New trends in predictive modelling - the uplift models success story

Montserrat Guillén[1] and Leo Guelman[1,2]

[1]University of Barcelona and [2]Royal Bank of Canada

R in Insurance Conference
London, UK, July 14, 2014

1 Introduction
- Motivation
- The problem

2 Customers who react

3 uplift R package

4 Price elasticity

Pricing, retaining, enhancing

✓ Pricing (calculating expected loss + margin + profits)

Pricing, retaining, enhancing

- ✓ Pricing (calculating expected loss + margin + profits)
- ✓ Retaining (reducing policy lapse)

Pricing, retaining, enhancing

- ✓ Pricing (calculating expected loss + margin + profits)
- ✓ Retaining (reducing policy lapse)
- ✓ Enhancing (cross-selling additional products to existing customers)

Pricing, retaining, enhancing

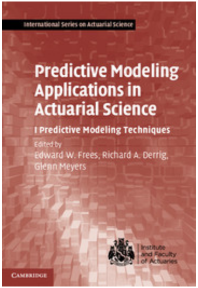
- ✓ Pricing (calculating expected loss + margin + profits)
- ✓ Retaining (reducing policy lapse)
- ✓ Enhancing (cross-selling additional products to existing customers)

  Classical approach: Predictive modeling (Negative Binomial model, Gamma model, Logistic regression, Cox,...)

Pricing, retaining, enhancing

- ✓ Pricing (calculating expected loss + margin + profits)
- ✓ Retaining (reducing policy lapse)
- ✓ Enhancing (cross-selling additional products to existing customers)

Classical approach: Predictive modeling (Negative Binomial model, Gamma model, Logistic regression, Cox,...)



Cambridge University Press, July 2014

The classical approach

✓ Price depends on expected loss and occasionally on claims experience (experience rating or bonus malus)

The classical approach

- ✓ Price depends on expected loss and occasionally on claims experience (experience rating or bonus malus)
- ✓ Price changes adjust risk transfer variations exclusively

The classical approach

- ✓ Price depends on expected loss and occasionally on claims experience (experience rating or bonus malus)
- ✓ Price changes adjust risk transfer variations exclusively
- ✓ Modeling losses by line of business is central

The classical approach

- ✓ Price depends on expected loss and occasionally on claims experience (experience rating or bonus malus)
- ✓ Price changes adjust risk transfer variations exclusively
- ✓ Modeling losses by line of business is central
- ✓ Customer retention is examined later

The classical approach

- ✓ Price depends on expected loss and occasionally on claims experience (experience rating or bonus malus)
- ✓ Price changes adjust risk transfer variations exclusively
- ✓ Modeling losses by line of business is central
- ✓ Customer retention is examined later

The classical approach

- ✓ Price depends on expected loss and occasionally on claims experience (experience rating or bonus malus)
- ✓ Price changes adjust risk transfer variations exclusively
- ✓ Modeling losses by line of business is central
- ✓ Customer retention is examined later

Observational data are available

But....do insurers have historical information that can be understood as experimental data?

Data

- ✓ Is treatment data in the insurance portfolio available?
- ✓ Have partial marketing actions been performed in the past?
- ✓ Is it possible to collect "action-response" data?

Data

- ✓ Is treatment data in the insurance portfolio available?
- ✓ Have partial marketing actions been performed in the past?
- ✓ Is it possible to collect “action-response” data?

An example:

Direct mail campaign in a bank (L=6256)
Proportion of purchase and non purchase in each treatment group

	Control	Promotion
No purchase	85.17%	61.60%
Purchase	14.83%	38.40%

Average treatment effect (uplift)=23.57%

Questions for experimental data

- ✓ Many factors influence customer decisions, so it is difficult to predict the probability of a customer lapse and the impact of loosing a customer. We should take into account the relationship between events affecting one particular contract and customer's decisions regarding other contracts held in the same company

- ✓ Policy holders expect services from the insurer. The aim is to find a personalized treatment for each customer.

- ✓ Which specific actions should a company design?
- ✓ What is the optimal price to be charged?
- ✓ Which groups of customers should be targeted in order to increase profits (reduce lapses and control price rebates)?

Model and notation

We assume price $P^*_{\ell m}$ charged to policy holder $\ell = \{1, \ldots, L\}$ for a given contract in year $m = \{1, \ldots, M\}$ is the sum of three components:

$$P^*_{\ell m} = LC_{\ell m} + SR_{\ell m} + B_{\ell m}, \quad \ell = \{1, \ldots, L\} \; m = \{1, \ldots, M\}$$

- a fair premium ($LC_{\ell m}$), resulting from an evaluation of the policy holder's risk characteristics, that is, an estimation of expected claims compensation or loss;
- a price loading ($SR_{\ell m}$), capturing solvency requirements, managerial efficiency or caution; and, finally,
- profits ($B_{\ell m}$), reflecting a minimum level of return to the company's shareholders or to the insurance company's owner.

