SAS® GLOBAL FORUM 2018

USERS PROGRAM

April 8 – 11 | Denver, CO
Colorado Convention Center

#SASGF
Prescriptive Analytics: Using Optimization with Predictive Models to find the Best Action

Dr. Mamdouh Refaat, Angoss Software (Datawatch)

Mamdouh Refaat is the SVP and Chief Data Scientist at Angoss Software where has been since 1999. He Leads the data science group in Angoss, a provider of data science software modeling tools.

He has used for more than 20 years and he authored two books on using SAS in data preparation for data mining, and for the development of credit risk scorecards.
Agenda

• About Optimization
• About Collections
• Collections Optimization
• A detailed example & a real case study
• Q&A
Why now?

What, why, how ...

Evolution of “Analytics”
Reports / Charts

What happened?

A Pie chart from William Playfair's "Statistical Breviary", 1801
Regression models, correlation analysis, ANOVA

How did/will it happen?

The original Galton’s diagram explaining the frequency distribution (~1880)
OLAP – Dynamic Reports

OLAP applications, Pivot Tables, Dynamic reporting

What happened (again!)?

Computers / Graphics → dynamic reporting / charting possible
Then and Now...The Analytics Journey

(Time) Reactive - Proactive - Actionable

Business Value

Diagnosis

- Predictive Modeling
- EDA/statistical analysis
- OLAP & ad-hoc queries

Insight
What Will Happen?

Foresight
What Should I Do?

Prescriptive

Optimization

Hindsight
What Happened & Why?

Alerts, Query, Reports

Predictive

Predictive Modeling, Forecasting, Statistical Analysis
Optimization – what is the best action?

Customer Portfolio

Letters  SMS  Calls

Models predicting likelihood of successful marketing

- Cost of each channel
- Channel capacity
- Minimum payment / loan value

Optimization

Optimal Channel per account
Optimization 101
Overview

Value

Make the best decisions under business constraints

Examples

• Marketing channel optimization
• Best of Collections Channel
• Optimization of Credit Offer – optimal price
• Best Product Offer for Acquisition
Constrained Optimization

Example:

Find $\min F(x_1, x_2, x_3)$
subject to

$\begin{align*}
7x_1 - 2x_2 + 3x_3 &\leq 7 \\
8x_1 + 9x_2 - 2.5x_3 &\leq 25 \\
x_1, x_2, x_3 &\geq 0
\end{align*}$

Objective Function

Decision variables

constraints

We try to find the decision variables $x_1, x_2, x_3$ that will minimize $F$ while satisfying the given constraints
A more fun Example!

Joy vs. consumption

Weight / health concerns vs. consumption
Happiness = Joy - Concerns

Objective function

Decision variable

Joy consumption

Weight (concern) consumption

Happiness consumption

*
Constraints

Happiness

Objective function

Feasible domain

Constraints

Decision variable

consumption
Profit = Revenue - Cost
Marketing Channel ➔ Maximum profit

Model 1: Response to a Telephone call $P_T$ $4.00$

Model 2: Response to an email $P_E$ $0.02$

Model 3: response to a letter $P_L$ $1.15$

$3,000/day$

$200,000/day$

$3,000/day$

$8,000$

<table>
<thead>
<tr>
<th>Cust. ID</th>
<th>$P_T$</th>
<th>$P_E$</th>
<th>$P_L$</th>
</tr>
</thead>
<tbody>
<tr>
<td>000-001</td>
<td>0.98</td>
<td>0.20</td>
<td>0.00</td>
</tr>
<tr>
<td>000-002</td>
<td>0.41</td>
<td>0.21</td>
<td>0.74</td>
</tr>
<tr>
<td>000-003</td>
<td>0.31</td>
<td>0.39</td>
<td>0.89</td>
</tr>
<tr>
<td>000-004</td>
<td>0.67</td>
<td>0.27</td>
<td>0.14</td>
</tr>
<tr>
<td>000-005</td>
<td>0.08</td>
<td>0.12</td>
<td>0.00</td>
</tr>
<tr>
<td>000-006</td>
<td>0.00</td>
<td>0.32</td>
<td>0.45</td>
</tr>
<tr>
<td>000-007</td>
<td>0.48</td>
<td>0.00</td>
<td>0.12</td>
</tr>
</tbody>
</table>

