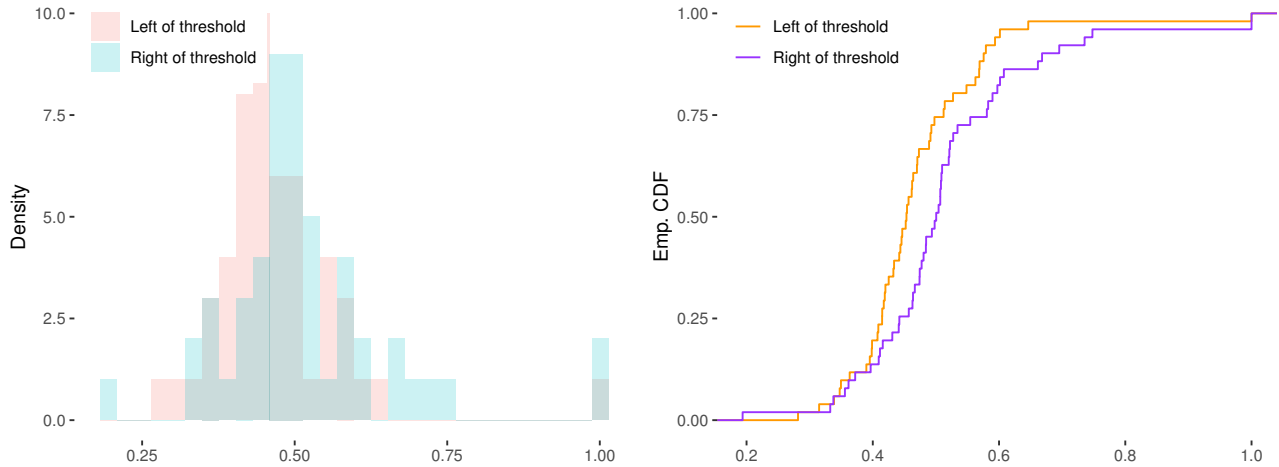
10

Figure 3: Histogram and CDF for Democrat vote share t-1

```
Test Statistic: CvM
Number of Permutations: 600
Number of Obs: 6558


***********************************************************
H0: 'Continuity of the baseline covariates at the cutoff'
***********************************************************


Estimates:
            T(Sn)  Pr(>|z|)  q
demshareprev  0.08  0        32  ***
demwinprev    0.08  0        35  ***
demofficeexp  0.06  0.01     45  **
Joint.Test    0.11  0        32  ***
---
Signif. codes:  0.01 '***' 0.05 '**' 0.1 '*'
```

A plot function is also available for objects of the class `RDperm`. It works as the base `plot` function, but it needs the specification of the desired baseline covariate to be plotted. The output can be a `ggplot` histogram (`hist`), CDF `cdf` or both. The default is `both`.

```
R> plot(permtest,w="demshareprev")
```

7   Conclusions

In this paper we describe the `RATest` package in `R`, which allows the practitioner to test the null hypothesis of continuity of the distribution of the baseline covariates in the RDD, as developed by Canay and Kamat (2017). Based on a result on induced order statistics, the `RATest` package implements a permutation test based on the Cramér-von Mises test statistic.

This paper also revisits the theory of permutation tests and the asymptotic framework to restore the validity of such procedures when an approximate group invariance assumption holds. Under this assumption, the aforementioned permutation test has several advantages, say, the

ability to control the type-I error in large samples, as well as its flexibility since we need not to assume a parametric distribution generating the data.

The main functionalities of the package have been illustrated by applying them to the celebrated RDD of the U.S. House elections in Lee (2008).

References

Canay, I. A. and Kamat, V. (2017). Approximate permutation tests and induced order statistics in the regression discontinuity design. Technical report, CeMMAP working paper CWP21/17, Centre for Microdata Methods and Practice.

Canay, I. A., Romano, J. P., and Shaikh, A. M. (2017). Randomization tests under an approximate symmetry assumption. *Econometrica*, 85(3):1013–1030.

Chung, E. and Romano, J. P. (2013). Exact and asymptotically robust permutation tests. *The Annals of Statistics*, 41(2):484–507.

Fisher, R. A. (1934). *Statistical methods for research workers*. Edinburgh.

Hahn, J., Todd, P., and Van der Klaauw, W. (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*, 69(1):201–209.

Hoeffding, W. (1952). The large-sample power of tests based on permutations of observations. *The Annals of Mathematical Statistics*, pages 169–192.

Koenker, R. (2016). *quantreg: Quantile Regression*. R package version 5.26.

Koenker, R. and Xiao, Z. (2002). Inference on the quantile regression process. *Econometrica*, 70(4):1583–1612.

Lee, D. S. (2008). Randomized experiments from non-random selection in us house elections. *Journal of Econometrics*, 142(2):675–697.

Lehmann, E. L. and Romano, J. P. (2006). *Testing statistical hypotheses*. Springer Science & Business Media.

Olivares-Gonzalez, M. and Sarmiento-Barbieri, I. (2017). *RATest: Randomization Tests*. R package version 0.1.0.

Portnoy, S. and Koenker, R. (1989). Adaptive l-estimation for linear models. *The Annals of Statistics*, pages 362–381.

Romano, J. P. (1990). On the behavior of randomization tests without a group invariance assumption. *Journal of the American Statistical Association*, 85(411):686–692.

Silverman, B. W. (1986). *Density estimation for statistics and data analysis*, volume 26. CRC press.