Model and notation

- We define renewal $D_{\ell m}$ as a binary variable which equals 1 if policy holder ℓ renews his policy in year m, and 0 otherwise.
- Renewal $D_{\ell m}$ depends on marketing actions.
- Renewal $D_{\ell m}$ depends on external competitors.
- Renewal ($D_{\ell m}$) and price ($P^*_{\ell m}$) are mutually dependent.
- If the price increases many policy holders will abandon the company, but if the price falls then renewal is more likely than lapsing.

Our goal

- We estimate the expected change in customer value due to personalized actions (marketing campaign, price change,....).
- Or we estimate the global expected profit change due to personalized action.

We use personalized treatment models, where price change is a "treatment" (action) that predicts a "response" and combines information on :

- **risk**
- **behaviour**

General framework

- There are L policy holders in a portfolio and that they may hold more than one policy.
- We indicate each type of insurance product by j, where $j = 1, ..., K$ and K is the total number of possible insurance products.
- The company can control prices, so let us call $A_{\ell jm}$ the action (price change) to be offered to policy holder ℓ in year m for policy j before renewal.

General framework

- We define the set of all individual strategies as $A_m = \{A_{\ell jm};\ \ell = 1, ..., L;\ \ j = 1, ..., K\}$.
- The total value at m, $V(A_m)$, is the sum of the expected profits over all customers generated from year m to M.

Value: multi-product and multi-year

- The indicator $I_{\{D_{\ell jm}=1\}}$ equals one if policy holder ℓ holds product j in year m, and 0 otherwise.
- Additionally, let $S_{\ell js}$ be the probability that customer ℓ keeps policy j in year s, namely $P(D_{\ell js} = 1)$ for $s = m, ..., M$.
- Let $B_{\ell jm}$ be the profit of policy j from policy holder ℓ in year m, and r is the interest discount factor. So the total value of a portfolio at m is:

$$V(A_m) = \sum_{\ell=1}^{L} \sum_{j=1}^{K} I_{\{D_{\ell jm}=1\}} B_{\ell jm} \sum_{s=m}^{M} S_{\ell js} r^{s-m}.$$

Profit: only one product and one year case

$$\underset{Z_{\ell t} \forall \ell \forall t}{\text{Max}} \sum_{\ell=1}^{L} \sum_{t=1}^{T} Z_{\ell t} \left[P_{\ell}(1 + RC_t)(1 - \hat{LR}_{\ell t})(1 - \hat{r}_{\ell t}) \right]$$

with restrictions:

$$\sum_{t=1}^{T} Z_{\ell t} = 1, \quad Z_{\ell t} \in \{0, 1\}, \quad \sum_{\ell=1}^{L} \sum_{t=1}^{T} Z_{\ell t} \hat{r}_{\ell t} / L \leq \alpha$$

where P_{ℓ} is price paid by ℓ, $\ell = \{1, 2, \ldots, L\}$, L is the total number of customers, RC_t is price change rate which is categorized in T ordered values, $t = \{1 < 2 < \ldots < T\}$, $\hat{LR}_{\ell t}$ is the loss ratio, namely, cost divided by premium, $\hat{r}_{\ell t}$ is the probability of lapse for customer ℓ if price change t is applied ($Z_{\ell t} = 1$) and α is the maximum lapse rate that is allowed for this portfolio (so, $1 - \alpha$ is the minimum retention rate).

Motivation

- The values chosen for the actionable attributes have important implications for the ultimate profitability of the insurance company
- There is no "global" better action $\Rightarrow$ Relevant in the context of treatment heterogeneity effects
- The objective is NOT to predict a response variable with high accuracy (as in predictive modeling), but to select the optimal action or treatment for each client
- Optimal personalized treatment $\Rightarrow$ the one that maximizes the probability of a desirable outcome (e.g., Profits)
- Not addressed by traditional predictive modeling techniques (GLMs, CART, SVM, Neural Nets, etc.).

Customer loyalty and duration

Households are customer units

- ✓ Brockett, P. L. et al. (2008) Survival Analysis of Household Insurance Policies: How Much Time Do You Have to Stop Total Customer Defection, **Journal of Risk and Insurance** 75, 3, 713-737.
- ✓ Guillen, M., Nielsen, J. P., Scheike, T. and Perez-Marin, A. M. (2011a) Time-varying effects in the analysis of customer loyalty: a case study in insurance, **Expert Systems with Applications**, 39, 3551-3558.

Cross-selling

Selling more policies to existing policyholders

- ✓ Guillen, M., Perez-Marin, A.M. and Alcañiz, M. (2011) A logistic regression approach to estimating customer profit loss due to lapses in insurance, **Insurance Markets and Companies: Analyses and Actuarial Computations**, 2, 2, 42-54.
- ✓ Thuring, F., Nielsen, J.P., Guillen, M. and Bolance, C. (2012) Selecting prospects for cross-selling financial products using multivariate credibility, **Expert Systems with Applications**, 39, 10, 8809-8816.