$3,000/day$

$200,000/day$

$3,000/day$

$8,000$
## Additional constraints

- Each customer is contacted by only **ONE CHANNEL**
- Some customers should not be contacted at all
- Any customer cannot be contacted more than **twice a quarter**

### Cust. ID | $P_T$ | $P_E$ | $P_L$
---|---|---|---
000-001 | 0.98 | 0.20 | 0.00
000-002 | 0.41 | 0.21 | 0.74
000-003 | 0.31 | 0.39 | 0.89
000-004 | 0.67 | 0.27 | 0.14
000-005 | 0.08 | 0.12 | 0.00
000-006 | 0.00 | 0.32 | 0.45
000-007 | 0.48 | 0.00 | 0.12

| Cust. ID | Marketing Channel - today |
---|---
000-001 | Telephone call |
000-002 | Letter |
000-003 | ---- (no contact) --- |
000-004 | ... |
000-005 | ... |
000-006 | ... |
000-007 | ... |
The Marketing Channel Optimization Problem

Model 1: Response to a Telephone call

Model 2: Response to an email

Model 3: Response to a letter

Max Profit

Constraints:
- Total budget
- Capacity of each channel

Revenue per unit
Cost per unit for each channel

Solution

Best Channel per customer

<table>
<thead>
<tr>
<th>Customer ID</th>
<th>Best Channel</th>
</tr>
</thead>
<tbody>
<tr>
<td>100001</td>
<td>Telephone</td>
</tr>
<tr>
<td>100002</td>
<td>Email</td>
</tr>
<tr>
<td>100003</td>
<td>Email</td>
</tr>
<tr>
<td>100004</td>
<td>Letter</td>
</tr>
<tr>
<td>...</td>
<td></td>
</tr>
</tbody>
</table>
Classification of Optimization Problems

and Solution Techniques!
Components of Optimization

Decision variables

Objective Function

Constraints

Continuous
Integer
Binary

Linear
Quadratic
Nonlinear

Linear
Nonlinear
Conditional

Programming ↔ Optimization!
Decision variables – which type?

- **Continuous values (real numbers)**
  - Balance, Profit, loss
  - Interest rate
  - Temperature
  - Volume / weight / times ...

- **Discrete (Integer) values (any non-divisible entity)**
  - Number of units
  - Number of products
  - Number of customers

- **Binary variables**
  - Actions (Call/Don’t call)
  - Mutually exclusive states (Good/Bad status)
Objective Function

- Price-demand is (linear, quadratic, other?)
- Production Cost - volume is (linear, piecewise linear, quadratic)
- Shipping cost –volume is (constant, piecewise constant, linear)
- Cost per calls (in a call center)- volume is (constant, piecewise constant, linear, other)
Creating Objective Functions

Conflicting objectives

Represents the business objective as a “Function” of the decision variables

- The units of the objective function are the **units** of the business objective
- We attempt to translate conflicting relations using a **single unit**
- **Money/time/counts** are the most common units of objective functions
Constraints

Reality of the business problem

Limitations of real life

- Resources are not unlimited
- Actions are not independent
  - If we call the customer, no need for SMS
  - We cannot call a customer more than a certain number of times every quarter
### Classification of Optimization Problems (1)

<table>
<thead>
<tr>
<th>Decision Variables</th>
<th>Objective Function</th>
<th>Constraints</th>
<th>Problem Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Continuous</td>
<td>Linear</td>
<td>Linear</td>
<td>Linear Programming (LP)</td>
</tr>
<tr>
<td>Continuous</td>
<td>Quadratic</td>
<td>Linear</td>
<td>Quadratic programming (QP)</td>
</tr>
<tr>
<td>Integer / binary</td>
<td>Linear</td>
<td>Linear</td>
<td>Integer programming (IP)</td>
</tr>
<tr>
<td>Mixed integer and continuous</td>
<td>Linear</td>
<td>Linear</td>
<td>Mixed Integer-linear programming</td>
</tr>
<tr>
<td>Continuous</td>
<td>Nonlinear – convex</td>
<td>Nonlinear – convex</td>
<td>Convex programming (CP)</td>
</tr>
<tr>
<td>Continuous</td>
<td>Nonlinear</td>
<td>Nonlinear</td>
<td>Nonlinear programming (NLP)</td>
</tr>
<tr>
<td>Mixed integer and continuous</td>
<td>Nonlinear</td>
<td>Nonlinear</td>
<td>General mixed nonlinear programming</td>
</tr>
</tbody>
</table>

