Customer Analytics

Logistic regression, churn management

Tilburg University

Agenda

- Case: proactive churn management
- Logistic regression
- Deviance & R^2
- Overfitting and cross-validation
- ROC/Lift curves
- Optimal targeting

Churn management

Blattberg, Kim and Neslin (2008) Ch. 24

Tilburg University

from lecture 1: Customer lifecycle

Customer **development**: change in behavior over time: buying more (up-selling) or different things (cross-selling)

Customer **acquisition**: how customers are “born” or first contact with the firm.

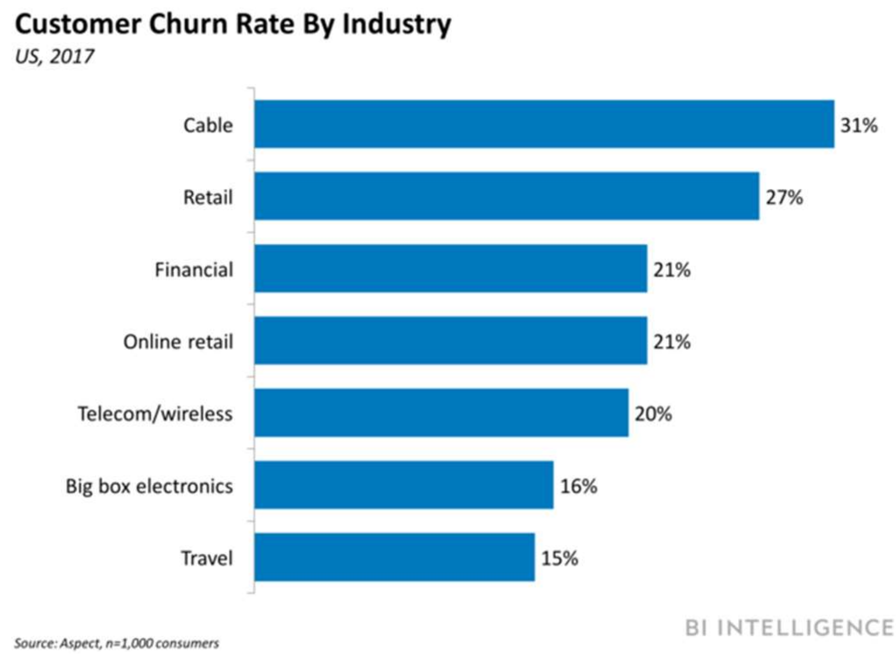
Customer **retention**: preventing customer “death” or churn.

Marketing is about acquiring, developing and retaining customers

Case: preventing churn

Customer churn (customer quitting)

Typically, it is cheaper to retain existing customers than to acquire new customers



Case: proactive churn management

- Reactive churn management
  - Customer already almost out of the door
- **Proactive**: identify customers who have high probability of churning.
  - Need model to predict churn
- Contact customer ahead of churn (quit)
  - Offer incentive or service to prevent customer from churning

How can companies predict who will churn?

- Use model to predict quit or stay based on all variables you have at your disposal
  - Past usage
  - Contract type
  - Marketing
  - Demographics

- Use the results of this model to predict likelihood of quitting for new/current customers.

Logistic regression

[Blattberg, Kim and Neslin (2008) Ch. 15: 377-85]

Quantitative_Models_in_Marketing_Research: Chapter 4, a binomial dependent variable

Modeling response

Dependent variable is <u>binary</u>

$$Y \in \{0\,,1\}$$

- Mutually exclusive: can't be both
- Collectively exhaustive: can't be neither
- Arbitrary what we call "1" and "0"

examples

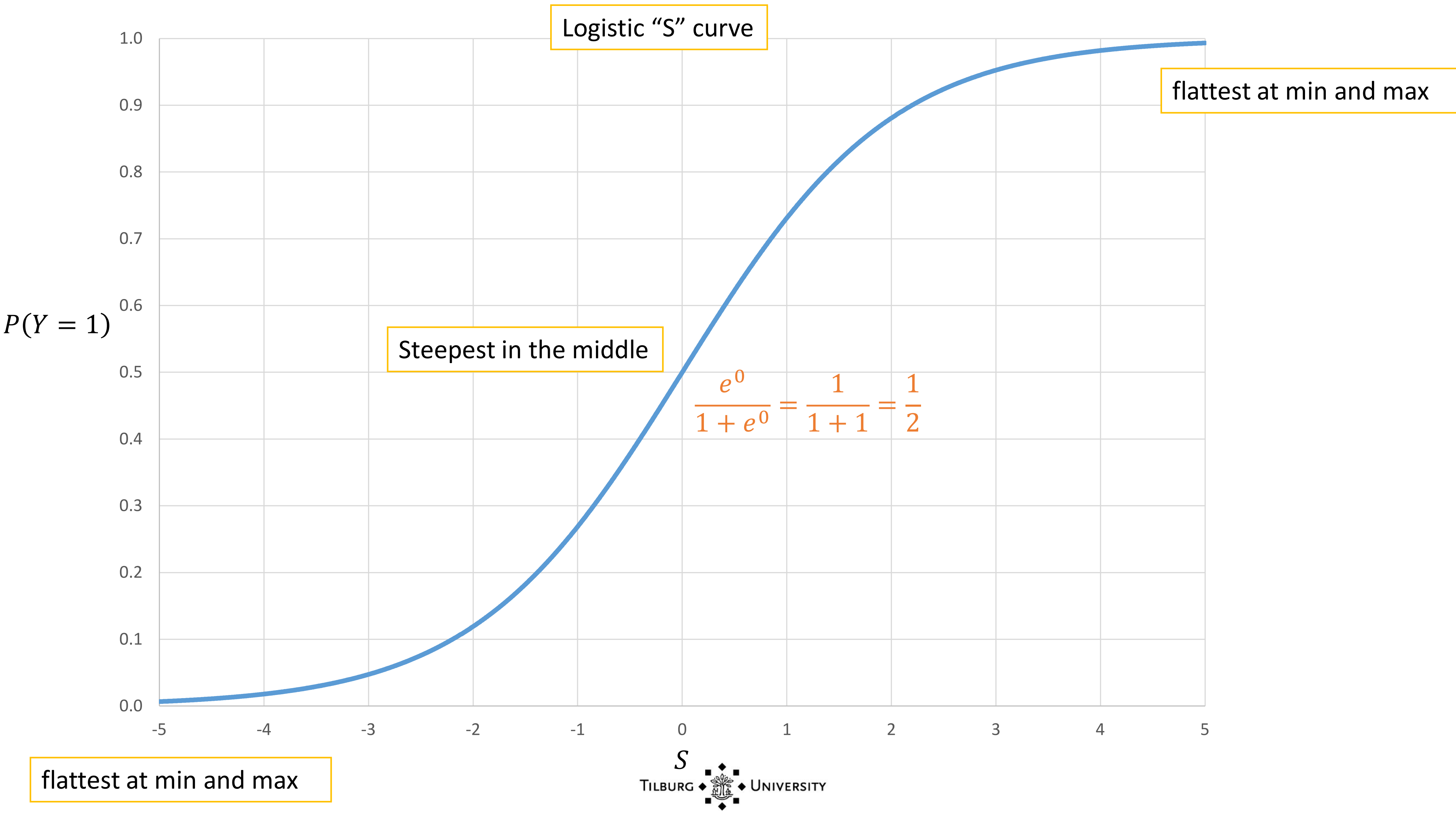
$y_i = \mathbf{1}$	$y_i = \mathbf{0}$
Churn/quit	Didn't churn/quit
Joined/subscribed	Didn't join/subscribe
Responded to marketing	Didn't respond
Bought	Didn't buy
Chose option A	Chose option B
Upgraded	Didn't upgrade
Changed service	Didn't change service

Modeling the probability

$$P(Y=1) = \frac{\exp(S)}{1+\exp(S)}$$

$$S = \beta_0 + \sum_{j=1}^{p} \beta_j X_j = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \cdots \beta_p x_p$$

$$\exp(1) = e \approx 2.72$$


Logistic "S" curve
flattest at min and max
Steepest in the middle
$P(Y = 1)$
$\frac{e^0}{1+e^0} = \frac{1}{1+1} = \frac{1}{2}$
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
-5
-4
-3
-2
-1
0
1
2
3
4
5
S
flattest at min and max
Tilburg University

Linear model for log odds

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \cdots \beta_p x_p$$

where $p = \mathrm{P}(\mathrm{Y} = 1)$.

$$p = \frac{1}{4} \Leftrightarrow \frac{p}{1-p} = \frac{1}{3}$$

Generalized linear models

$$g(\mu) = \beta_0 + \beta_1 x_1 + \cdots$$

where μ is the mean and $g(\mu)$ is the link function.