1 Introduction
- Motivation
- The problem

2 Customers who react

3 uplift R package

4 Price elasticity

Treatment-response: a new perspective

- ✓ Guelman, L., Guillen, M. and Perez-Marin, A. M. (2012) Random forest for uplift modeling: an insurance customer retention case, **Lecture Notes in Business Information Processing**, 115, 123-133.
- ✓ Guelman, L., Guillen, M. and Perez-Marin, A. M. (2013) Uplift random forests, **Cybernetics and Systems: an International Journal**, accepted.
- ✓ Guelman, L., Guillen, M. and Perez-Marin, A. M. (2014) A survey of personalized treatment models for pricing strategies in insurance, **Insurance: Mathematics and Economics**, accepted.

Targeting the right customers

- An insurance company is interested in increasing the retention rate of its customers.
- The point is to decide which customers should be targeted by some retention action.
- Instead of targeting the most likely to leave customers, the authors advocate that the company should target those customers with a higher expected increase in the retention probability as a result of the marketing action by using uplift modeling.

If targeted by retention action	If NOT targeted by retention action	Remark
Churn	Churn	Unnecesary costs
Renew	Renew	Unnecesary costs
Churn	Renew	Negative effects
Renew	Churn	Best targets!

Methodology:

Notation:

- $X = \{X_1, ..., X_p\}$ a vector of predictor variables,
- Y = binary response variable (1=renew, 0=lapse)
- t refers to the treatment $(t = 1)$ and control $(t = 0)$
- L = a collection of observations $\{(y_\ell, x_\ell, t_\ell)\,; \ell = 1, ..., L\}$
- Uplift model $\widehat{f}^{uplift}(x_\ell) = E\left(Y_\ell | x_\ell; t_\ell = 1\right) - E\left(Y_\ell | x_\ell; t_\ell = 0\right)$

Uplift model: indirect estimation

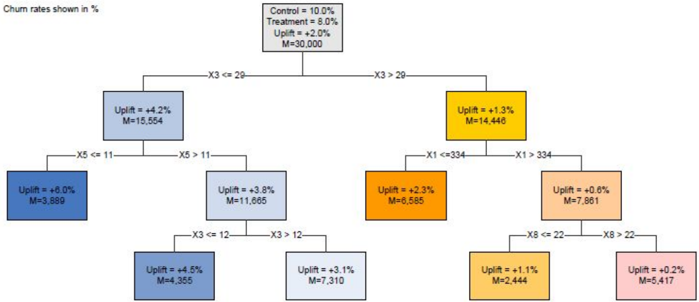
There are two general approaches: indirect and direct estimation

- Indirect uplift estimation:
  - Build two separate models, one using the treatment $(t = 1)$ subset and another one using control data $(t = 0)$.
    Predicted uplift is estimated by subtracting the class probabilities from the two models
    $P(Y = 1|x; t = 1) - P(Y = 1|x; t = 0)$
  - Alternatively, a single model can be obtained including an interaction term for every predictor in $X = \{X_1, ..., X_p\}$ and treatment t.
    This method does not work very well in practice, as the relevant predictors for the response are likely to be different from the relevant uplift predictors and the functional form of the predictors are likely to be different as well.

Uplift model: direct estimation

- Modeling uplift directly:
  - Requires modifying existing methods/algorithms or designing novel ones
  - Intuitively, tree-based algorithms are appropriate as they partition the input space into subgroups
  - Rzepakowski and Jaroszewicz (2011) and Radcliffe and Surry (2011) have proposed estimation algorithms
  - Our proposed algorithm: uplift Random Forests

Methodology: illustration



Methodology: uplift Random Forests

- In Guelman et al. (2012 and 2013) the proposed algorithm for modeling uplift directly is based on maximizing the distance in the class distributions between treatment and control groups
- Relative Entropy or *Kullback-Leibler distance KL* between two probability mass functions $P_t(Y)$ and $P_c(Y)$ is given by

$$KL\left(P_t(Y)||P_c(Y)\right) = \sum_{y \in Y} P_t(y) \log \frac{P_t(y)}{P_c(y)}$$

.

Methodology

- Conditional on a given split Ω, KL becomes

  $$KL\left(P_t(Y)||P_c(Y)|\Omega\right) = \sum_{\omega \in \Omega} \frac{M(\omega)}{M} KL\left(P_t(Y|\omega)||P_c(Y|\omega)\right)$$

  where $M = M_t + M_c$ (the sum of the number of training cases in treatment and control groups) and $M(\omega) = M_t(\omega) + M_c(\omega)$ (the sum of the number of training cases in which the outcome of the uplift Ω is ω in treatment and control groups).
- Define KL_{gain} as the increase in the KL divergence from a split Ω relative to the KL divergence in the parent node

  $$KL_{gain}(\Omega) = KL(P_t(Y)||P_c(Y)|\Omega) - KL(P_t(Y)||P_c(Y))$$

Methodology

- Final split criterion is

$$KL_{ratio}(\Omega) = \frac{KL_{gain}(\Omega)}{KL_{norm}(\Omega)}$$

where KL_{norm} is a normalization factor that punishes:

