

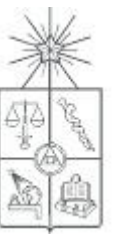


INGENIERIA INDUSTRIAL
UNIVERSIDAD DE CHILE

IN4402: Aplicaciones de Probabilidades y Estadística

CAUSAL RANDOM FOREST - APPLICATION

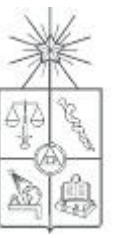
ANDRÉS FERNÁNDEZ



- In summary:
 - When there is need for a treatment estimation, we can use Trees to estimate a Conditional Average Treatment Effect (CATE)
 - Because leaves provide a good similarity in conditional covariables
 - We modify the splitting criterion to maximize **heterogeneity**
 - Decision Trees produce **biased** and **not consistent** estimators
 - We use $EMSE_{\tau}$ as the criterion to maximize heterogeneity and balance
 - We estimate many **causal** trees to produce a **causal** random forest
 - We use splitting and estimate samples to produce an **honest** result
 - We estimate an ATE that is a function of covariables

CAUSAL RANDOM FORESTS

INTRODUCTION AND MAIN CONCEPTS



- Why study heterogeneity? -> Target of treatment more efficiently
- Application to Summer Jobs randomly assigned
 - For crime prevention, Chicago offers a part-time job during the summer to 14-22 aged vulnerable young people
 - Data of criminal records, employment and school attendance is followed
 - Outcome: (1) violent-crime arrests, (2) employment
 - Covariates: 14 (demographics, neighborhood, educational, criminal history)
 - $N = 6850$



- Does CRF identify correctly heterogeneous effects?
- Data randomly split in two (in sample and out of sample)
 - They use in-sample data to train HCRF and out-of-sample for testing
 - They use bagging procedure sampling 20% of the observations
 - They estimate $\hat{\tau}_i$ (CATE) for every individual
 - They split each sample in two groups: $(\hat{\tau}_i > 0)$ & $(\hat{\tau}_i < 0)$
 - For every group, they estimate the regression

$$Y_i = \beta_0 + \beta_1 1_{(\hat{\tau}_i > 0)} + \beta_2 T \cdot 1_{(\hat{\tau}_i > 0)} + \beta_3 T \cdot (1 - 1_{(\hat{\tau}_i > 0)}) + \Phi X^k + a_b + \varepsilon_i$$

- Is $\beta_2 = \beta_3$?

CAUSAL RANDOM FORESTS

INTRODUCTION AND MAIN CONCEPTS



- Does CRF identify correctly heterogeneous effects?

CAUSAL RANDOM FORESTS

INTRODUCTION AND MAIN CONCEPTS



- Does CRF identify correctly heterogeneous effects?

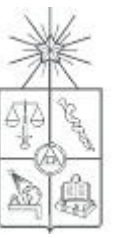
TABLE 1—TREATMENT EFFECTS BY PREDICTED RESPONSE

Subgroup	No. of violent crime arrests	Any formal employment
<i>Panel A. In sample</i>		
$\hat{\tau}_i^{CF}(x) > 0$	0.22 (0.05)	0.19 (0.03)
$\hat{\tau}_i^{CF}(x) < 0$	−0.05 (0.02)	−0.14 (0.03)
H_0 : subgroups equal, $p =$	0.00	0.00
<i>Panel B. Out of sample</i>		
$\hat{\tau}_i^{CF}(x) > 0$	−0.01 (0.05)	0.08 (0.03)
$\hat{\tau}_i^{CF}(x) < 0$	−0.02 (0.02)	−0.01 (0.03)
H_0 : subgroups equal, $p =$	0.77	0.02

- Is it maybe overfitting?

CAUSAL RANDOM FORESTS

INTRODUCTION AND MAIN CONCEPTS



- Does CRF identify correctly heterogeneous effects?
- Even more honest estimation:

CAUSAL RANDOM FORESTS

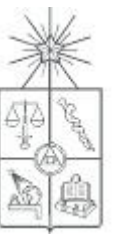
INTRODUCTION AND MAIN CONCEPTS



TABLE 1—TREATMENT EFFECTS BY PREDICTED RESPONSE

Subgroup	No. of violent crime arrests	Any formal employment
<i>Panel B. Out of sample</i>		
$\hat{\tau}_i^{CF}(x) > 0$	−0.01 (0.05)	0.08 (0.03)
$\hat{\tau}_i^{CF}(x) < 0$	−0.02 (0.02)	−0.01 (0.03)
H_0 : subgroups equal, $p =$	0.77	0.02
<i>Panel C. Adjusted in sample</i>		
$\hat{\tau}_i^{CF}(x) > 0$	−0.06 (0.04)	0.05 (0.03)
$\hat{\tau}_i^{CF}(x) < 0$	−0.02 (0.02)	−0.04 (0.03)
H_0 : subgroups equal, $p =$	0.41	0.02

- Does CRF identify correctly heterogeneous effects?
- Even more honest estimation of in sample estimation:
- There were some overfitting in violent-crime outcome



- Summary:
 - HCRF are more successful in identifying different subgroups than linear regression
 - There are heterogeneity in effect in future employment but not in violent-crime reduction
 - HCRF are still subject to overfitting
- By characterizing the young people more affected by summer jobs we can be more efficient in
 - Sending more targeted invitations
 - Change the purpose of the program