Multiplicative effect on odds

Another way of writing previous equation

$$\left(\frac{p}{1-p}\right) = \exp(\beta_0)\exp(\beta_1 x_1)\exp(\beta_2 x_2)\exp(\beta_3 x_3)\ldots\exp(\beta_p x_p)$$

Multiplicative effect for a unit increase in odds x_k:

$$\exp(\beta_k)$$

% change in odds, $(\exp(\beta_k) - 1) * 100$

Maximum Likelihood Estimation

Likelihood = probability of the data given parameters. $i = 1, \dots I$ observations

$$L\ (\beta) = \prod_i p_i^{y_i}\ (1 - p_i)^{1-y_i}$$

$$LL\ (\beta) = \sum_i y_i \log(p_i) + (1 - y_i) \log(1 - p_i)$$

As with frequentist and Empirical Bayes methods, we choose β to maximize log-likelihood:

$$\hat{\beta} = \arg\max_{\beta} LL\ (\beta)$$

Telco churn (inspecting)

7000 obs., 20 variables

RStudio: Notebook Output

	gender <fctr>	SeniorCitizen <fctr>	Partner <fctr>	Dependents <fctr>	tenure <int>	PhoneService <fctr>	MultipleLines <fctr>	InternetService <fctr>	OnlineSecurity <fctr>	OnlineBackup <fctr>	DeviceProtection <fctr>
1	Female	0	Yes	No	1	No	No	DSL	No	Yes	No
2	Male	0	No	No	34	Yes	No	DSL	Yes	No	Yes
3	Male	0	No	No	2	Yes	No	DSL	Yes	Yes	No
4	Male	0	No	No	45	No	No	DSL	Yes	No	Yes
5	Female	0	No	No	2	Yes	No	Fiber optic	No	No	No
6	Female	0	No	No	8	Yes	Yes	Fiber optic	No	No	Yes

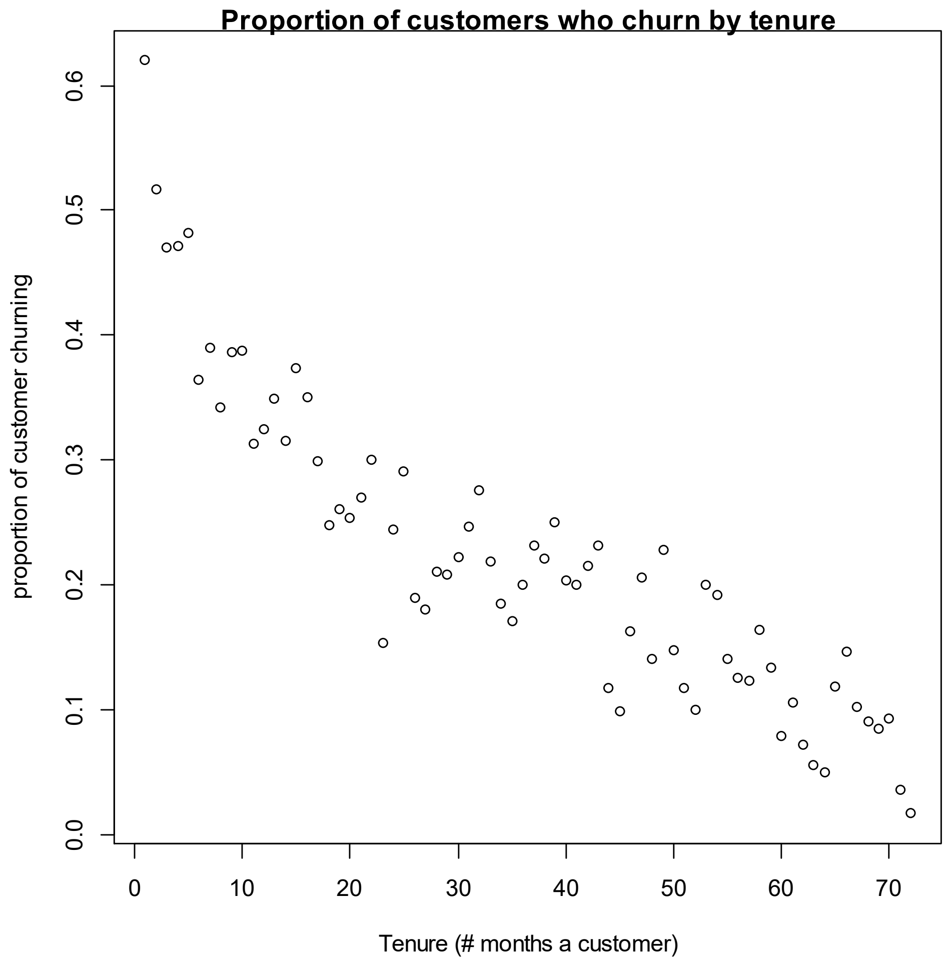
6 rows | 1-12 of 20 columns

Telco churn

RStudio: Notebook Output

TechSupport <fctr>	StreamingTV <fctr>	StreamingMovies <fctr>	Contract <fctr>	PaperlessBilling <fctr>	PaymentMethod <fctr>	MonthlyCharges <dbl>	TotalCharges <dbl>	Churn <fctr>
No	No	No	Month-to-month	Yes	Electronic check	29.9	0.0299	No
No	No	No	One year	No	Mailed check	57.0	1.8895	No
No	No	No	Month-to-month	Yes	Mailed check	53.9	0.1082	Yes
Yes	No	No	One year	No	Bank transfer (automatic)	42.3	1.8408	No
No	No	No	Month-to-month	Yes	Electronic check	70.7	0.1517	Yes
No	Yes	Yes	Month-to-month	Yes	Electronic check	99.7	0.8205	Yes

