Institute of Actuaries of India

Pricing in Health Insurance including Multi- Variate Analysis

Capacity Building Seminar in Health Insurance

Presentation by Raunak Jha

3rd September 2012

Mumbai

TOWERS WATSON

Contents

Market Overview

Market Overview

- Health insurance currently contributes nearly 24 % of the total GI industry turnover.
- It's been the fastest growing class of business. The growth rate for FY 11-12 over 10-11 was 20 %.
- Health insurance was a highly unprofitable business till FY 2007-08.The business now appears to be headed for profitability, going by latest trends. Underwriting experience for FY 10-11 was 92%.
- There has been very limited development from a "scope of coverage" perspective. Current offerings are basically multiple variants of an indemnity based Mediclaim hospitalization product.
- On the Retail side, major changes taken place are inclusion of coverage for Day-care procedures, Sum Insured Restoration option.
- On the Group side, many extras were thrown, e.g. maternity, no waiting period, no pre-existing exclusion etc. But on the other hand, to control losses, many counter measures were introduced (and rightly so) like sub-limits, co-pay etc.
- Industry still has a long way to cover a wider portion of health value chain covering aspects like primary care, wellness and preventive healthcare..

Current Pricing Practices

Current Practice

Operational Companies having own Data:

- Data analysis often limited by quality and quantity of available data
- Univariate analyses, at aggregate level, mostly considering rate-making for different segments

Companies lacking data (New Entrants)

- Pricing follows market range, with some margins for uncertainty of future risk experience
- Periodic reviews in pricing based on one's own emerging experience.

Overview of Univariate Techniques followed

TOWERS WATSON

Overview of Univariate Techniques followed

Basic Premium risk relationship equation:

Premium = Claim Cost (including LAE) + Expenses + Margins

The two basic approaches followed (which realigns the above equation):

Pure Premium Approach:

This approach determines a computed average rate per exposure given by formula:

$$\text{Indicated Average Rate} = \frac{(\text{Pure risk cost + Fixed Expense) Per Exp}}{(1.0 - \text{Variable Expense \%} - \text{Target Profit \%})}$$

Loss ratio Method:

This method compares sum of the projected claims ratio and projected fixed expense ratio to the variable permissible loss ratio

$$\text{Indicated Change Factor} = \frac{\text{Claims Ratio+ Fixed Expense Ratio}}{(1\text{-Variable Expense \%} - \text{Target UW Profit \%})} - 1$$

Segmental Ratemaking

Segmental Ratemaking

This is the process of categorizing and grouping risks with similar loss potential and charging different rates to reflect differences in loss potential among the various groups.

It involves:

- Determining the risk parameters that are effective for segmenting risks into different groups, each with similar expected loss experience.
- Subdividing insured population into appropriate levels
- Calculating the indicated *rate differential relative to the base level* for each level being priced.

Segmental Ratemaking :

Basic Statistics:

Exposure (No. of Lives Insured) Distribution (In Lacs)				
	Territory			
Sum Insured	1	2	3	Total
Low	7	130	143	280
Medium	108	126	126	360
High	179	129	40	348
Total	294	385	309	988

Base level

Sum Insured	True Relativities
Low	0.7300
Medium	1.0000
High	1.4300

Territory	True Relativities
1	0.6312
2	1.0000
3	1.2365

Classification Ratemaking using Pure Premium Approach (Learning by Examples)

The Indicated Differential for a given level is equal to the projected ultimate pure (risk) premium for that level divided by the projected ultimate pure premium for the base class.

Territory	Exposure	Loss *	Indicated Pure Premium	Indicated Relativity	Indicated Relativity to Base
1	294	15,235	51.82	0.7877	0.7526
2	385	26,510	68.86	1.0466	1.0000
3	309	23,255	75.26	1.1439	1.0930
Total	988	65,000	65.79	1.0000	0.9555

* **Including LAE**

Result Comparison		
Territory	**True Relativities**	**Pure Premium Relativities**
1	0.6312	0.7526
2	1.0000	1.0000
3	1.2365	1.0930

Classification Ratemaking using Loss Ratio Approach (Learning by Examples)