  - splits with different treatment/control proportions on each branch
  - splits with unbalanced number of cases on each branch

Empirical study: targeting customers that react to campaigns

- Auto insurance portfolio from a large Canadian insurer
- A sample of approx. 12,000 customers coming up for renewal were randomly allocated into two groups:
    - Renewal letter+courtesy call: aim was to maximize customer retention
    - A control group: no retention efforts
    - Treatment is not much effective if targets are selected randomly

	Attrition rates by group		
	Overall	Letter + Call	Control
Retained policies	10857	7492	3365
Cancelled policies	1111	757	354
Attrition rate	9.3%	9.2%	9.5%

Empirical study

We compare four uplift models:

- Uplift Random Forest Algorithm (upliftRF)
- The Two-Model Approach by using logistic regression (two-model)
- A Single Uplift Tree with Pruning (single-tree)
- and the approach based on explicitly adding an interaction term between each predictor and the treatment indicator by using logistic regression (int-model)

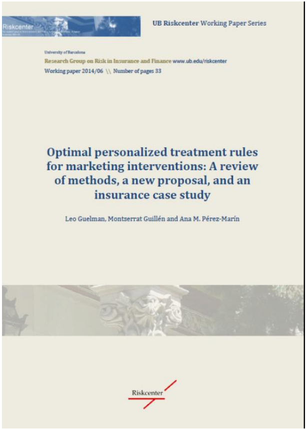
Top decile uplift

	Attrition rate (%)		
	Control	Treatment	Uplift
upliftRF	21.24	9.21	12.03
two-model	33.60	23.03	10.57
single-tree	13.98	5.21	8.77
int-model	27.41	20.60	6.81
random	9.50	9.20	0.30

Empirical study: conclusions

- None of the models dominates the others at all target volumes
- The *upliftRF* performs best in this application, specially for low target volumes: it is able to identify a 30 percent of customers for whom the retention program was highly effective (any additional targeted customer would result in a smaller reduction in attrition, as a result of negative effects of the campaign on the remaining customers)
- The *int-model* and *two-model* are able to identify the top 10 percent customers with highest attrition rate, but not those most impacted by the retention activity

Working paper


UB Riskcenter Working Paper Series

University of Barcelona

Research Group on Risk in Insurance and Finance www.ub.edu/riskcenter

Working paper 2014/06 \\ Number of pages 33

Optimal personalized treatment rules for marketing interventions: A review of methods, a new proposal, and an insurance case study

Leo Guelman, Montserrat Guillén and Ana M. Pérez-Marín

Riskcenter

http://www.ub.edu/riskcenter/research/WP/UBriskcenterWP201406.pdf

1 Introduction
- Motivation
- The problem

2 Customers who react

3 uplift R package

4 Price elasticity

uplift Package Highlights

- **First R package implementing uplift models**
- **Exploratory Data Analysis (EDA) tools customized for uplift**
  - Check balance of covariates (`checkBalance`)
  - Univariate uplift analysis (`explore`)
  - Preliminary variable screening (`niv`)
- **Uplift estimation methods**
  - Causal conditional inference forests (`ccif`)
  - Uplift random forests (`upliftRF`)
  - Modified covariate method (`tian_transf`)
  - Modified outcome method (`rvtu`)
  - Uplift k-nearest neighbor (`upliftKNN`)
- **Performance assessment for uplift models**
  - Uplift by decile (`performance`)
  - Qini curve and Qini-coefficient (`qini`)
- **Other functionality**
  - Profiling uplift models (`modelProfile`)
  - Monte-Carlo uplift simulations (`sim_pte`)

Package Documentation and Key Papers

- Guelman, L. (2014). uplift: Uplift Modeling. R package version 0.3.5. Available from the CRAN: http://www.cran.r-project.org/package=uplift
- Guelman, L., Guillén, M. and Pérez-Marín, A.M. (2014). “A survey of personalized treatment models for pricing strategies in insurance". *Insurance: Mathematics and Economics*. *Accepted*.
- Guelman, L., Guillén, M. and Pérez-Marín, A.M. (2014). “Uplift random forests". *Cybernetics & Systems*, Special issue on “Intelligent Systems in Business and Economics". *Accepted*.
- Guelman, L., Guillén, M. and Pérez-Marín, A.M. (2014). “Optimal personalized treatment rules for marketing interventions: A review of methods, a new proposal, and an insurance case study". *Submitted*.