*Programming ≡ Optimization*
Classification of Optimization Problems (2)

- Optimization
  - Stochastic Programming
    - Deterministic
    - Multi-objective
  - Unconstrained optimization
    - Least squares problems
    - Nonlinear equations
    - Nonlinear programming
    - Global Optimization
  - Continuous
  - Discrete
    - Mixed Integer programming
    - Linear programming
    - Quadratic programming
    - Combinatorial Problems
    - Integer programming
Optimization Methodology
Realism vs. simplicity

Approximation of a cow's geometry

The rectangular cow model

The idealized spherical cow
Real World ↔ Mathematical Model

Complexity

Real world problem

Approximation/
Ease of solution

Mathematical Model

Right balance?

Approximate description of reality
Implications – Accuracy, validation, reporting, ...

Real world problem

Mathematical Model (Optimization)

Apply (Score)

Predictions / expected results

Reports!

Actual results

Real world data
The Solution of Optimization Problems

Find: decision variables
\[ x_1, x_2, \ldots \]

Min (max) \[ F(x_1, x_2, \ldots) \]

Subject to
\[ g_1(x_1, x_2, \ldots) \leq 0 \]
\[ g_2(x_1, x_2, \ldots) \leq 0 \]
\[ \vdots \]

• The solution
  1. Existence of a solution (exists, bounded, feasible)
  2. Value of decision variables \[ x_1, x_2, \ldots \]
  3. Value of objective function at optimal solution
  4. State of constraints (free or active)

Can't we always find a solution? !!
No Optimal Value: Minimum = Maximum!

Saddle Point
How can a real problem have a saddle point?

\[ x = \text{Balance} \]
\[ y = \text{Interest rate} \]

\[ \text{Risk} = (x-1000)^2 - (y-1.2)^2 \]

Saddle Point at
\[ x = 1000 \]
\[ y = 1.2 \]
Infeasible Problems

Infeasible problem

Max \( F = 9 x_1 \)

Subj \( x_1 > 10, \)
\( x_1 < 0, \)
How can a real problem be infeasible?

Maximize

Cows ≥ 1

Not Feasible
Unbounded Objective Function

Unbounded problem
Max \( F = x_1 + x_2 \)
How can a real problem be unbounded?

Maximize

Cows ≥1

Unbounded!!
Source of these problems is ...
The Solution of Optimization Problems

Find: decision variables $x_1, x_2, \ldots$

Min(max) $F(x_1, x_2, \ldots)$

Subject to

$g_1 (x_1, x_2, \ldots) \leq 0$
$g_2 (x_1, x_2, \ldots) \leq 0$

...$

• The solution
  1. Existence of a solution (exists, bounded, feasible)
  2. Value of decision variables $x_1, x_2, \ldots$
  3. Value of objective function at optimal solution
  4. State of constraints (free or active)
Lessons after the solution – Sensitivity Analysis

- Value of decision variables $x_1, x_2, ...$
- Value of objective function at optimal solution
- State of constraints (free or active)

Values and trends

$F(x_1, x_2, ...)$

$g_1(x_1, x_2, ...) \leq 0$
$g_2(x_1, x_2, ...) \leq 0$

...
Sensitivity analysis – Objective Function

Effect of change in the objective function
Sensitivity analysis - constraints

Effect of change in the Constraints
More on Sensitivity Analysis

The optimization problem may turn into an unconstrained problem!
How to solve optimization problems