Territory	Premium at Current Rate Level	Loss *	Loss Ratio	Indicated Relativity Change Factor	Current Relativities	Indicated Relativity	Indicated Relativity to Base
1	21,315	15,235	71%	1.0996	0.6000	0.6598	0.6538
2	40,414	26,510	66%	1.0092	1.0000	1.0092	1.0000
3	38,271	23,255	61%	0.9348	1.3000	1.2153	1.2043
Total	100,000	65,000	65%	1.0000			

Indicated Relativity is Indicated Relativity Change Factor* Current Relativity

Result Comparison			
Territory	True Relativities	Pure Premium Indication	Loss Ratio Indication
1	0.6312	0.7526	0.6538
2	1.0000	1.0000	1.0000
3	1.2365	1.0930	1.2043

Minimum Bias Approach

TOWERS WATSON

Classification Ratemaking using Minimum Bias Approach (Learning by Examples)

This example assumes only two rating variables: gender and territory. Gender has values male (with a rate relativity expressed as g1) and female (g2). Territory has values urban (t1) and rural (t2). The base levels, relative to which all multiplicative indications will be expressed, are female and rural (hence g2 = 1.00 and t2 = 1.00). The actual loss costs (or pure premiums) are as follows:

The Actual Loss Costs (In Lacs):			
Gender	**Urban**	**Rural**	**Total**
Male	**650**	**300**	**950**
Female	**250**	**240**	**490**
Total	**900**	**540**	**1440**

Exposure Distribution (In Thousands):			
Gender	**Urban**	**Rural**	**Total**
Male	**170**	**90**	**260**
Female	**105**	**110**	**215**
Total	**275**	**200**	**475**

Base Rate = 100

	Actual Experience		Expected Experience
MALE :	170*650+90*300	=	100*170*g1*t1+100*90*g1*t2
FEMALE :	105*250+110*240	=	100*105*g2*t1+100*110*g2*t2
URBAN :	170*650+105*250	=	100*170*g1*t1+100*90*g2*t1
RURAL :	90*300+110*230	=	100*90*g1*t2+100*110*g2*t2

First Seed is usually univariate pure premium relativity:

t1 = 900/540 = 1.667   t2 = 900/900 = 1.000

Classification Ratemaking using Minimum Bias Approach (Learning by Examples)

On substitution:

MALES : 170*650+90*300=100*170*g1*1.667+100*90*g1*1.000

137500 =28333*g1+9000*g1

g1= 3.683

FEMALE : 105*250+110*240=100*105*g2*t1+100*110*g2*t2

52650 =10500*g2+11000*g2

g2= 2.449

URBAN : 170*650+105*250=100*170*3.683*t1+100*90*2.449*t1

136750 =62612*t1+22040*t1

t1= 1.615

RURAL : 90*300+110*240=100*90*3.683*t2+100*110*2.449*t2

53400 =25713*t2+14539*t2

t2= 1.327

Base Loss Cost =100*2.449*1.327

Classification Ratemaking using Minimum Bias Approach (Learning by Examples)

g1= 3.698/2.449=1.504

g2= 2.449/2.449=1.000

t1= 1.615/1.327=1.218

t2= 1.327/1.327=1.000

Several iterations of univariate analysis are performed on rating variables, each time adjusting for the exposure weight and the indication of the previous variable in the sequence.

The minimum bias procedures are not technically multivariate methods, and they were not necessarily based directly on statistical theory.

Many of the minimum bias procedures are actually a subset of the statistical method, generalized linear models (GLMs).

Iterating the minimum bias procedure a sufficient number of times may result in convergence with GLM results.

Multivariate Methods and their benefits

Benefits of Multivariate Methods

- Consider all rating variables simultaneously and automatically adjusts for exposure correlations between rating variables
- Remove unsystematic effects in the data (also known as noise) and capture only the systematic effects (also known as signal) as much as possible
- Produce model diagnostics, additional information about the certainty of results and the appropriateness of the model fitted
- Allows consideration of the interaction, or interdependency, between two or more rating variables.
- Examples of Multivariate Classification Methods:
  - Linear Models
  - Generalised Linear Models
  - Cluster Analysis
  - Classification and Regression Trees (CART)
  - Multivariate Adaptive Regression Spline(MARS)
  - Neural Networks

Multivariate Methods: Generalised Linear Models(GLMs)