6 rows | 13-21 of 20 columns

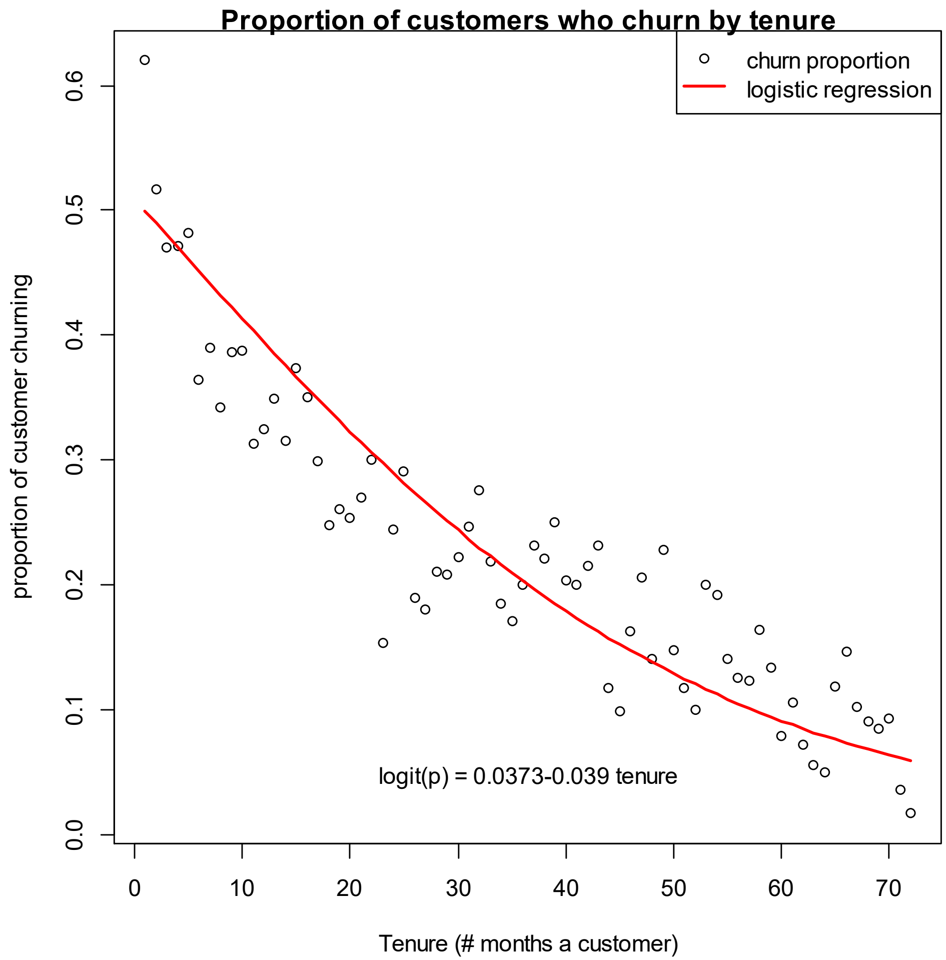


TILBURG UNIVERSITY

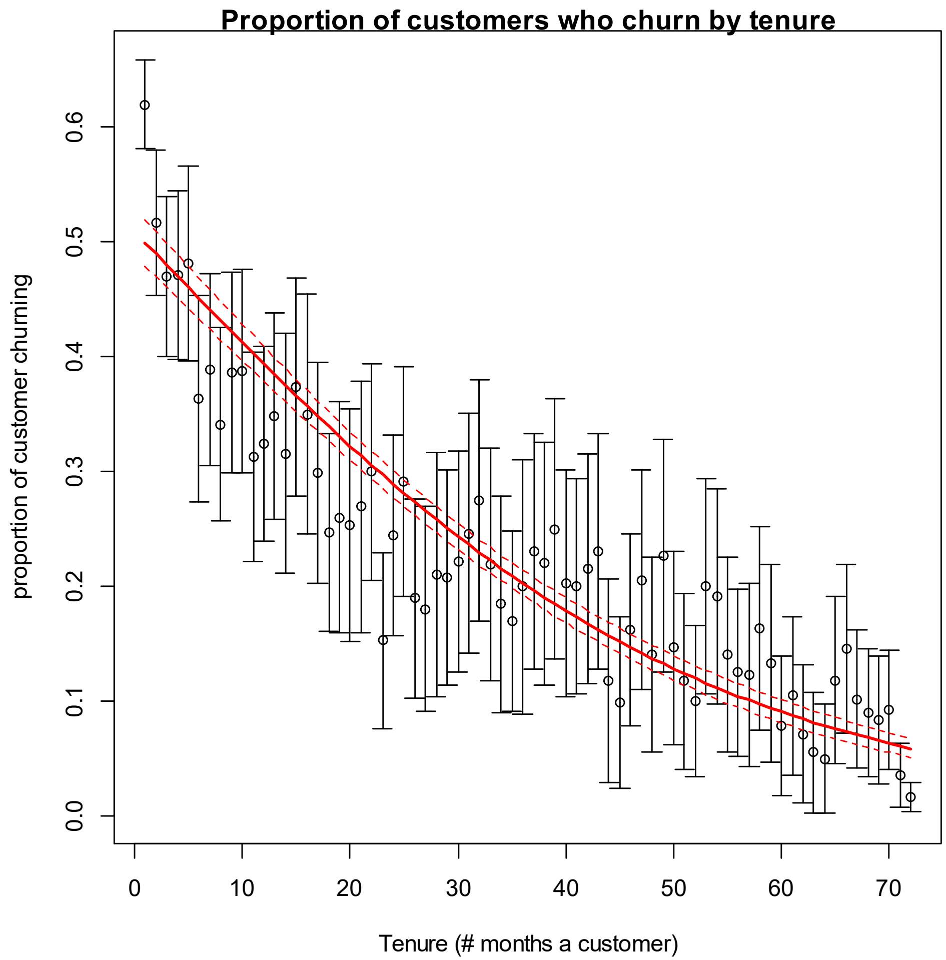
Model 0:

$$y = \begin{cases} 0, & \text{Churn} = \text{"No "} \\ 1, & \text{Churn} = \text{"Yes"} \end{cases}$$

$$P(\mathrm{y} = 1) = \frac{\exp(\beta_0 + \beta_1 \text{ Tenure})}{1 + \exp(\beta_0 + \beta_1 \text{ Tenure})}$$



TILBURG ◆ UNIVERSITY



TILBURG UNIVERSITY

Model 1

$$P(\text{Churn} = 1) = \frac{\exp(x'\beta)}{1 + \exp(x'\beta)}$$

$$x'\beta = \beta_0 + \beta_1 \text{Male} + \beta_2 \text{ Senior} + \beta_3 \text{Partner} + \cdots + \beta_{23} \text{Tenure}$$

In R,

```
model_1 <- glm(Churn ~ . , data=telco, family="binomial")
```

"everything else"

Face validity:
Do the signs of significant coefficients make sense?

```
Call:
glm(formula = Churn ~ ., family = "binomial", data = telco)

Deviance Residuals:
   Min      1Q  Median      3Q     Max
-1.918  -0.679  -0.286   0.728   3.430

Coefficients:
                                     Estimate Std. Error z value             Pr(>|z|)
(Intercept)                          1.165287   0.815135    1.43               0.1528
genderMale                          -0.021833   0.064804   -0.34               0.7362
SeniorCitizen1                       0.216775   0.084530    2.56               0.0103 *
PartnerYes                          -0.000384   0.077830    0.00               0.9961
DependentsYes                       -0.148488   0.089731   -1.65               0.0980 .
tenure                              -0.060588   0.006236   -9.72 < 0.0000000000000002 ***
PhoneServiceYes                      0.171468   0.648692    0.26               0.7915
MultipleLinesYes                     0.448395   0.177256    2.53               0.0114 *
InternetServiceFiber optic           1.747475   0.798080    2.19               0.0286 *
InternetServiceNo                   -1.786295   0.807268   -2.21               0.0269 *
OnlineSecurityYes                   -0.205420   0.178688   -1.15               0.2503
OnlineBackupYes                      0.026042   0.175401    0.15               0.8820
DeviceProtectionYes                  0.147375   0.176374    0.84               0.4034
TechSupportYes                      -0.180497   0.180602   -1.00               0.3176
StreamingTVYes                       0.590507   0.326309    1.81               0.0703 .
StreamingMoviesYes                   0.599296   0.326684    1.83               0.0666 .
ContractOne year                    -0.660795   0.107585   -6.14     0.000000000814591 ***
ContractTwo year                    -1.357106   0.176445   -7.69     0.000000000000015 ***
PaperlessBillingYes                  0.342354   0.074495    4.60     0.000004314313182 ***
PaymentMethodCredit card (automatic) -0.087792  0.114079   -0.77               0.4416
PaymentMethodElectronic check        0.304467   0.094497    3.22               0.0013 **
PaymentMethodMailed check           -0.057587   0.114911   -0.50               0.6163
MonthlyCharges                      -0.040344   0.031755   -1.27               0.2039
TotalCharges                         0.328940   0.070628    4.66     0.000003202625160 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 8143.4  on 7031  degrees of freedom
Residual deviance: 5826.3  on 7008  degrees of freedom
AIC: 5874

Number of Fisher Scoring iterations: 6
```

Interpreting coefficients

What is the effect of these variables on the odds of churn?

