



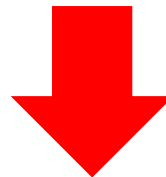
## **The R User Conference 2008**

August 12-14, Technische Universität Dortmund, Germany

**The “DeltaR” package:  
A flexible way to compare  
regression models on independent samples  
using a bootstrap approach**

# Premise

- As a statistician working with social scientists, I often deal with stimulating problems.
- In the past year, a number of psychology researchers asked me whether it was possible to test the difference in the amount of variance explained by two regression models performed on independent samples (e.g., males vs. females).
- Given the lack of specific software applications and the relatively small amount of literature available on this topic ...



... I just tried to “pull myself up by my own bootstraps” ...

# General aim

To propose a flexible method to test the difference between explained variance ( $\Delta$  *R-squares*) of two multiple linear regression models on **two independent samples**.

- From a theoretical perspective, the method is based upon a stratified, non-parametric **bootstrap approach**.
- From an applied viewpoint, the method is implemented in an R package called "**DeltaR**".

# Specific aims

Given two regression models (not necessarily equal in their formulation) performed on two independent samples, DeltaR will produce:

- **a confidence interval (CI)** at a given *1-alpha* probability level for the true value of  $\Delta R$ -squares;
- **an observed *p-value*** associated with the value of  $\Delta R$ -squares hypothesized under  $H_0$  (i.e., if this value is set to 0,  $H_0$  states that the two models explain the same amount of variance);
- **a graphical representation** of both the analyzed data and the results.

# Introduction

Bootstrap is a computer-intensive method of statistical analysis that allows to draw inferences on the real population and its parameters.

The key idea is to build a sampling distribution for a statistic of interest by resampling with replacement from the original sample.

Basic notations used:

- $\theta$  = the unknown parameter of interest
- $t$  = statistic used to estimate the unknown parameter
- $t^*$  = bootstrapped  $t$
- $R$  = number of bootstrap samples

# Main issues to deal with

1. Choosing a  $t$  statistic
2. Developing a bootstrap strategy
3. Building a bootstrap hypothesis test

# 1. Choosing a t statistic

- A first type of  $t^*$  statistic could simply be:

$$t^* = R_1^{2*} - R_2^{2*}$$

i.e., the difference between bootstrapped R-squares of the two regression models.

- Another type of  $t^*$  statistic could be:

$$t^* = \frac{(R_1^{2*} - R_2^{2*})}{\sqrt{\widehat{Var}(R_1^{2*} - R_2^{2*})}}$$

- The subsequent applications are based on the first type of  $t^*$  statistic.

## 2. Developing a bootstrap strategy

Properties of the bootstrap approach developed for our problem:

- **non-parametric:** the empirical distribution function obtained by bootstrapping is used to estimate the real population distribution.

All original data values have the same probability to be resampled.

- **stratified:** resampling is done within each independent sample.
- **random-X resampling:** assume that we want to fit a regression model with  $y$  as dependent variable and  $x_1, x_2, \dots, x_k$  as predictors.

Let us call  $\mathbf{z}_i' = (y_i, x_{1i}, x_{2i}, \dots, x_{ki})$ ,  $i = 1 \dots n$ , each multivariate observation.

To create bootstrap samples, we simply resample  $\mathbf{z}_i'$ .

- **balanced:** within each independent sample, each observation is overall resampled the same number of times.



### 3. Building a bootstrap two-tailed hypothesis test

In general, we use the following rule to calculate an equal-tail  $p$ -value for the hypothesized  $\Delta R$ -squares value:

$$\hat{p}_{eq}(\mathcal{G}_0) = 2 \min \left( \frac{\sum_{j=1}^R I(t^* \leq \mathcal{G}_0)}{R}, \frac{\sum_{j=1}^R I(t^* > \mathcal{G}_0)}{R} \right)$$

where :

$t^*$  = the bootstrapped statistic

$R$  = number of bootstrap replicates

$I(\cdot)$  = the indicator function

$\mathcal{G}_0$  = value of the parameter hypothesized under  $H_0$

# The deltaR function: Arguments

- **data1, data2:** two data frames containing the variables of the first and of the second sample, respectively;
- **formula1, formula2:** the typical R symbolic description of the models to be fitted on the first and on the second sample;
- **iter:** the number of bootstrap replicates;
- **Ctype:** a character string representing the type of confidence interval required;
- **conf:** a scalar specifying the confidence level of the required interval;
- **theta0:** a number indicating the hypothesized true value of the  $\Delta R$ -squares;
- **graph:** logical. If TRUE a graphical representation of results is created.

