The package RATest contains a collection of randomization tests, data sets and examples. The current version focuses on two testing problems and their implementation in empirical work, mostly related to economics. First, it facilitates the empirical researcher to test for particular hypotheses, such as comparisons of means, medians, and variances from k populations using robust permutation tests, which asymptotic validity holds under very weak assumptions, while retaining the exact rejection probability in finite samples when the underlying distributions are identical. Second, it implements Canay and Kamat (2017) permutation test for testing the continuity assumption of the baseline covariates in the sharp regression discontinuity design (RDD).

In this post we’ll focus on the implementation of Canay and Kamat (2017) test presening a summarized version of the [vignette](https://cran.r-project.org/web/packages/RATest/vignettes/RDperm.pdf) that is available on [CRAN](https://cran.r-project.org/).

**Set up of the problem**

One of the assumptions in which a Sharp Regression Discontinuity Design rests is that baseline covariates are continuous at the cutoff. However, this hypothesis of continuity is intrinsically intestable. Practitioners have relied on graphical inspections, or on *conditional means* tests, neither of which is a test for the continuity of the baseline covariates at the cutoff.[1]

Canay and Kamat (2017) propose a permutation test approach to test this hypothesis, where the null hypothesis states the 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 also can control the limiting rejection probability under general assumptions.

We organize this post like the [vignette](https://cran.r-project.org/web/packages/RATest/vignettes/RDperm.pdf) proceding in a similar but for succinct fashion.

**Testable Hypothesis**

**Potential Outcomes**

Consider the simplest model for a randomized experiment with subject *i*’s (continuous) response to a binary treatment . The treatment assignment in the sharp RDD follows the rule , where is the so called running variable, and is the cutoff at which the discontinuity arises. This threshold is conveniently assumed to be equal to 0.

For every subject , there are two mutually exclusive potential outcomes – either subject gets the treateatment or not. If subject *i* receives the treatment (), we will say the potential outcome is . Similarly, if subject *i* belongs to the control group (), the potential outcome is . We are interested in the average treatment effect (ATE) at the cutoff, i.e.

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

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

Condition (1) is equivalent to being right continuous at and

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

**Induced Order Statistics**

Consider a random sample from a distribution of . The order statistics of the sample of the running variable, will in sample of the baseline covariate, say, according to the rule: if then for all . It is worth mentioning that the values of this induced order statistics are not necessarily ordered.

**The test statistic**

The test statistic exploits the behavior of the closest units to the left and right of the cutoff . More precisely, fix *q* ∈ ℕ and take the *q* closest values of the order statistics of to the right, and the closest values to the left:

and

respectively. The induced order for the baseline covariates is then

    and

respectively. The random variabes can be viewed as an independent sample of conditional on being close to the cutoff from the left. Analogously, can be thought of an independent sample of $ conditional on $ being close to the cutoff from the right. Let and be the empirical CDFs of the two samples of size , respectively,

    and

Stack all the observations of the baseline covariates into

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

**Computing *p*-values**

We argued that the permutation test rejects the null hypothesis (2) if is bigger than the quantile of the randimization distribution. Alternatively, we can define the *p*-value of a permutation test, , as

where is the observed sample, and is the cardinality of . It can be shown

therefore, the test that rejects when is level .

**Stochastic approximation**

When computing the permutation distribution, we often encounter the situation that the cardinality of might be large such that it becomes computationally prohibitive. In this situation, it is possible to approximate the -values the following way. Randomly sample permutations from with or without replacement. Suppose the sampling is with replacement, then are i.i.d. and uniformly distributed on . Then

is such that

where this $P$ takes into account the randomness of and the sampling of the . Like in the previous case, the test that rejects when is level .

It is worth noticing that the approximation satisfies the above equation regardless of , although a bigger will improve the approximation. As a matter of fact, as . The RATest package uses by default.

**Tuning parameter *q***

The implementation of the test statistic heavily relies on , 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 , and remains fixed as . Canay and Kamat (2017) recomend the rule of thumb

where is the smallest integer greater or equal to , is the density if at zero, is the coefficient of correlation and , and is the variance of . For the scalar case, it can be estimated from sample. The feasible tuning parameter is

where , and . The lower bound, represents situations in which the randomized and non-randomized versions of the permutation test differ, whereas the upper bound, 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 with . The density function is 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 are obtained directly from the R package quantreg. Finally, and are estimated by their sample counterparts.