remember $\exp(\beta)$ is the multiplicative effect of unit change of x on odds of y.

```
Call:
glm(formula = Churn ~ ., family = "binomial", data = telco)

Deviance Residuals:
   Min      1Q  Median      3Q     Max
-1.918  -0.679  -0.286   0.728   3.430

Coefficients:
                                      Estimate Std. Error z value            Pr(>|z|)
(Intercept)                           1.165287   0.815135    1.43              0.1528
genderMale                           -0.021833   0.064804   -0.34              0.7362
SeniorCitizen1                        0.216775   0.084530    2.56              0.0103 *
PartnerYes                           -0.000384   0.077830    0.00              0.9961
DependentsYes                        -0.148488   0.089731   -1.65              0.0980 .
tenure                               -0.060588   0.006236   -9.72 < 0.0000000000000002 ***
PhoneServiceYes                       0.171468   0.648692    0.26              0.7915
MultipleLinesYes                      0.448395   0.177256    2.53              0.0114 *
InternetServiceFiber optic            1.747475   0.798080    2.19              0.0286 *
InternetServiceNo                    -1.786295   0.807268   -2.21              0.0269 *
OnlineSecurityYes                    -0.205420   0.178688   -1.15              0.2503
OnlineBackupYes                       0.026042   0.175401    0.15              0.8820
DeviceProtectionYes                   0.147375   0.176374    0.84              0.4034
TechSupportYes                       -0.180497   0.180602   -1.00              0.3176
StreamingTVYes                        0.590507   0.326309    1.81              0.0703 .
StreamingMoviesYes                    0.599296   0.326684    1.83              0.0666 .
ContractOne year                     -0.660795   0.107585   -6.14   0.000000000814591 ***
ContractTwo year                     -1.357106   0.176445   -7.69   0.000000000000015 ***
PaperlessBillingYes                   0.342354   0.074495    4.60   0.000004314313182 ***
PaymentMethodCredit card (automatic) -0.087792   0.114079   -0.77              0.4416
PaymentMethodElectronic check         0.304467   0.094497    3.22              0.0013 **
PaymentMethodMailed check            -0.057587   0.114911   -0.50              0.6163
MonthlyCharges                       -0.040344   0.031755   -1.27              0.2039
TotalCharges                          0.328940   0.070628    4.66   0.000003202625160 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 8143.4  on 7031  degrees of freedom
Residual deviance: 5826.3  on 7008  degrees of freedom
AIC: 5874

Number of Fisher Scoring iterations: 6
```

```
Call:
glm(formula = Churn ~ ., family = "binomial", data = telco)

Deviance Residuals: 
   Min      1Q  Median      3Q     Max  
-1.918  -0.679  -0.286   0.728   3.430  

Coefficients:
                                     Estimate Std. Error z value              Pr(>|z|)    
(Intercept)                          1.165287   0.815135    1.43                0.1528    
genderMale                          -0.021833   0.064804   -0.34                0.7362    
SeniorCitizen1                       0.216775   0.084530    2.56                0.0103 *  
PartnerYes                          -0.000384   0.077830    0.00                0.9961    
DependentsYes                       -0.148488   0.089731   -1.65                0.0980 .  
tenure                              -0.060588   0.006236   -9.72 < 0.0000000000000002 ***
PhoneServiceYes                      0.171468   0.648692    0.26                0.7915    
MultipleLinesYes                     0.448395   0.177256    2.53                0.0114 *  
InternetServiceFiber optic           1.747475   0.798080    2.19                0.0286 *  
InternetServiceNo                   -1.786295   0.807268   -2.21                0.0269 *  
OnlineSecurityYes                   -0.205420   0.178688   -1.15                0.2503    
OnlineBackupYes                      0.026042   0.175401    0.15                0.8820    
DeviceProtectionYes                  0.147375   0.176374    0.84                0.4034    
TechSupportYes                      -0.180497   0.180602   -1.00                0.3176    
StreamingTVYes                       0.590507   0.326309    1.81                0.0703 .  
StreamingMoviesYes                   0.599296   0.326684    1.83                0.0666 .  
ContractOne year                    -0.660795   0.107585   -6.14    0.000000000814591 ***
ContractTwo year                    -1.357106   0.176445   -7.69    0.000000000000015 ***
PaperlessBillingYes                  0.342354   0.074495    4.60    0.000004314313182 ***
PaymentMethodCredit card (automatic) -0.087792  0.114079   -0.77                0.4416    
PaymentMethodElectronic check        0.304467   0.094497    3.22                0.0013 ** 
PaymentMethodMailed check           -0.057587   0.114911   -0.50                0.6163    
MonthlyCharges                      -0.040344   0.031755   -1.27                0.2039    
TotalCharges                         0.328940   0.070628    4.66    0.000003202625160 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 8143.4  on 7031  degrees of freedom
Residual deviance: 5826.3  on 7008  degrees of freedom
AIC: 5874

Number of Fisher Scoring iterations: 6
```

Deviance

- Measures the distance between data and fit

$$\mathrm{Dev}(\beta) = -2\ \mathrm{LL}(\hat{\beta}) + C$$

- Residual deviance, $\mathrm{D} = \mathrm{Dev}(\hat{\beta})$

- Null deviance, $\mathrm{D}_0 = \mathrm{Dev}(\beta_{1:p} \ = 0)$ (only estimate the intercept, β_0)

- Proportion of deviance explained by all variables (except intercept):

$$R^2 = \frac{D_0 - D}{D_0} = 1 - \frac{D}{D_0}$$

$$D = 5826$$
$$D_0 = 8143$$

$$R^2 = 1 - \frac{5826}{8143} = .28$$

```
Call:
glm(formula = Churn ~ ., family = "binomial", data = telco)

Deviance Residuals:
   Min      1Q  Median      3Q     Max
-1.918  -0.679  -0.286   0.728   3.430

Coefficients:
                                      Estimate Std. Error z value            Pr(>|z|)
(Intercept)                           1.165287   0.815135    1.43              0.1528
genderMale                           -0.021833   0.064804   -0.34              0.7362
SeniorCitizen1                        0.216775   0.084530    2.56              0.0103 *
PartnerYes                           -0.000384   0.077830    0.00              0.9961
DependentsYes                        -0.148488   0.089731   -1.65              0.0980 .
tenure                               -0.060588   0.006236   -9.72 < 0.0000000000000002 ***
PhoneServiceYes                       0.171468   0.648692    0.26              0.7915
MultipleLinesYes                      0.448395   0.177256    2.53              0.0114 *
InternetServiceFiber optic            1.747475   0.798080    2.19              0.0286 *
InternetServiceNo                    -1.786295   0.807268   -2.21              0.0269 *
OnlineSecurityYes                    -0.205420   0.178688   -1.15              0.2503
OnlineBackupYes                       0.026042   0.175401    0.15              0.8820
DeviceProtectionYes                   0.147375   0.176374    0.84              0.4034
TechSupportYes                       -0.180497   0.180602   -1.00              0.3176
StreamingTVYes                        0.590507   0.326309    1.81              0.0703 .
StreamingMoviesYes                    0.599296   0.326684    1.83              0.0666 .
ContractOne year                     -0.660795   0.107585   -6.14    0.00000000814591 ***
ContractTwo year                     -1.357106   0.176445   -7.69    0.00000000000015 ***
PaperlessBillingYes                   0.342354   0.074495    4.60    0.00000431431318 2 ***
PaymentMethodCredit card (automatic) -0.087792   0.114079   -0.77              0.4416
PaymentMethodElectronic check         0.304467   0.094497    3.22              0.0013 **
PaymentMethodMailed check            -0.057587   0.114911   -0.50              0.6163
MonthlyCharges                       -0.040344   0.031755   -1.27              0.2039
TotalCharges                          0.328940   0.070628    4.66    0.00000320262516 0 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 8143.4  on 7031  degrees of freedom
Residual deviance: 5826.3  on 7008  degrees of freedom
AIC: 5874

Number of Fisher Scoring iterations: 6
```

In-sample vs. out-of-sample

- So far we've calculated deviance and R^2 using the same data we used to fit the model: **in-sample**

- But, all that matters is how well the model predicts new data, **out-of-sample**.

- Why is "new data" so important here?

  - Generalizeability (external validity) = the ability of the results from one data set to carry over to other data sets?

Reading: Blattberg, Kim and Neslin (2008) 11.4

TILBURG UNIVERSITY

Overfitting

- when a model estimated on a particular data set predicts well for <u>only</u> that data set, not others.