How to **NOT** solve optimization problems!
## Optimal Channel Assignment

**Solution 1**

<table>
<thead>
<tr>
<th>ID</th>
<th>Ch</th>
<th>ID</th>
<th>Ch</th>
<th>ID</th>
<th>Ch</th>
<th>ID</th>
<th>Ch</th>
<th>ID</th>
<th>Ch</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>E</td>
<td>1</td>
<td>C</td>
<td>1</td>
<td>E</td>
<td>1</td>
<td>E</td>
<td>1</td>
<td>E</td>
</tr>
<tr>
<td>2</td>
<td>E</td>
<td>2</td>
<td>E</td>
<td>2</td>
<td>C</td>
<td>2</td>
<td>E</td>
<td>2</td>
<td>E</td>
</tr>
<tr>
<td>3</td>
<td>E</td>
<td>3</td>
<td>E</td>
<td>3</td>
<td>C</td>
<td>3</td>
<td>E</td>
<td>3</td>
<td>E</td>
</tr>
<tr>
<td>4</td>
<td>E</td>
<td>4</td>
<td>E</td>
<td>4</td>
<td>E</td>
<td>4</td>
<td>C</td>
<td>4</td>
<td>E</td>
</tr>
<tr>
<td>5</td>
<td>E</td>
<td>5</td>
<td>E</td>
<td>5</td>
<td>E</td>
<td>5</td>
<td>E</td>
<td>5</td>
<td>C</td>
</tr>
</tbody>
</table>

**Solution 2**

<table>
<thead>
<tr>
<th>ID</th>
<th>Ch</th>
<th>ID</th>
<th>Ch</th>
<th>ID</th>
<th>Ch</th>
<th>ID</th>
<th>Ch</th>
<th>ID</th>
<th>Ch</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>E</td>
<td>1</td>
<td>E</td>
<td>1</td>
<td>E</td>
<td>1</td>
<td>E</td>
<td>1</td>
<td>C</td>
</tr>
<tr>
<td>2</td>
<td>E</td>
<td>2</td>
<td>E</td>
<td>2</td>
<td>E</td>
<td>2</td>
<td>E</td>
<td>2</td>
<td>C</td>
</tr>
<tr>
<td>3</td>
<td>E</td>
<td>3</td>
<td>E</td>
<td>3</td>
<td>E</td>
<td>3</td>
<td>E</td>
<td>3</td>
<td>C</td>
</tr>
<tr>
<td>4</td>
<td>E</td>
<td>4</td>
<td>E</td>
<td>4</td>
<td>E</td>
<td>4</td>
<td>C</td>
<td>4</td>
<td>E</td>
</tr>
<tr>
<td>5</td>
<td>C</td>
<td>5</td>
<td>E</td>
<td>5</td>
<td>C</td>
<td>5</td>
<td>C</td>
<td>5</td>
<td>C</td>
</tr>
</tbody>
</table>

**Solution 3**

<table>
<thead>
<tr>
<th>ID</th>
<th>Ch</th>
<th>ID</th>
<th>Ch</th>
<th>ID</th>
<th>Ch</th>
<th>ID</th>
<th>Ch</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>E</td>
<td>1</td>
<td>E</td>
<td>1</td>
<td>E</td>
<td>1</td>
<td>E</td>
</tr>
<tr>
<td>2</td>
<td>E</td>
<td>2</td>
<td>E</td>
<td>2</td>
<td>E</td>
<td>2</td>
<td>E</td>
</tr>
<tr>
<td>3</td>
<td>E</td>
<td>3</td>
<td>E</td>
<td>3</td>
<td>E</td>
<td>3</td>
<td>E</td>
</tr>
<tr>
<td>4</td>
<td>E</td>
<td>4</td>
<td>E</td>
<td>4</td>
<td>E</td>
<td>4</td>
<td>C</td>
</tr>
<tr>
<td>5</td>
<td>C</td>
<td>5</td>
<td>C</td>
<td>5</td>
<td>C</td>
<td>5</td>
<td>C</td>
</tr>
</tbody>
</table>