Illustrative Data: Cross-Sell Intervention from a Major Bank

- **Pilot direct mail campaign** to sell a financial product to existing bank clients
- **Randomized experiment** with **N=6256 clients** assigned in **equal proportions** to treatment and control groups
- Treated clients received a **promotion** to buy the product. Clients in the control group did NOT receive the promotion
- Overall uplift of **23.6%** (38.4% - 14.8%), significantly higher than usual, but cost of promotion was very high as well $\Rightarrow$ **cross-sell initiative still not cost effective if all clients are targeted**

Table: Cross-sell rates by group

	Treatment	Control
Purchased product = N	1927	2664
Purchased product = Y	1201	464
Cross-sell rate	**38.4%**	**14.8%**

The Business Question

- Can we identify a **subgroup of clients for which the cross-sell intervention was more effective that the average**?
- If so, **target only those clients** in the post-pilot campaign deployment
- **Different from traditional predictive modeling methods** which attempt to predict

$$\mathit{Prob}(\text{Product "B"} | \text{Product "A"}, \mathbf{X})$$

- Here we attempt to estimate the **causal effect** of the intervention **at the individual client level**

Data Splitting

- bankDM dataset contains the cross-sell outcome (response), the treatment indicator (treatment), and **13 predictors** describing various demographic and behavioral client characteristics

- Partition data into train set (bankDM.train) and test set (bankDM.test) in 70/30 proportions.

```
set.seed(455)
samp.ind <- sample(1:nrow(bankDM), 0.7 * nrow(bankDM), replace = FALSE)
bankDM.train <- bankDM[samp.ind, ]
bankDM.test <- bankDM[-samp.ind, ]
```

Check Balance of Predictors Between Treatment/Control

Given predictors, a treatment variable, and (optionally) a stratifying factor, `checkBalance` calculates standardized mean differences along each predictor, and tests for conditional independence of the treatment variable and the covariates.

```
balForm <- as.formula(paste("treatment ~", paste("X", 1:13, sep = "",
    collapse = "+")))
cb <- checkBalance(balForm, data = bankDM.train)
round(cb$results[, c(1:3, 6:7), ], 2)

##       stat
## vars   treatment=0 treatment=1 adj.diff     z    p
##   X1         35.39       35.38    -0.01 -0.02 0.98
##   X2        100.86      100.27    -0.60 -1.01 0.31
##   X3        179.31      179.48     0.17  0.10 0.92
##   X4         30.38       30.49     0.11  0.38 0.71
......

cb$overall

##           chisquare df p.value
## unstrat        5.56 13  0.9607
```

Univariate Uplift Analysis

The function `explore` computes the average value of the response variable for each predictor by treatment indicator

A convenient formula interface used by most functions in `uplift` includes a special term of the form `trt()` to mark the treatment variable: `response ~ trt(treatment) +  var1 + var2 + ...`

Let's look at an example:

```
eda <- explore(response ~ trt(treatment) + X1, nbins = 4,
    data = bankDM.train)
eda

## $X1
##          N(Treat=0) N(Treat=1) Ybar(Treat=0) Ybar(Treat=1) Uplift
## [20,27]         612        636        0.1438        0.3412 0.1974
## (27,34]         535        512        0.1533        0.3691 0.2159
## (34,43]         537        539        0.1583        0.3673 0.2091
## (43,61]         505        503        0.1406        0.4712 0.3306
......
```

Preliminary Variable Screening

The function `niv` produces a **net information value (NIV)** for each predictor (Larsen, 2010)

Extension of the **information value (IV)**, commonly used in credit risk scorecard applications (Anderson, 2007)

Helpful exploratory tool to (preliminary) determine the predictive power of each variable for uplift.

```
niv_res <- niv(modelForm, B = 20, nbins = 4, plotit = FALSE,
    data = bankDM.train)
niv_res$niv_val[order(niv_res$niv_val[, 3], decreasing = TRUE),
    ]

##        niv penalty adj_niv
## X13 4.8885  0.5397  4.3488
## X1  1.5945  0.1202  1.4743
## X6  1.4480  0.1770  1.2710
## X9  1.2260  0.2450  0.9810
## X11 1.0755  0.1444  0.9311
......
```

Causal Conditional Inference Forest (CCIF) - Pseudocode

Algorithm 1 Causal conditional inference tree

```
for each terminal node do
    Test the global null hypothesis H_0 of no interaction effect between the
    treatment A and any of the p predictors at a level of significance α based
    on a permutation test (Strasser and Weber, 1999)
    if the null hypothesis H_0 cannot be rejected then
        Stop
    else
        Select the j*th predictor X_{j*} with the strongest interaction effect (i.e.,
        the one with the smallest adjusted P value)
        Choose a partition Ω* of the covariate X_{j*} in two disjoint sets M ⊂
        X_{j*} and X_{j*} \ M based on the G^2(Ω) split criterion
    end if
end for
```

$$G^2(\Omega) = \frac{(L-4)\{\overbrace{(\bar{Y}_{n_L}(1) - \bar{Y}_{n_L}(0))}^{\text{Left Node}} - \overbrace{(\bar{Y}_{n_R}(1) - \bar{Y}_{n_R}(0))}^{\text{Right Node}}\}^2}{\hat{\sigma}^2\{1/L_{n_L}(1) + 1/L_{n_L}(0) + 1/L_{n_R}(1) + 1/L_{n_R}(0)\}}$$