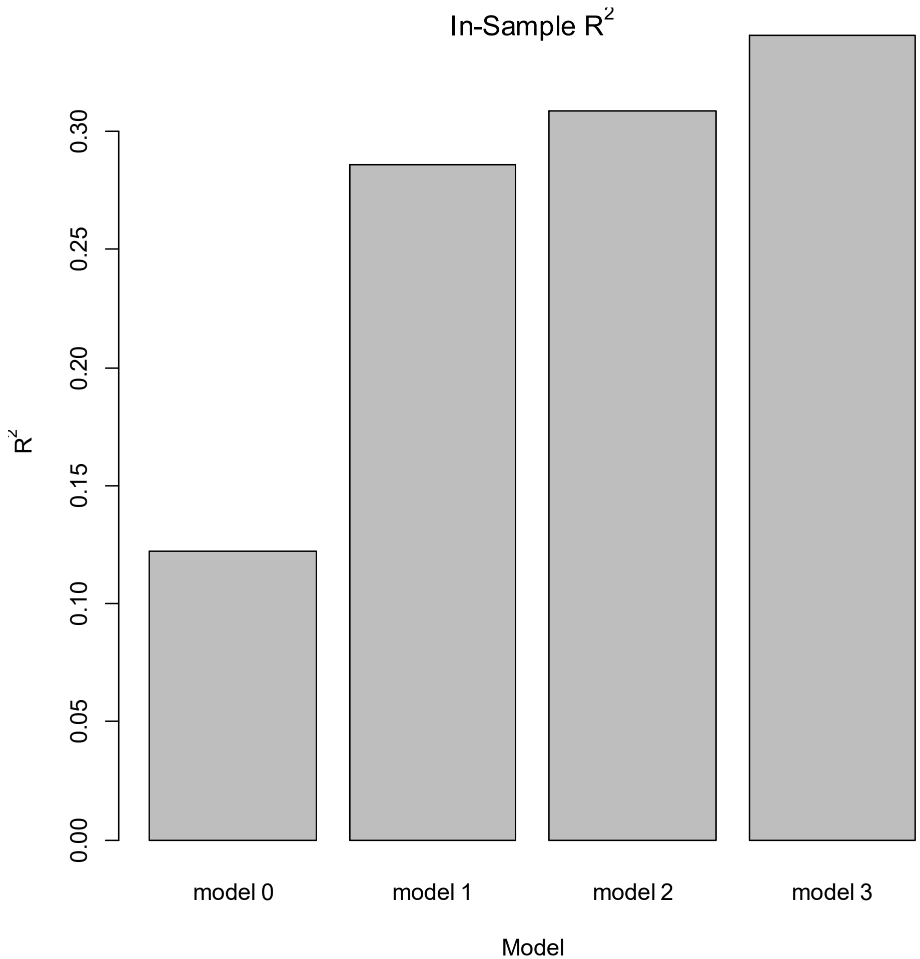
$$\mathrm{Dev}(\hat{\beta} \mid \mathrm{data}_{\mathrm{IS}}) < \mathrm{Dev}\,(\hat{\beta} \mid \mathrm{data}_{\mathrm{OOS}})$$

<u>Overfitting is a failure of generalizeability</u>

Example

- **Model 0:** already seen (1 coefficients, always excluding intercept)
- **Model 1**: already seen (24 coefficients)
- **Model 2**: Start with model 1, but instead of one coefficient for tenure, we have dummy variables for all but one of 72 levels (94 coefficients)
$$= \beta_0 + \cdots \beta_{23} 1\{\text{Tenure} = 2\} + \beta_{24} 1\{\text{Tenure} = 3\} + \cdots \beta_{93} 1\{\text{Tenure} = 72\}$$
- **Model 3**: Start with model 2, but add all interactions between tenure and payment type (307 coefficients):

$$= \beta_0 + \cdots \beta_{94} 1\{\text{Tenure} = 2\} 1\{\text{Payment} = \text{credit}\} + \beta_{95} 1\{\text{Tenure} = 2\} 1\{\text{Payment} = \text{echeck}\} + \beta_{96} 1\{\text{Tenure} = 2\} 1\{\text{Payment} = \text{check}\} + \cdots$$


In-Sample R^2
0.00
0.05
0.10
0.15
0.20
0.25
0.30
R^2
model 0
model 1
model 2
model 3
Model

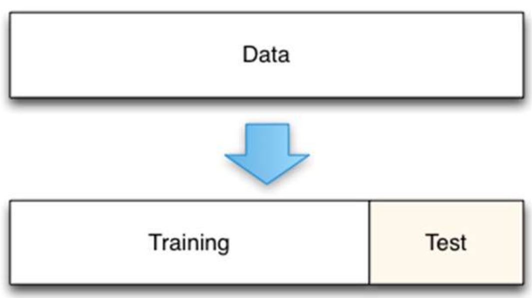
Out-of-sample

We mimic “new data” by holding out a part of the data.

Fit the model on one part, see how well it does on another part.

Cross-validation

- Simplest method: randomly split data into two parts:
  - "Training" (in-sample, calibration) = use this to fit model (typically 70-80%)
  - "Testing" (out-of-sample validation) = use this to make predictions



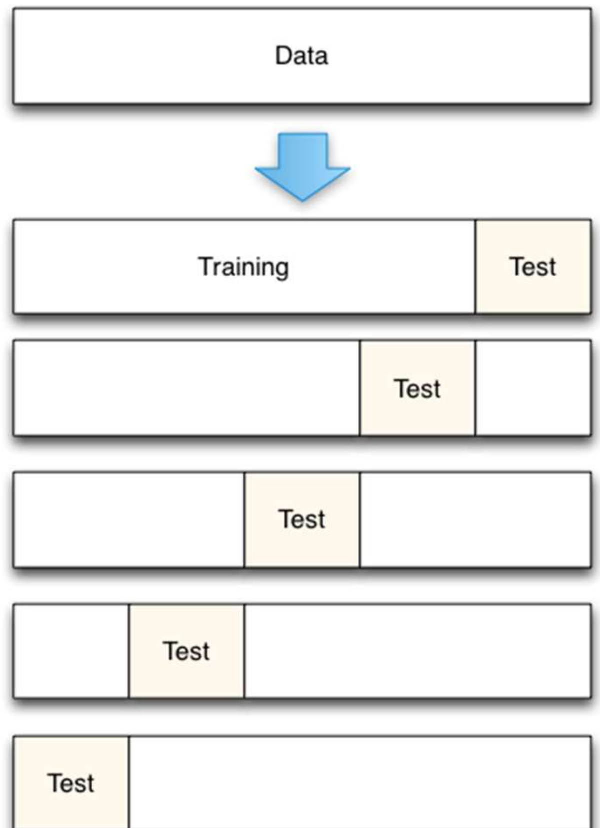
Problems

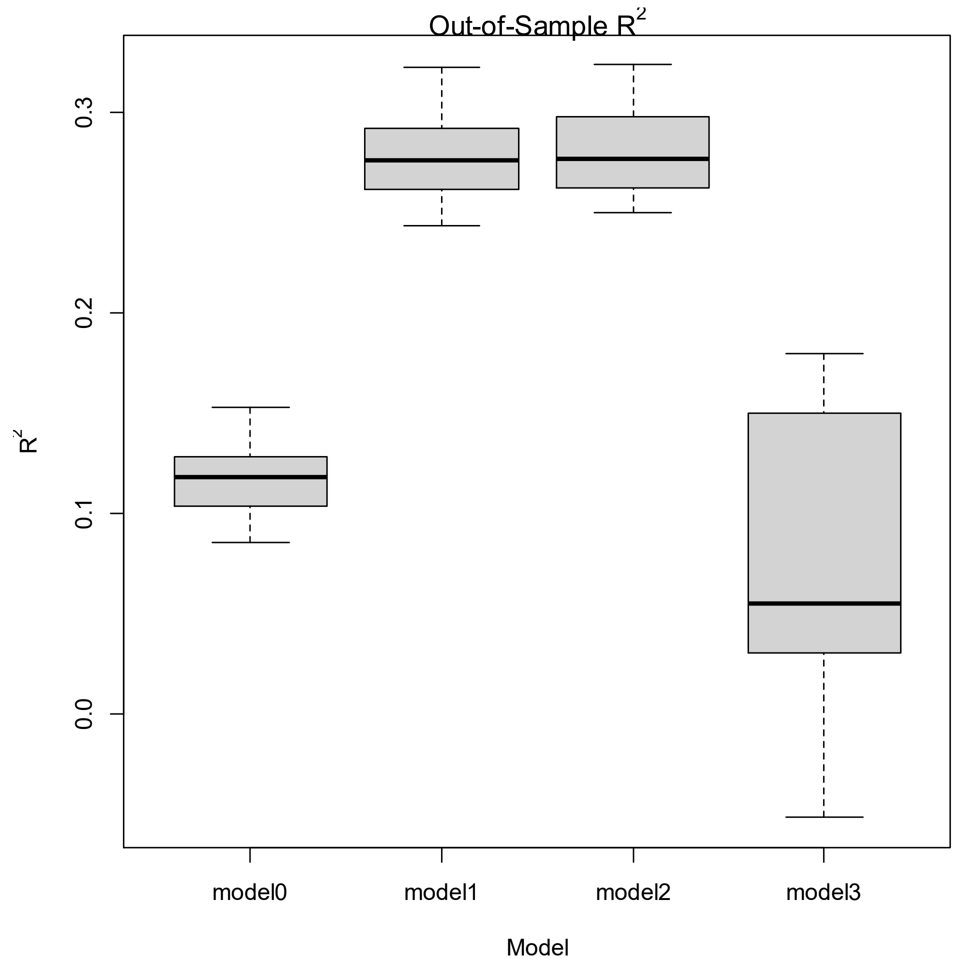
- Efficiency: some of the data is not used in model fitting, and maybe can be thought of as wasted

- Stability: what if, by chance, the validation sample is full of outliers?