**Solution 4**

<table>
<thead>
<tr>
<th>ID</th>
<th>Ch</th>
<th>ID</th>
<th>Ch</th>
<th>ID</th>
<th>Ch</th>
<th>ID</th>
<th>Ch</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>E</td>
<td>1</td>
<td>E</td>
<td>1</td>
<td>E</td>
<td>1</td>
<td>E</td>
</tr>
<tr>
<td>2</td>
<td>E</td>
<td>2</td>
<td>E</td>
<td>2</td>
<td>E</td>
<td>2</td>
<td>E</td>
</tr>
<tr>
<td>3</td>
<td>E</td>
<td>3</td>
<td>E</td>
<td>3</td>
<td>E</td>
<td>3</td>
<td>E</td>
</tr>
<tr>
<td>4</td>
<td>E</td>
<td>4</td>
<td>E</td>
<td>4</td>
<td>E</td>
<td>4</td>
<td>C</td>
</tr>
<tr>
<td>5</td>
<td>E</td>
<td>5</td>
<td>E</td>
<td>5</td>
<td>E</td>
<td>5</td>
<td>E</td>
</tr>
</tbody>
</table>

**Solution 5**

<table>
<thead>
<tr>
<th>ID</th>
<th>Ch</th>
<th>ID</th>
<th>Ch</th>
<th>ID</th>
<th>Ch</th>
<th>ID</th>
<th>Ch</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>E</td>
<td>1</td>
<td>E</td>
<td>1</td>
<td>E</td>
<td>1</td>
<td>E</td>
</tr>
<tr>
<td>2</td>
<td>E</td>
<td>2</td>
<td>E</td>
<td>2</td>
<td>E</td>
<td>2</td>
<td>E</td>
</tr>
<tr>
<td>3</td>
<td>E</td>
<td>3</td>
<td>E</td>
<td>3</td>
<td>E</td>
<td>3</td>
<td>E</td>
</tr>
<tr>
<td>4</td>
<td>E</td>
<td>4</td>
<td>E</td>
<td>4</td>
<td>E</td>
<td>4</td>
<td>C</td>
</tr>
<tr>
<td>5</td>
<td>E</td>
<td>5</td>
<td>E</td>
<td>5</td>
<td>E</td>
<td>5</td>
<td>E</td>
</tr>
</tbody>
</table>

**Solution 32**

<table>
<thead>
<tr>
<th>ID</th>
<th>Ch</th>
<th>ID</th>
<th>Ch</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>C</td>
<td>1</td>
<td>C</td>
</tr>
<tr>
<td>2</td>
<td>C</td>
<td>2</td>
<td>C</td>
</tr>
<tr>
<td>3</td>
<td>C</td>
<td>3</td>
<td>C</td>
</tr>
<tr>
<td>4</td>
<td>C</td>
<td>4</td>
<td>C</td>
</tr>
<tr>
<td>5</td>
<td>C</td>
<td>5</td>
<td>C</td>
</tr>
</tbody>
</table>

**Number of possible solutions**

\[ \text{Number of possible solutions} = 2 \times 2 \times 2 \times 2 \times 2 = (2)^5 = (\text{channels})^{\text{records}} \]
A Realistic Small Collections Portfolio

~ 200,000 accounts

Number of solutions \((2)^{200,000} = (2 \times 2 \times 2 \times 2)^{50,000} = (16)^{50,000} > (10)^{50,000}\)

This is a **VERY LARGE NUMBER** of solutions to test

How large is \(10^{50,000}\) ?
Atoms are everywhere!

How big is an atom?
Atoms are really, really **SMALL**

The atoms in a tablespoon of salt would create a chain that is **3.5 times** the circumference of the planet Earth.
Putting things in perspective!

5 x \((10)^{23}\) Atoms in a tablespoon of water

\((10)^{82}\) Atoms in the entire Universe

Our **Small** Collections problem has more than \((10)^{50,000}\) Solutions

That is \((10)^{49,918}\) times more than the number of atoms in the universe
So, what is the lesson?