Details in Guelman, Guillén and Pérez-Marín, 2014, **IME**. *Accepted.*

Fitting a CCIF

`ccif` implements recursive partitioning in a causal conditional inference framework.

```
ccif_fit1 <- ccif(modelForm, data = bankDM.train, ntree = 1000,
    split_method = "Int", distribution = approximate(B = 999),
    verbose = TRUE)
```

Table: Some **ccif** options

`ccif` argument	Description
`mtry`	Number of variables to be tested in each node
`ntree`	Number of trees in the forest
`split_method`	Split criteria: `"KL"`, `"ED"`, `"Int"` or `"L1"`
`interaction.depth`	The maximum depth of variable interactions
`pvalue`	Maximum acceptable p-value required to make a split
`bonferroni`	Apply Bonferroni adjustment to pvalue
`minsplit`	Minimum number of obs. for a split to be attempted
`...`	Additional args. passed to `independence_test{coin}`.

Standard Generic Functions for "ccif" Objects

`summary` and `predict` S3 methods for objects of class `"ccif"`

```
class(ccif_fit1)

## [1] "ccif"

summary(ccif_fit1)$importance

##    var rel.imp
## 1  X13  31.292
## 2   X1  21.136
## 3  X10   9.516
## 4  X12   7.206
## 5   X8   4.133
......

pred_ccif <- predict(ccif_fit1, bankDM.test)
head(pred_ccif, 4)

##      pr.y1_ct1 pr.y1_ct0
## [1,]    0.3513    0.1508
## [2,]    0.3541    0.1480
## [3,]    0.3543    0.1493
## [4,]    0.3478    0.1528
```

Evaluating Model Performance

Once we have a set of predictions, we can use `performance` to compute the uplift by decile.

```
perf_ccif <- performance(pred_ccif[, 1], pred_ccif[, 2],
    bankDM.test$response, bankDM.test$treatment, groups = 10)
perf_ccif

##       group n.ct1 n.ct0 n.y1_ct1 n.y1_ct0 r.y1_ct1 r.y1_ct0  uplift
##  [1,]     1   103    85       57        8   0.5534  0.09412 0.45928
##  [2,]     2    89    99       39        9   0.4382  0.09091 0.34729
##  [3,]     3    94    93       45       15   0.4787  0.16129 0.31743
##  [4,]     4    96    92       36       17   0.3750  0.18478 0.19022
##  [5,]     5    95    93       40       16   0.4211  0.17204 0.24901
##  [6,]     6    87   100       37       15   0.4253  0.15000 0.27529
##  [7,]     7   102    86       34       16   0.3333  0.18605 0.14729
##  [8,]     8    90    97       24       19   0.2667  0.19588 0.07079
##  [9,]     9    87   101       22       14   0.2529  0.13861 0.11426
## [10,]    10    95    93       26        9   0.2737  0.09677 0.17691
......
```

Evaluating Model Performance

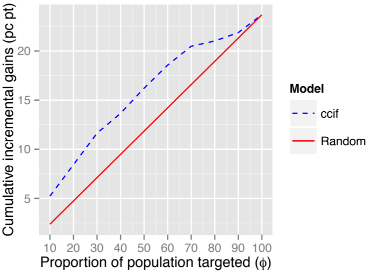
The **Qini curve** (Radcliffe, 2007) is a two-dimensional depiction of model performance for uplift models (extension of *Gains curve*).

The **Qini-coefficient** defined as the *area between the Qini curve and the Random curve*, and gives a single estimate of uplift model performance.

The function `qini` can be used to obtain both the Qini curve and the Qini-coefficient from a `"performance"` object.

```
qini(perf_ccif, plotit = TRUE)$Qini

## [1] 0.02906
```



Implementing Alternative Uplift Methods

upliftRF implements **Uplift Random Forest**

```
upliftRF_fit1 <- upliftRF(modelForm, data = bankDM.train, ntree = 1000,
    interaction.depth = 3, split_method = "KL", minsplit = 50)
```

upliftKNN implements **Uplift k-nearest neighbor**

```
upliftKNN_fit1 <- upliftKNN(bankDM.train[, 1:13], bankDM.test[,
    1:13], bankDM.train$response, bankDM.train$treatment,
    k = 5, dist.method = "euclidean", p = 2, ties.meth = "min",
    agg.method = "mean")
```

rvtu implements the **Modified outcome method**

```
bankDM.train.mom <- rvtu(modelForm, data = bankDM.train,
    method = "undersample")
glm.mom <- glm(modelForm.mom, data = bankDM.train.mom,
    family = "binomial")
```

tian_transf implements the **Modified covariate method**

```
bankDM.train.mcm <- tian_transf(modelForm,
    data = bankDM.train, method = "undersample")
glm.mcm <- glm(modelForm.mcm, data = bankDM.train.mcm,
    family = "binomial")
```