The RATest package also computes the corresponding rule of thumb that changes a little bit in when *W* is a *K*-dimensional vector, since the variances and correlations are not scalars.

**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).

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 where , the running variable, is the vote shares between Democrats and Republicans.

Figure 1 shows Lee (2008) sharp RD strategy. The figure illustrates the sharp change in probability of a Democrat winning against the difference in vote share in the previous election. The data used here and contained in the package have six covariates and 6,558 observations with information on the Democrat runner and the opposition. The data set is named , 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)

#Load Package

require("RATest")

# Load the data set

lee2008<-lee2008

lee2008$d<- ifelse(lee2008$difdemshare >= 0,1,0)

# Predict with local polynomial logit of degree 4

logit.a <- glm(formula = demsharenext ~ poly(difdemshare, degree = 4) +

poly(difdemshare, degree = 4) \* d,

family = binomial(link = "logit"),

data = lee2008)

lee2008$demsharenexthat<-predict(logit.a, lee2008, type = "response")

# Create local average by 0.005 interval of the running variable (share)

breaks <- round(seq(-1, 1, by = 0.005), 3)

lee2008$i005<-as.numeric(as.character(cut(lee2008$difdemshare,

breaks = breaks,

labels = head(breaks, -1),

right = TRUE)))

m\_next<-tapply(lee2008$demsharenext,lee2008$i005,mean)

m\_next<-data.frame(i005=rownames(m\_next), m\_next=m\_next)

mp\_next<-tapply(lee2008$demsharenexthat,lee2008$i005,mean)

mp\_next<-data.frame(i005=rownames(mp\_next), mp\_next=mp\_next)

panel.a<-merge(m\_next,mp\_next,by=c("i005"), all=T)

panel.a$i005<-as.numeric(as.character(panel.a$i005))

# Plot panel (a) MHE

panel.a <- panel.a[which(panel.a$i005 > -0.251 & panel.a$i005 < 0.251), ]

plot.a <- ggplot(data = panel.a, aes(x = i005)) +

geom\_point(aes(y = m\_next)) +

geom\_line(aes(y = mp\_next, group = i005 >= 0)) +

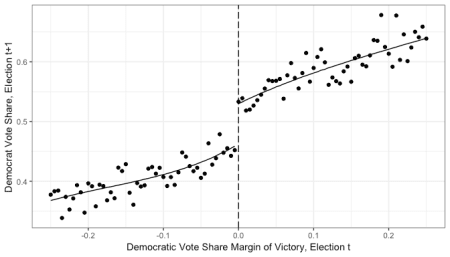
geom\_vline(xintercept = 0, linetype = 'longdash') +

xlab('Democratic Vote Share Margin of Victory, Election t') +

ylab('Democrat Vote Share, Election t+1') +

theme\_bw()

plot.a



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

# Predict with local polynomial logit of degree 4

logit.b <- glm(formula = demshareprev ~ poly(difdemshare, degree = 4) +

poly(difdemshare, degree = 4) \* d,

family = binomial(link = "logit"),

data = lee2008)

lee2008$demshareprevhat<-predict(logit.b, lee2008, type = "response")

# Create local average by 0.005 interval of the running variable (share)

breaks <- round(seq(-1, 1, by = 0.005), 3)

lee2008$i005<-as.numeric(as.character(cut(lee2008$difdemshare,

breaks = breaks,

labels = head(breaks, -1),

right = TRUE)))

m\_next<-tapply(lee2008$demshareprev,lee2008$i005,mean)

m\_next<-data.frame(i005=rownames(m\_next), m\_next=m\_next)

mp\_next<-tapply(lee2008$demshareprevhat,lee2008$i005,mean)

mp\_next<-data.frame(i005=rownames(mp\_next), mp\_next=mp\_next)

panel.b<-merge(m\_next,mp\_next,by=c("i005"), all=T)

panel.b$i005<-as.numeric(as.character(panel.b$i005))