- Combinatorial problems are HARD

- We don’t solve them by trying EVERY possible solution

- **We need special algorithms ➔ Software tools**
Few things about Collections

(difficulties!)
How? (Channels)

• Not all channels have the same effectiveness
• Not all respond same way to contacts using the different channels
• There are regulatory restrictions on writing off debts
• Collections budgets are limited
• There exist a large number of *accepted wisdom* in the form of “Collections Strategies”
• Collections strategies focus on sequence and timing of actions (30 DPD: Letter → 45 DPD: Call → ...)

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration. Other brand and product names are trademarks of their respective companies.
Definition of Successful Collection

• Spoke / Connected / Engaged with account holder
• Promised to pay
• Agreed on a payment plan
• Made a partial payment
• Made a full payment

When to remove the account from the collections queue?
Which action caused (triggered) payment?
# Vantage Analysis

## Current Period (month)

<table>
<thead>
<tr>
<th>Previous Period (month)</th>
<th>30-60 DPD</th>
<th>60-90 DPD</th>
<th>90-120 DPD</th>
<th>120-180 DPD</th>
<th>&gt; 180 DPD</th>
</tr>
</thead>
<tbody>
<tr>
<td>30-60 DPD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60-90 DPD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>90-120 DPD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>120-180 DPD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt; 180 DPD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Bad

- Arrows indicate a decrease in DPD across periods.

### Good

- Arrows indicate an increase in DPD across periods.
Models ... Models ...

- Likelihood to Self-cure
- Likelihood to pay back (any amount, %, all)
- Estimate of account to recover
- Likelihood to Write-off
- Likelihood to accept restructuring
- Likelihood to promise to pay
- Likelihood to connect/engage with the right person
- Likelihood to “respond” using a specific channel

Which model?
Channel ROI vs DPD

30-60 DPD
Letter, Email, Call, SMS

60-90 DPD
Letter, Email, Call, SMS

90-120 DPD
Letter, Email, Call, SMS

120-180 DPD
Letter, Call, SMS
Collections Optimization

Finally !!!
The framework

Collections Optimization

Account Portfolio

Data Preparation

Raw data
Collections results

Optimal “Actions”

Optimization

Constraints

Models predicting likelihood of successful collections using different Actions
Optimization problems consist of

1. Decision variables
2. Objective function
3. Constraints
Decision Variables

- Optimal channel → Channel(s) (1/0)
- When to contact → DPD / time of day (1/2/3)
- Message intensity → Level (1/2/3/...)
- Best action → collect, write-off, restructure
- Who Collects → internal, outsource (a/b)
- Agent assignment → which agent (1, ..., 86)

Contact strategy → Channel + when + intensity + who
Problem Setup

Complexity

Real world problem

Mathematical Model

Approximation/
Ease of solution

Right balance?

Approximate description of reality

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration. Other brand and product names are trademarks of their respective companies.
From Simple → Complex!

Result Acceptable?

Yes

No

Result Acceptable?

Yes

No

Deploy
A Detailed Example
Collections Channel Optimization

Accounts Data

- Cost of each channel / action
- Channel capacity
- One channel only

Likelihood of successful collections using channel

 Calls

 SMS

 Optimization

 Optimal channel per Account
# A Two-channel Collections Problem

<table>
<thead>
<tr>
<th>ID</th>
<th>Response Probability</th>
<th>Actions</th>
<th>Revenue (R)</th>
<th>Cost per action</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Call (α)</td>
<td>SMS (β)</td>
<td>Call (x)</td>
<td>SMS (y)</td>
</tr>
<tr>
<td>1</td>
<td>α1</td>
<td>β1</td>
<td>x1</td>
<td>y1</td>
</tr>
<tr>
<td>2</td>
<td>α2</td>
<td>β2</td>
<td>x2</td>
<td>y2</td>
</tr>
<tr>
<td>3</td>
<td>α3</td>
<td>β3</td>
<td>x3</td>
<td>y3</td>
</tr>
<tr>
<td>4</td>
<td>α4</td>
<td>β4</td>
<td>x4</td>
<td>y4</td>
</tr>
<tr>
<td>5</td>
<td>α5</td>
<td>β5</td>
<td>x5</td>
<td>y5</td>
</tr>
</tbody>
</table>