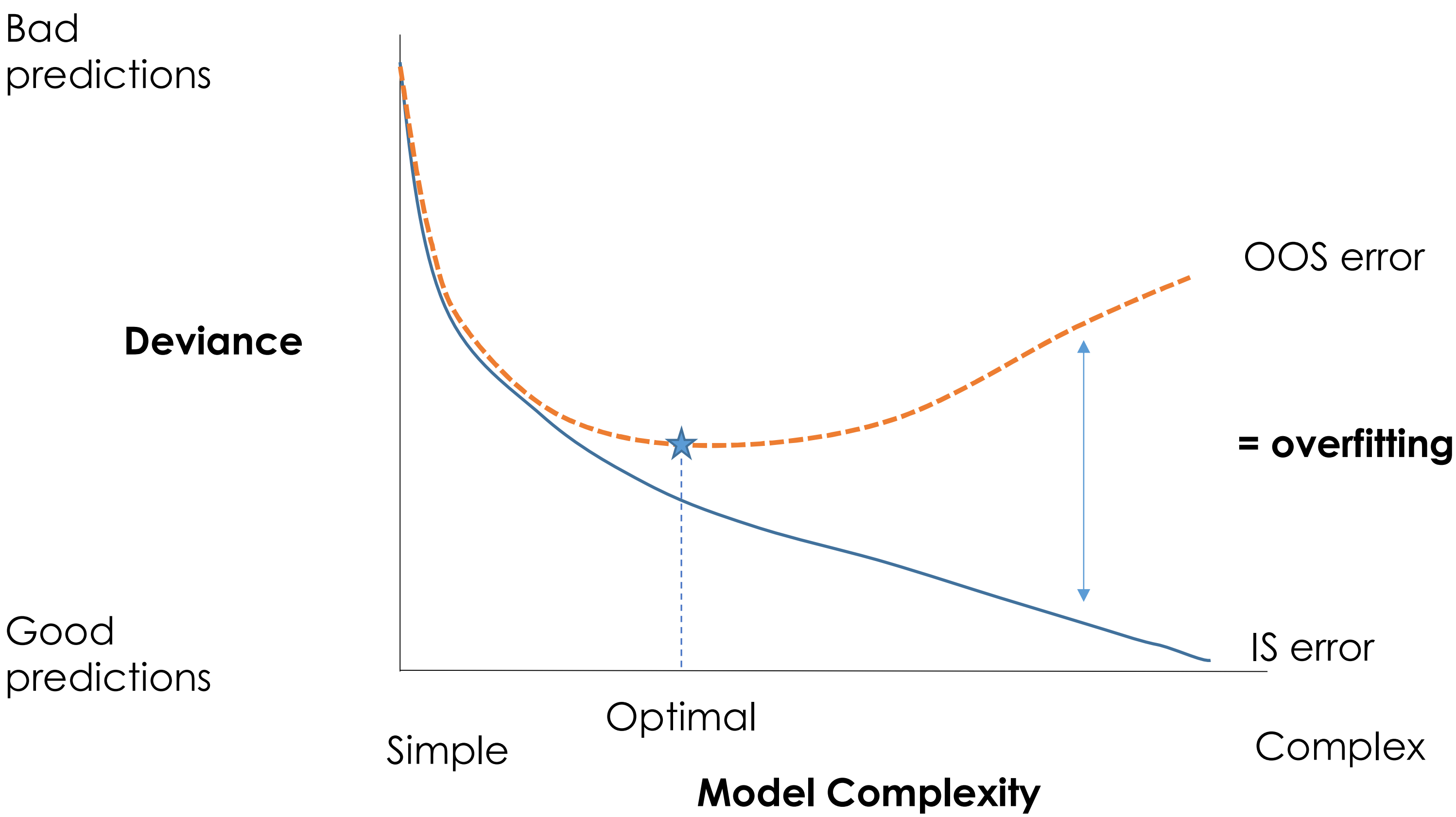
K-fold cross validation

- Here K = 5.
- Data randomly split into 5 equally sized groups of 20% each.
- 4 groups used to fit, one group to validate.
- Repeat so that all data is used.
- We now have a sample of K OOS R^2





TILBURG UNIVERSITY


Bad predictions
Deviance
Good predictions
OOS error
= overfitting
IS error
Optimal
Simple
Complex
Model Complexity

Other diagnostics

Classification (aka "confusion") matrix

		Predicted*		% correct
		0	**1**	
Observed	**0**	*True Negative (TN)*	*False Positive (FP)*	TN / (TN + FP)
	1	*False Negative (FN)*	*True Positive (TP)*	TP / (TP + FN)

Predicted $= 1\ if\ p_i > \bar{p}$ usually ($\bar{p} = 0.5$)

Accuracy = (TP + TN) / Total

"Hit Rate" = "Sensitivity" = True positive rate = TP / (TP + FN)

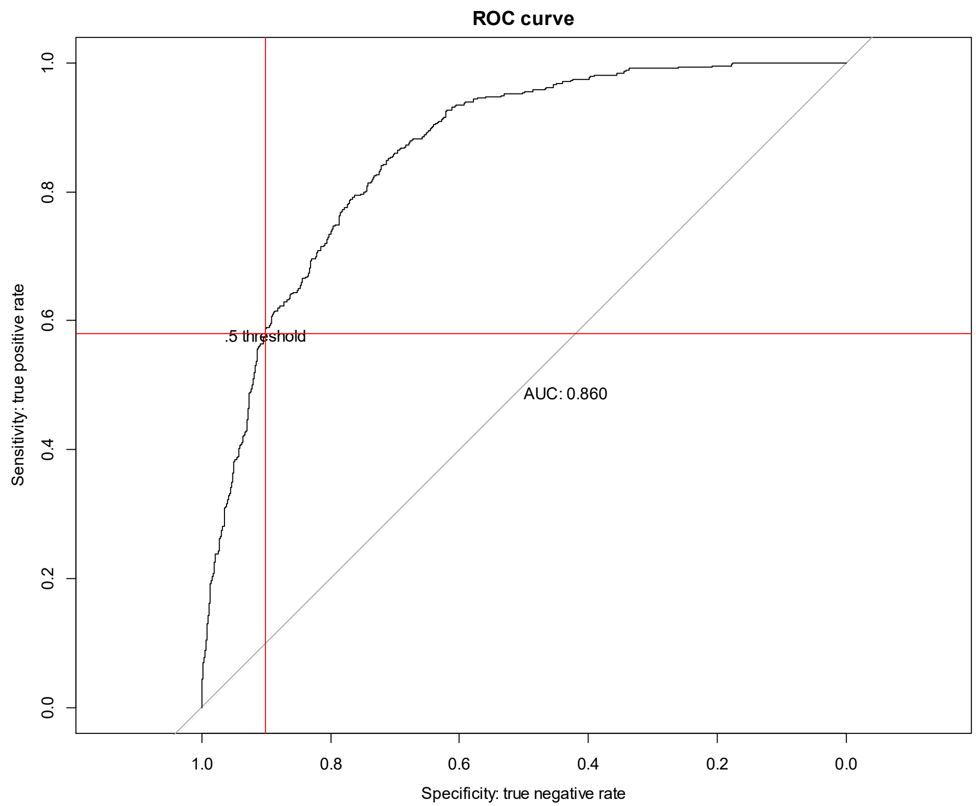
"Specificity" = True negative rate = TN / (TN + FP)