# Plot panel (b) MHE

panel.b <- panel.b[which(panel.b$i005 > -0.251 & panel.b$i005 < 0.251), ]

plot.b <- ggplot(data = panel.b, aes(x = i005)) +

geom\_point(aes(y = m\_next)) +

geom\_line(aes(y = mp\_next, group = i005 >= 0)) +

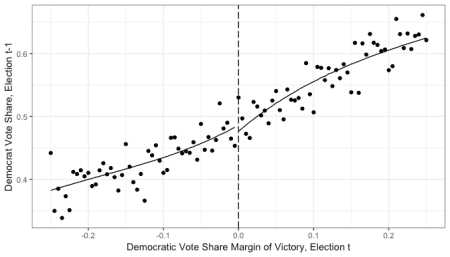
geom\_vline(xintercept = 0, linetype = 'longdash') +

xlab('Democratic Vote Share Margin of Victory, Election t') +

ylab('Democrat Vote Share, Election t-1') +

theme\_bw()

plot.b



This package however, implements Canay and Kamat (2017) in the function . The following code performs the test for the continuity of the named in the data set at the threshold. The function requires the name of the baseline covariate to be tested, the running variable , 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 () to the right and to the left of the cutoff. As default, the function uses the Cram’er-von Mises test . The function is available for a concise summary of the result.

# Lee2008

set.seed(101)

permtest<-RDperm(W="demshareprev", z="difdemshare",data=lee2008,q\_type=51)

summary(permtest)

#>

#> \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

#> \*\* RD Distribution Test using permutations \*\*

#> \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

#> 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 0.01 51 \*\*

#> ---

#> Signif. codes: 0.01 '\*\*\*' 0.05 '\*\*' 0.1 '\*'

The function reports the vaule of the test statistics (), 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 code shows how to do this

permtest\_rot<-RDperm(W=c("demshareprev","demwinprev", "demofficeexp"),

z="difdemshare",data=lee2008,q\_type='rot', n.perm=600)

summary(permtest\_rot)

#>

#> \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

#> \*\* RD Distribution Test using permutations \*\*

#> \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

#> Running Variable: difdemshare

#> Cutoff: 0

#> q: Rule of Thumb

#> 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 \*\*

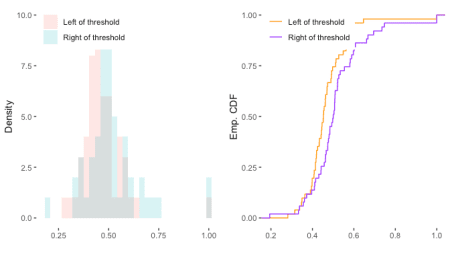
#> Joint.Test 0.11 0 32 \*\*\*

#> ---

#> 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 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.

plot(permtest,w="demshareprev")



**Conclusions**

In this blog post we described the RDperm funcion on 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.

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, Ivan A, and Vishal Kamat. 2017. “Approximate Permutation Tests and Induced Order Statistics in the Regression Discontinuity Design.” *The Review of Economic Studies* 85 (3). Oxford University Press: 1577–1608.

Hahn, Jinyong, Petra Todd, and Wilbert Van der Klaauw. 2001. “Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design.” *Econometrica* 69 (1). Wiley Online Library: 201–9.

Koenker, Roger, and Zhijie Xiao. 2002. “Inference on the Quantile Regression Process.” *Econometrica* 70 (4). Wiley Online Library: 1583–1612.

Lee, David S. 2008. “Randomized Experiments from Non-Random Selection in Us House Elections.” *Journal of Econometrics* 142 (2). Elsevier: 675–97.

Portnoy, Stephen, and Roger Koenker. 1989. “Adaptive L-Estimation for Linear Models.” *The Annals of Statistics*. JSTOR, 362–81.

Silverman, Bernard W. 1986. *Density Estimation for Statistics and Data Analysis*. Vol. 26. CRC press.

[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 ’s observed outcome, is , whereas the identification assumption in Hahn et al (2001) requires that both

[3] See condition 2b in Lee (2008).