**Decision variables**

**Objective function**

**Constraints**
### Step 1 – Decision Variables

<table>
<thead>
<tr>
<th>ID</th>
<th>Response Probability</th>
<th>Actions</th>
<th>Revenue (R)</th>
<th>Cost per action</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Call (α)  SMS (β)</td>
<td>Call (x)</td>
<td>SMS (y)</td>
<td>Call (c)  SMS(s)</td>
</tr>
<tr>
<td>1</td>
<td>α1      β1</td>
<td>x1</td>
<td>y1</td>
<td>R1</td>
</tr>
<tr>
<td>2</td>
<td>α2      β2</td>
<td>x2</td>
<td>y2</td>
<td>R2</td>
</tr>
<tr>
<td>3</td>
<td>α3      β3</td>
<td>x3</td>
<td>y3</td>
<td>R3</td>
</tr>
<tr>
<td>4</td>
<td>α4      β4</td>
<td>x4</td>
<td>y4</td>
<td>R4</td>
</tr>
<tr>
<td>5</td>
<td>α5      β5</td>
<td>x5</td>
<td>y5</td>
<td>R5</td>
</tr>
</tbody>
</table>

**Decision variables**
- x1, x2, ..., y1, y2, ... are binary (0/1)
- x1=0 → No Call
- x1=1 → Call
- ...
- Y1=0 → No SMS
- Y2=1 → SMS

**Channels to optimize**

---

© 2019, SAS Institute Inc. All rights reserved.
Step 2 – Objective Function: Profit

<table>
<thead>
<tr>
<th>ID</th>
<th>Response Probability</th>
<th>Actions</th>
<th>Revenue (R)</th>
<th>Cost per action</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Call (α)_SMS (β)</td>
<td></td>
<td>Call (x)_SMS (y)</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>α1_β1</td>
<td>x1</td>
<td>y1</td>
<td>R1_c_s</td>
</tr>
<tr>
<td>2</td>
<td>α2_β2</td>
<td>x2</td>
<td>y2</td>
<td>R2_c_s</td>
</tr>
<tr>
<td>3</td>
<td>α3_β3</td>
<td>x3</td>
<td>y3</td>
<td>R3_c_s</td>
</tr>
<tr>
<td>4</td>
<td>α4_β4</td>
<td>x4</td>
<td>y4</td>
<td>R4_c_s</td>
</tr>
<tr>
<td>5</td>
<td>α5_β5</td>
<td>x5</td>
<td>y5</td>
<td>R5_c_s</td>
</tr>
</tbody>
</table>

Account 1
Expected Revenue = max(α1 * x1 + β1 * y1) * R1
Constraint: x1 + y1 = 1
Cost = x1 * c + y1 * s

All Accounts
Profit = Σ{(αi * xi + βi * yi) * Ri - (xi * c + yi * s)}
## Step 3 – Constraints

<table>
<thead>
<tr>
<th>ID</th>
<th>Response Probability</th>
<th>Actions</th>
<th>Revenue (R)</th>
<th>Cost per action</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Call (a)</td>
<td>SMS (β)</td>
<td>Call (x)</td>
<td>SMS (y)</td>
</tr>
<tr>
<td>1</td>
<td>α1</td>
<td>β1</td>
<td>x1</td>
<td>y1</td>
</tr>
<tr>
<td>2</td>
<td>α2</td>
<td>β2</td>
<td>x2</td>
<td>y2</td>
</tr>
<tr>
<td>3</td>
<td>α3</td>
<td>β3</td>
<td>x3</td>
<td>y3</td>
</tr>
<tr>
<td>4</td>
<td>α4</td>
<td>β4</td>
<td>x4</td>
<td>y4</td>
</tr>
<tr>
<td>5</td>
<td>α5</td>
<td>β5</td>
<td>x5</td>
<td>y5</td>
</tr>
</tbody>
</table>

### Account level constraint:
Use only one channel

\[
x_1 + y_1 = 1 \\
x_2 + y_2 = 1 \\
x_3 + y_3 = 1 \\
x_4 + y_4 = 1 \\
x_5 + y_5 = 1
\]