Performance Comparison

Table: Uplift from targeting **top 3 deciles** – Test sample

	Treatment xSell (%)	Control xSell (%)	uplift (%)
`ccif`	49.30	11.55	**37.75**
`mcm`	47.74	11.91	35.83
`mom`	47.18	11.83	35.35
`upliftRF`	46.44	11.56	34.88
`upliftKNN`	40.87	15.38	25.49
`Random`	38.40	14.80	23.60

Profiling Clients Based on Selected Model

modelProfile can be used to profile a fitted uplift model: given a vector of uplift predictions, it computes basic summary statistics for each predictor by score quantile (optionally, LaTex output).

```
modelProfile(uplift_pred_ccif ~ X1 + X10 + X12 + X8 + X4 +
    X2, data = bankDM.test, groups = 10, group_label = "D",
    digits_numeric = 1, LaTex = FALSE)[-2, ]

##
##          Group
##           1    2   3   4   5   6   7   8   9   10  All
##      n    188  188 187 188 188 187 188 187 188 188 1877
##  X1  Avg. 52   39  32  30  33  41  40  31  28  28   35
##  X10 Avg. 95   85  89  102 107 113 109 86  100 114 100
##  X12 Avg. 189  195 177 187 222 220 200 179 196 236 200
##  X8  Avg. 10   10  10  10  11  10  10  10  10  10   10
##  X4  Avg. 32   30  33  30  32  30  30  30  31  32   31
##  X2  Avg. 102  100 101 100 100 101 98  104 101 96  100
```

1 Introduction
- Motivation
- The problem

2 Customers who react

3 uplift R package

4 Price elasticity

Retention combined with price changes

Guelman, L. and Guillen, M. (2014) A causal inference approach to measure price elasticity in automobile insurance, **Expert Systems with Applications**, 41(2), 387-396.

The role of price in customer retention

- Understanding price sensitivities at the individual policy holder level is extremely valuable for insurers.
- A rate increase has a direct impact on the premiums customers are paying, but there is also a causal effect on the customers decision to renew the policy term.
- It is difficult to measure price elasticity from most insurance datasets, as historical rate changes are reflective of a risk-based pricing exercise, therefore they are not assigned at random across the portfolio of policyholders.
- We propose a causal inference framework to measure price elasticity in the context of auto insurance.

Data considerations

1. The gold standard for measuring causal effects (i.e., effects attributable to treatments) is to obtain experimental data
2. In the context of price-elasticity, this would involve randomizing policyholders to various rate change levels (the latter playing the role of the "treatments")
3. This condition rarely holds in practice, as usually rate changes are assigned to policyholders based on a risk-based pricing model. Thus we end up with observational data (as opposed to experimental)

Data considerations

1. The good news is that under certain data conditions (Rosenbaum and Rubin, 1983) it is still possible to obtain unbiased estimates of causal effect from observational data – that is, we can obtain unbiased estimates of price elasticities
2. Two key concepts come into play here: propensity scores and matching algorithms
3. These methods can be used to reconstruct a "sort of" randomized study from observational data

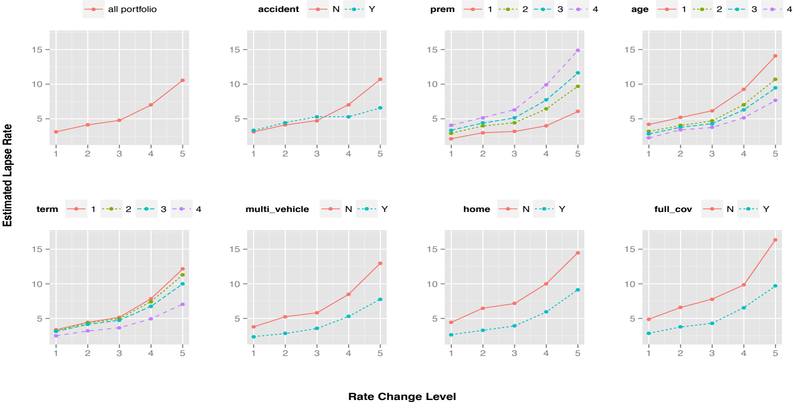
Methodology

- L policyholders, $\ell = \{1, 2, \ldots, L\}$.
- vector of pre-treatment covariates $\mathbf{x}_\ell$.
- ordered treatment variable t (rate change levels), which takes values $t = \{1 < 2 < \ldots < T\}$ on a set $\Im$.
- $Z_{\ell t}$ set of T binary treatment indicators, $Z_{\ell t} = 1$ if subject ℓ received treatment t, and $Z_{\ell t} = 0$ otherwise.
- potential responses $r_{\ell t}$, renewal outcome that would be observed from policyholder ℓ if assigned to treatment t .
- observed response for subject ℓ is $R_\ell = \sum_{t \in \Im} Z_{\ell t} r_{\ell t}$.
- Our interest is to estimate price elasticity, defined as the renewal outcomes that result and are caused by the price change interventions.