# The deltaR function: Returned objects

- **Main statistics:** the observed value of the  $\Delta R$ -squares statistic, the bootstrap estimates of bias, and the bootstrap estimate of standard error, ...
- **CI:** the type and values of the required confidence interval.
- **pval:** the observed *p-value* for the hypothesized value of  $\Delta R$ -squares.
- **a graphical representation** of both the analyzed data and the results.

# A case study

A sample of 694 Italian adolescents (50.4% boys) aged 11-13 years who participate in structured after-school activities completed self-reported measures of:

- Sense of loneliness ( $y$ )
- Parent-adolescent relationship ( $x_1$ )
- Perceived support from after-school activity leaders ( $x_2$ )

Hypothesized theoretical model:

$$y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2$$

Data source: Vieno A., Altoè G., Santinello M., & Vidotto G. (in press). *La partecipazione ad attività extrascolastiche come moderatore della solitudine durante la preadolescenza [Participation in Structured After-School Activities as a Moderator of Loneliness During Early Adolescence]*. Età Evolutiva.

# A case study: The first analysis

- The first analysis was conducted fitting the baseline model separately for boys and girls:

Sample	R-square	p value for interaction term	Bootstrapped 95% CI for $\Delta R\text{-squares}$ *	
			lower limit	upper limit
Boys (n=350)	.142	.005	-.169	.072
Girls (n=344)	.192	.650		

\* The Bootstrap percentile method was used.

- The resulting 95% CI suggests that the difference between R-squares is not statistically significant.
- Question: What happens if we delete the interaction term in the girls' model?

# A case study: deltaR code

```
> summary(data.boys)
      yb          x1b          x2b
Min.   :1.080   Min.   :1.50   Min.   :1.000
1st Qu.:1.850   1st Qu.:3.78   1st Qu.:2.500
Median :2.150   Median :4.17   Median :3.000
Mean   :2.196   Mean   :4.09   Mean   :3.097
3rd Qu.:2.460   3rd Qu.:4.56   3rd Qu.:4.000
Max.   :4.380   Max.   :5.00   Max.   :4.500
> summary(data.girls)
      yg          x1g          x2g
Min.   :1.150   Min.   :1.330   Min.   :1.000
1st Qu.:1.770   1st Qu.:3.670   1st Qu.:2.500
Median :2.080   Median :4.110   Median :3.000
Mean   :2.161   Mean   :4.006   Mean   :3.035
3rd Qu.:2.460   3rd Qu.:4.500   3rd Qu.:3.500
Max.   :4.230   Max.   :5.000   Max.   :4.500
> deltaR(data.boys,data.girls,yb~x1b+x2b+x1b:x2b,yg~x1g+x2g,iter=1000,graph=TRUE)
```

MAIN BOOTSTRAP STATISTICS:

```
      original      bias      st. error
t* -0.0495682 0.0026544 0.0609378
```

```
t = delta R-squares
bootstrap samples = 1000
```

95% PERCENTILE METHOD CI:

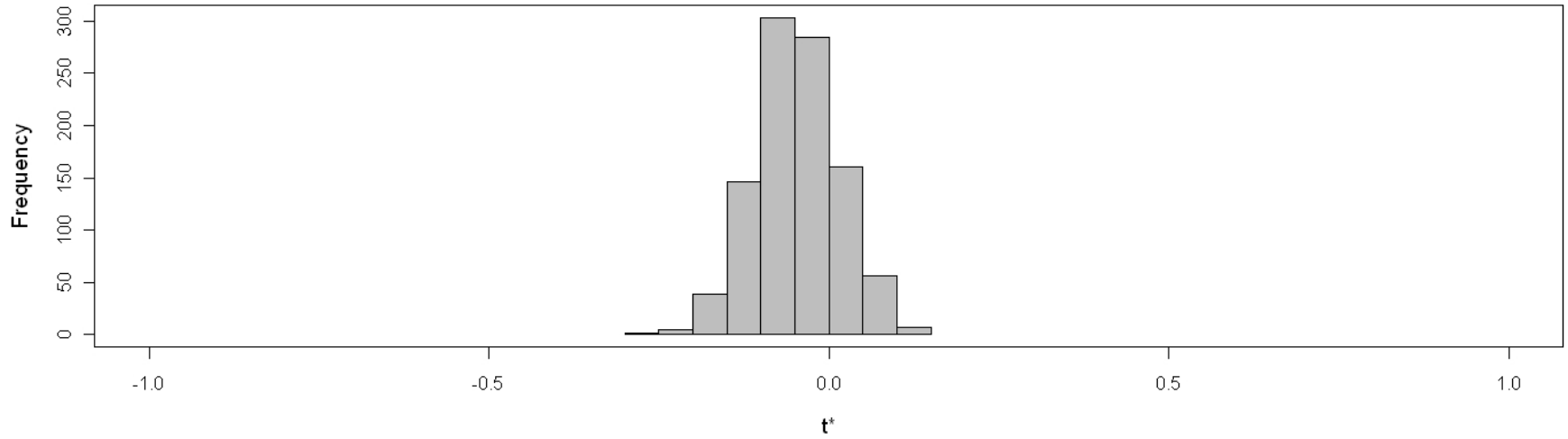
```
lower limit upper limit
-0.1618241  0.0744229
```