### Channel level constraint:
(Maximum Budget)

\[
x_1 + x_2 + x_3 + x_4 + x_5 \leq 3 \\
y_1 + y_2 + y_3 + y_4 + y_5 \leq 3
\]
Formal Problem Statement (Integer Programming)

Maximize
\[
\{ (\alpha_1 * x_1 + \beta_1 * y_1) * R_1 - (x_1 * c + y_1 * s) \\
+ (\alpha_2 * x_2 + \beta_2 * y_2) * R_2 - (x_2 * c + y_2 * s) \\
+ (\alpha_3 * x_3 + \beta_3 * y_3) * R_3 - (x_3 * c + y_3 * s) \\
+ (\alpha_4 * x_4 + \beta_4 * y_4) * R_4 - (x_4 * c + y_4 * s) \\
+ (\alpha_5 * x_5 + \beta_5 * y_5) * R_5 - (x_5 * c + y_5 * s) \}
\]

Subject to
\[
\begin{align*}
x_1 + y_1 &= 1; & x_1 + x_2 + x_3 + x_4 + x_5 &\leq 3; \\
x_2 + y_2 &= 1; \\
x_3 + y_3 &= 1; & y_1 + y_2 + y_3 + y_4 + y_5 &\leq 3; \\
x_4 + y_4 &= 1; \\
x_5 + y_5 &= 1; \\
\end{align*}
\]
\[
x_1, x_2, x_3, x_4, x_5, y_1, y_2, y_3, y_4, y_5 \text{ Integer}
\]
# Numerical Example

<table>
<thead>
<tr>
<th>ID</th>
<th>Response Probability</th>
<th>Revenue (R)</th>
<th>Cost per action</th>
<th>Actions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Call (α)</td>
<td>SMS (β)</td>
<td>Call (c)</td>
<td>SMS(s)</td>
</tr>
<tr>
<td>1</td>
<td>0.42</td>
<td>0.42</td>
<td>80</td>
<td>$1.00</td>
</tr>
<tr>
<td>2</td>
<td>0.40</td>
<td>0.72</td>
<td>200</td>
<td>$1.00</td>
</tr>
<tr>
<td>3</td>
<td>0.66</td>
<td>0.82</td>
<td>80</td>
<td>$1.00</td>
</tr>
<tr>
<td>4</td>
<td>0.35</td>
<td>0.41</td>
<td>150</td>
<td>$1.00</td>
</tr>
<tr>
<td>5</td>
<td>0.10</td>
<td>0.20</td>
<td>120</td>
<td>$1.00</td>
</tr>
</tbody>
</table>

Calls ≤ 3; SMS ≤ 3;
Observations

• We may not use all the assigned budget
• Some accounts will not be contacted
• Accounts may be assigned the channel with the least likelihood to respond!
• To take account of previous contacts, use them as variables in the models
• We optimize “day-by-day” $\Leftarrow \Rightarrow$ not the optimum value over a period of time
• Did not take account of write-offs
Actual Project – A North American Canadian Bank

Client data sources

Raw data

Collections results

Angoss

Models predicting likelihood of successful collections using different channels

Letters
SMS
Robot calls
Live calls
Agency

Optimal channel(s) per account

Optimization

• Cost of channel
• Channel capacity
• Minimum Payment
Results – Lessons Learned

• Higher dollars collected
• Higher channel ROI

• More attention to higher balances
• Issues with mixed portfolios (more focus on some at the expense of others)

• Revising the models to consider write-offs and effect on capital reserve
Final Words of Wisdom!

• Optimization is fun
• Use it if it makes sense
• Pay more attention to the problem set up than anything else
• Use a good software
• Don’t attempt to write your own algorithm before trying to use available tools (don’t re-invent the wheel)
Thank You all ...

mrefaat@Angoss.com
Your Feedback Counts!

Don't forget to complete the session survey in your conference mobile app.

1. Go to the Agenda icon in the conference app.
2. Find this session title and select it.
3. On the Sessions page, scroll down to Surveys and select the name of the survey.
4. Complete the survey and click Finish.