Classification matrix

		Predicted*		
		0	1	% Correct
Observed	0	*1167*	*119*	90.7
	1	*208*	*270*	56.5

Accuracy = 81.5%

ROC curve generalizes all cutoffs



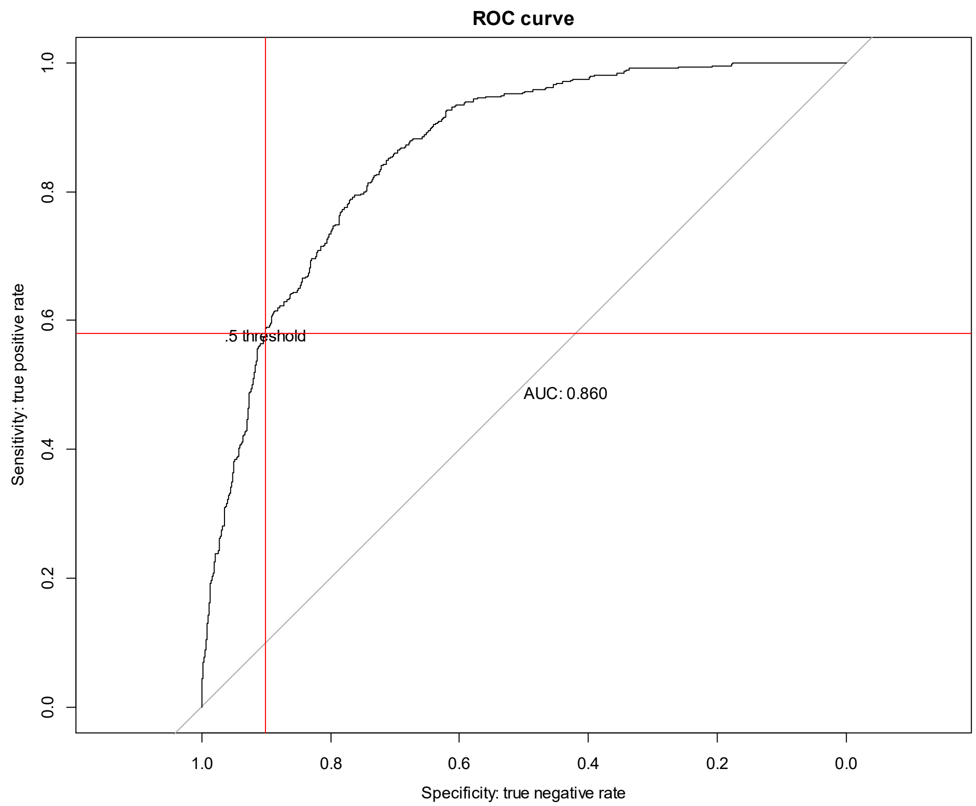
An ideal model will have a perfect true positive rate with no false positive rate

ROC curve will hug the top left corner

A bad model will make the same predictions for everyone, giving the diagonal line

Reading: [Blattberg, Kim and Neslin (2008) Ch 11]

Area under the curve (AUC)



AUC is a measure of how well the model distinguishes between two classes

One way of interpreting AUC is as the probability that the model ranks a random positive example more highly than a random negative example

Reading: Blattberg, Kim and Neslin (2008) Ch 11.

Lift

To get the "biggest bang for the buck", the marketer who can afford to target n persons picks the n highest-probability customers as targets.

Basic idea:

Based on predictions: pick top 10%, top 20%, … all 100%…

what % of churners will I get?

rank according to model, from most likely to least likely.

$$\lambda_1 = \frac{r_1}{\bar{r}}$$

Churn rate in top decile

Churn rate in entire sample (.27)

Reading: Blattberg, Kim and Neslin (2008) Ch 10.3.4

Lift

To get the "biggest bang for the buck", the marketer who can afford to target n persons picks the n highest-probability customers as targets.

Basic idea:

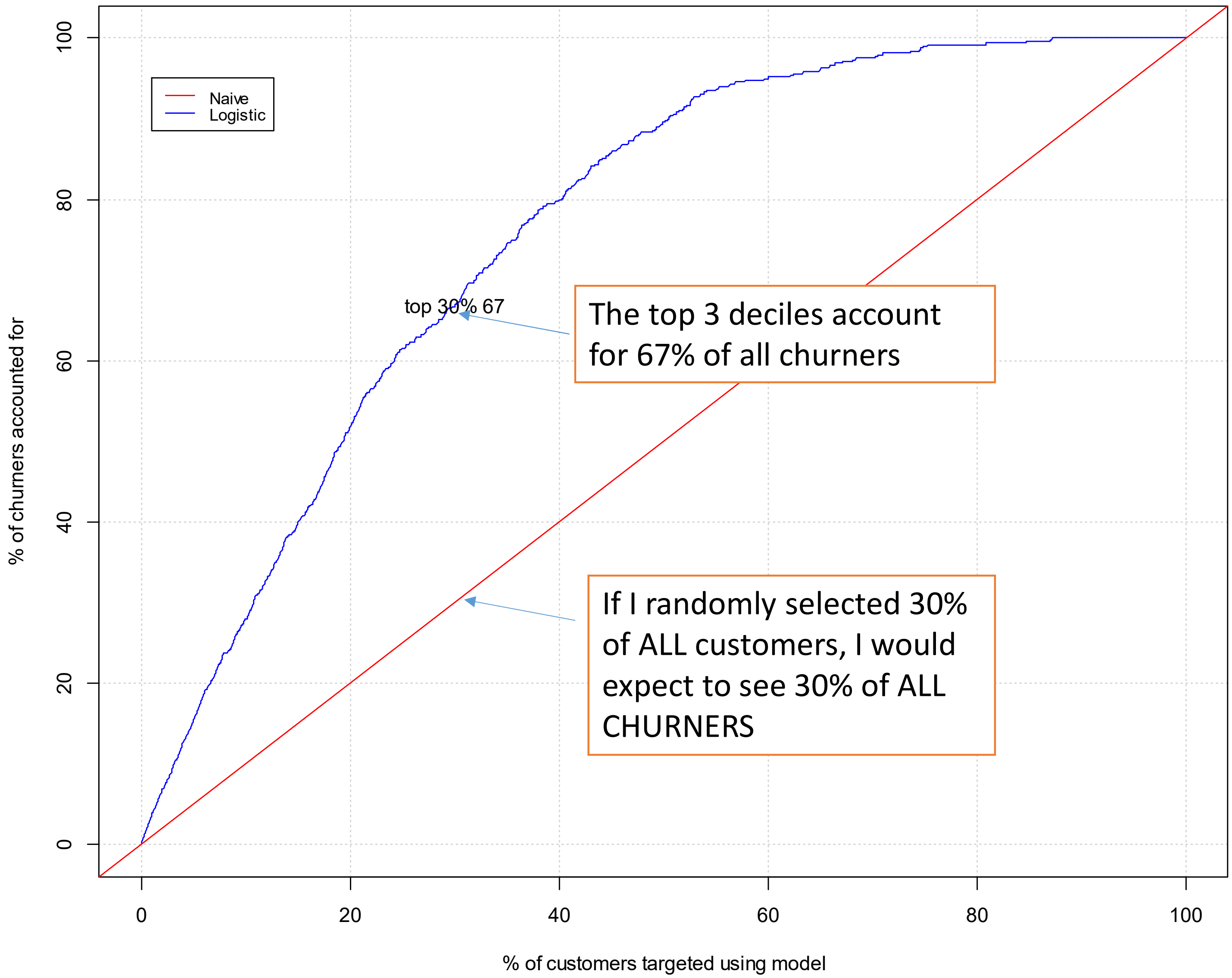
Based on predictions: pick top 10%, top 20%, … all 100%…

what % of churners will I get?

rank according to model, from most likely to least likely.

	decile	actual churn rate	lift	cumulative lift
Highest p	10	0.7684	2.8355	28.4
	9	0.6250	2.3065	51.5
	8	0.4350	1.6054	67.5
	7	0.3466	1.2791	80.3
	6	0.2500	0.9226	89.5
	5	0.1412	0.5212	94.8
	4	0.0795	0.2936	97.7
	3	0.0395	0.1459	99.2
Lowest p	2	0.0227	0.0839	100.0
	1	0.0000	0.0000	100.0