Empirical application: the data

- $L = 329,000$ auto insurance policies from a major Canadian insurer that have been given a renewal offer from June-2010 to May-2012 consisting on a new rate either lower, equal or higher than the current rate.
- more than 60 pre-treatment covariates (characteristics of the policy, the vehicle and driver).
- the treatment is the rate change: percentage change in premium from the current to the new rate, categorized into 5 ordered values $t = \{1 < 2 < \ldots < 5\}$.
- response variable: renewal outcome of the policy, measured 30 days after the effective date of the new policy term

Empirical application: estimated lapse rate



Empirical application: managerial implications

Which rate change should be applied to each policyholder to maximize the overall expected profit for the company subject to a fixed overall retention rate?

$$\underset{Z_{\ell t} \forall \ell \forall t}{\text{Max}} \sum_{\ell=1}^{L} \sum_{t=1}^{T} Z_{\ell t} \left[P_\ell (1 + RC_t)(1 - \hat{LR}_{\ell t})(1 - \hat{r}_{\ell t}) \right]$$

where P_ℓ is the current premium, RC_t is the actual rate change level associated with treatment t, $\hat{LR}_{\ell t}$ the predicted loss ratio (i.e., the ratio of the predicted insurance losses relative to premium), $\hat{r}_{\ell t}$ is the lapse probability of subject ℓ if exposed to rate change level t, and α the overall lapse rate of the portfolio.

Empirical application: managerial implications

The expected function to maximize is the expected profit of the portfolio

$$\underset{Z_{\ell t}\forall\ell\forall t}{\text{Max}} \sum_{\ell=1}^{L}\sum_{t=1}^{T} Z_{\ell t}\left[P_{\ell}(1+RC_t)(1-\hat{LR}_{\ell t})(1-\hat{r}_{\ell t})\right]$$

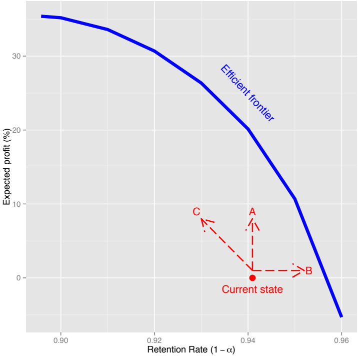
subject to the following constraints

$$\sum_{t=1}^{T} Z_{\ell t} = 1 \quad : \forall \ell$$

$$Z_{\ell t} \in \{0,1\}$$

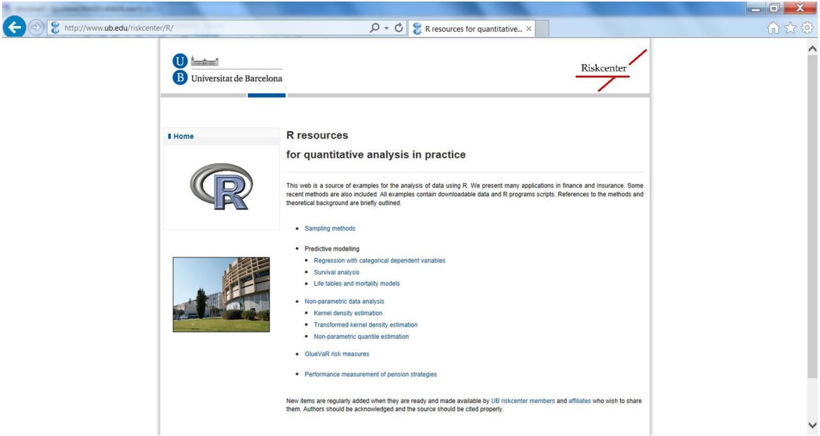
$$\sum_{\ell=1}^{L}\sum_{t=1}^{T} Z_{\ell t}\hat{r}_{\ell t}/L \leq \alpha$$

Empirical application: managerial implications



Conclusions

- We have presented an approach to estimate price elasticity functions which allows for heterogeneous causal effects as a result of rate change interventions
- The model can assist managers in selecting an optimal rate change level for each policyholder for the purpose of maximizing the overall profits for the company
- Valuable insights can be gained by knowing the current company's position of growth and profitability relative to the optimal values given by the efficient frontier
- The managerial decision is to determine in which direction the company should move towards the frontier, as each decision point places a different weight on each of these objectives.



http://www.ub.edu/riskcenter/R/

R resources for quantitative...

Universitat de Barcelona

Riskcenter

Home

R resources

for quantitative analysis in practice

This web is a source of examples for the analysis of data using R. We present many applications in finance and Insurance. Some recent methods are also included. All examples contain downloadable data and R programs scripts. References to the methods and theoretical background are briefly outlined.

- Sampling methods
- Predictive modelling
  - Regression with categorical dependent variables
  - Survival analysis
  - Life tables and mortality models
- Non-parametric data analysis
  - Kernel density estimation
  - Transformed kernel density estimation
  - Non-parametric quantile estimation
- GlueVaR risk measures
- Performance measurement of pension strategies

New items are regularly added when they are ready and made available by UB riskcenter members and affiliates who wish to share them. Authors should be acknowledged and the source should be cited properly.

www.ub.edu/riskcenter/R