ALTERNATIVE HYPOTHESIS: true difference in R-squares is not equal to 0

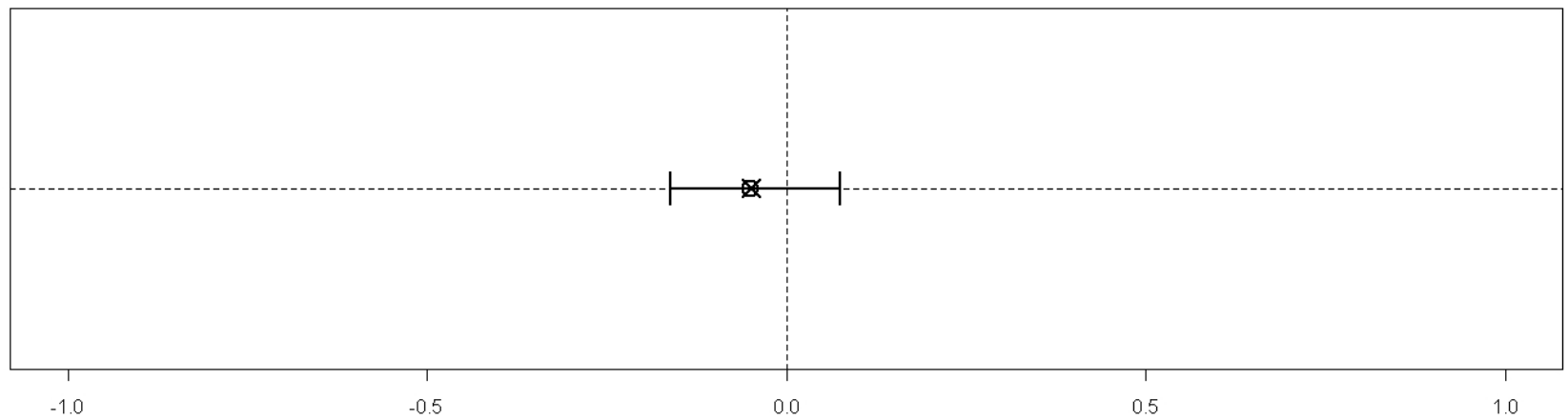
```
p-value = 0.446
```

# A case study: deltaR graphics

Histogram of bootstrapped t



Original t and bootstrapped 95% CI



# Discussion

- The proposed method may be considered as a first step to deal with the issue of comparing the amount of explained variance of regression models performed on independent samples.
- The choice of the models to be compared is a key aspect that needs to be strongly based on theoretical as well as statistical considerations.
- The technique is particularly useful for explorative purposes; furthermore, the obtained results may serve as a starting point for more complex analyses (e.g., multisample analyses).



# Limitations and future research

- Simulations are needed to evaluate the performance of the proposed method in terms of Type I and Type II error (they are currently running ...).
- The effect of observed outliers on method performance should be investigated.
- Other types of bootstrap CI and bootstrap strategies should be considered.

# References

Canty, A. & B. Ripley (2008). boot: Bootstrap R (S-Plus) Functions. *R package version 1.2-33*.

Davison, A. C. & D. V. Hinkley (1997). *Bootstrap Methods and their Application*. Cambridge: Cambridge University Press.

Efron, B. (1979). Bootstrap Methods: Another Look at the Jackknife. *Annals of Statistics* 7: 1-26.

Efron, B. & R. J. Tibshirani (1993). *An Introduction to the Bootstrap*. New York: Chapman and Hall.

Fox, J. (2008). Bootstrapping Regression Models. In “*An R and S-PLUS Companion to Applied Regression: A Web Appendix to the Book.*”, Sage, Thousand Oaks, CA.

URL <http://socserv.socsci.mcmaster.ca/jfox/Books/Companion/appendix-bootstrapping.pdf>

MacKinnon, J. (2007). *Bootstrap Hypothesis Testing*. Queen’s Economics Department Working Paper No. 1127.

URL [http://www.econ.queensu.ca/working\\_papers/papers/qed\\_wp\\_1127.pdf](http://www.econ.queensu.ca/working_papers/papers/qed_wp_1127.pdf)

**Thank you for your attention !!**



... so, let's Resample 😊 ...