Reading: Blattberg, Kim and Neslin (2008) Ch 10.3.4

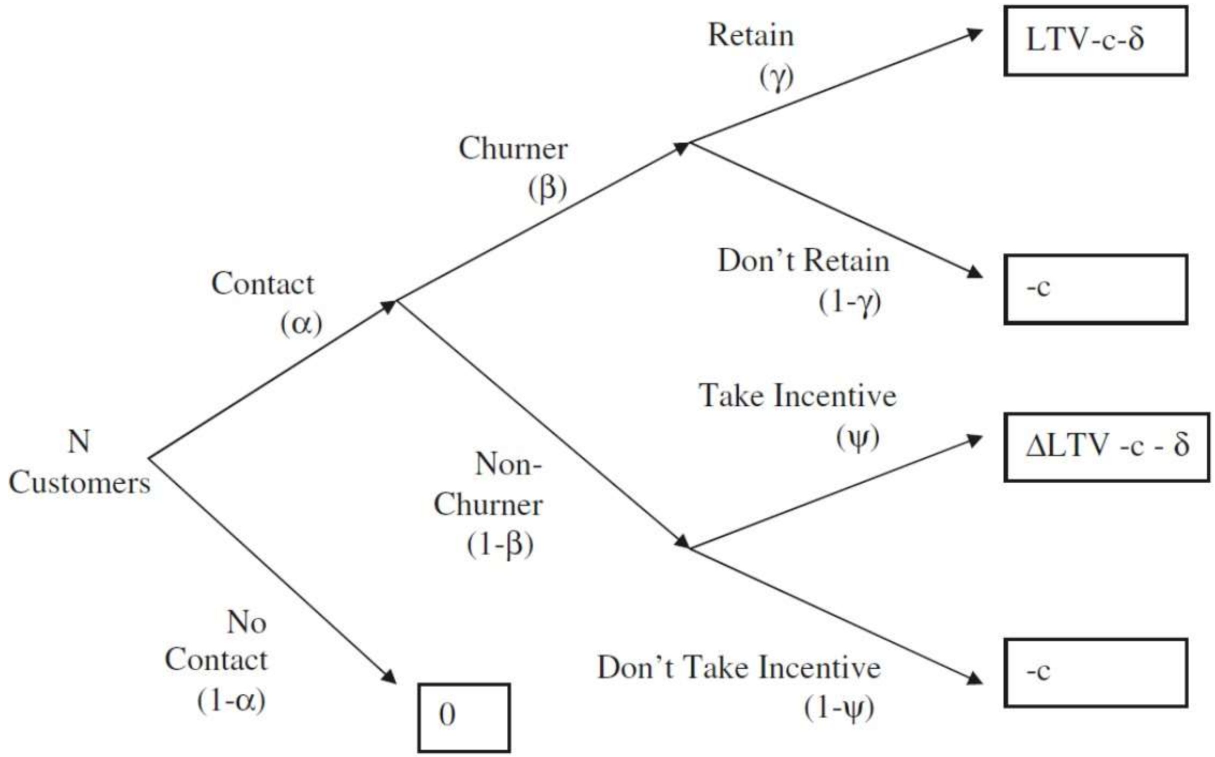


Naive
Logistic
top 30% 67
The top 3 deciles account for 67% of all churners
If I randomly selected 30% of ALL customers, I would expect to see 30% of ALL CHURNERS
% of churners accounted for
% of customers targeted using model
0
20
40
60
80
100

Logistic regression for proactive targeting

How many deciles should we target?

BKN 24.4.2



α proportion contacted
β proportion churners contacted
γ proportion rescued if a churner
c cost of contacting
δ cost of incentive to stay
LTV lifetime value of customer

ψ proportion non-churner takes incentive = 1
Δ percent increase in LTV of non-churners = 0

proportion

- If we contact the top K deciles, what is the proportion of actual churners contacted?

$$\beta_K = \frac{\sum_{k=1}^{K} r_k \, n_k}{\sum_{k=1}^{K} n_k}$$

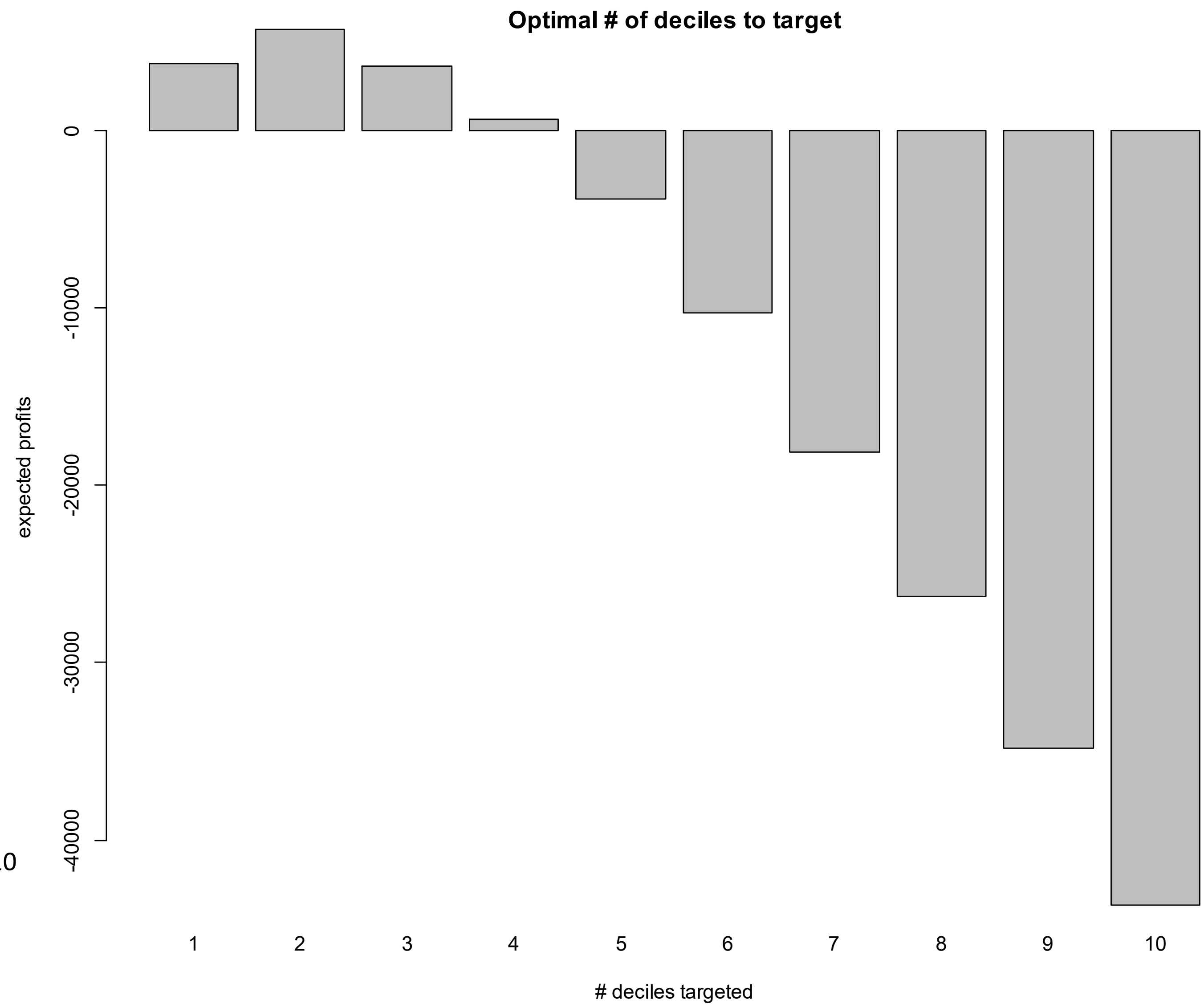
where r_k actual churn rate in decile k and n_k is the number of customers in decile k. These can be found in the lift table.

profits

$$\Pi = N\,\alpha\,\big(\beta\,\gamma\,(LTV\ \ - \delta - c) - \beta(1-\gamma)(c) - (1-\beta)(\delta + c)\big)$$

$$= N\,\alpha\,\big(\beta\,\big(\gamma\,LTV + \delta\,(1-\gamma)\big) - \delta - c\big)$$

choose $\alpha \in \{.1, .2, \ldots 1\}$ to maximize Π



γ proportion rescued if a churner = 0.10
c cost of contacting = 0.50
δ cost of incentive to stay = 50
LTV lifetime value of customer = 500