GLM is a multivariate method and considers all factors simultaneously

SIGNAL

NOISE

y = h(linear function of rating factors) + error

- Signal is measured by a function of the rating factors
- $g = h^{-1}$ is called the LINK function

- Error chosen should reflect the underlying process
- Error can come from the exponential family of distributions
  - e.g. Normal, Poisson, Gamma, Binomial

- $y = h(X\beta) + \text{error}$
- The link function is chosen to measure the “signal” most accurately – signal is determined from the parameters relating to the rating factors
- An appropriate error assumption ensures the correct signal is extracted even in sparse datasets

Mutivariate Methods: GLMs (Model Examples)

- The link function is chosen to measure the “signal” most accurately and is applied to parameters relating to the rating factors
- The error structure is intended to reflect the variability of the underlying process

Observed Response	Most Appropriate Error Structure	Most Appropriate Link Function
Claim Frequency	Poisson	Log
Average Cost or Severity	Gamma	Log
Risk Premium	Gamma or user-defined	Log
Renewal Rate	Binomial	Logit
Conversion or Hit Rate	Binomial	Logit

Comparison of techniques discussed so far

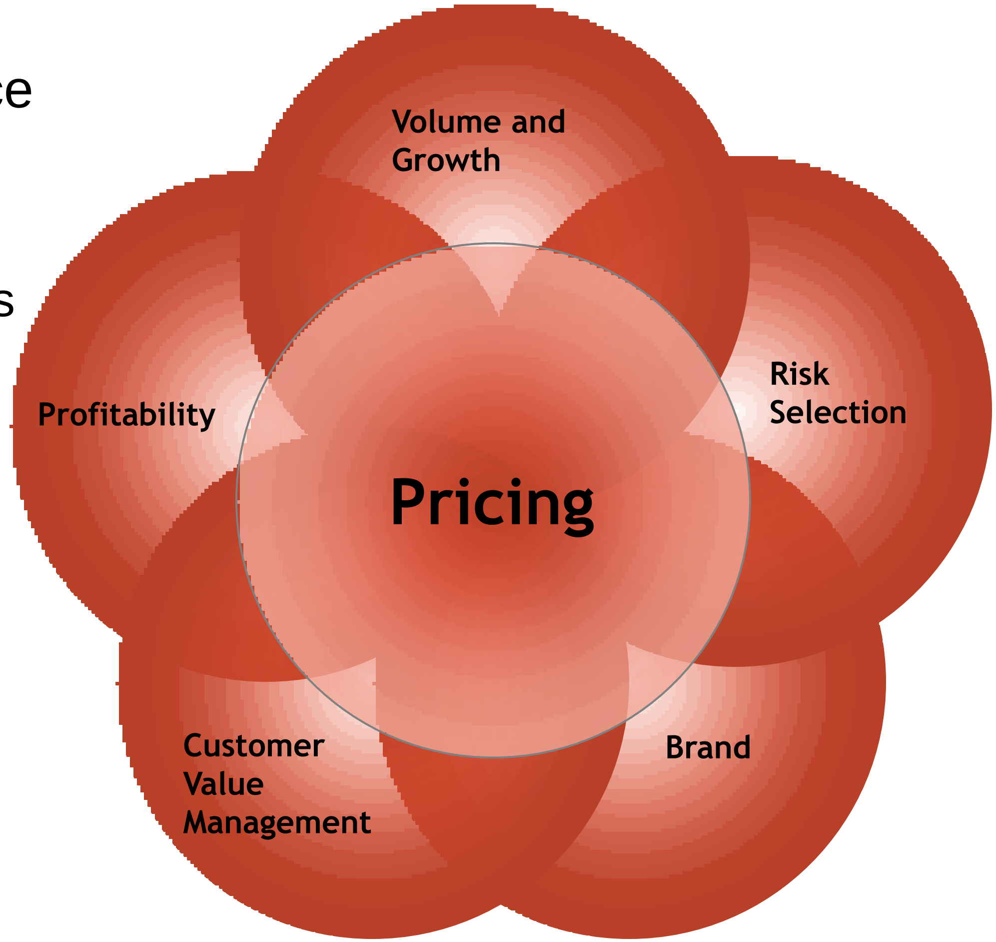
- Univariate Techniques
  - Limits use of explanatory variables which are continuous in nature
  - Ignores Correlation between variables(this may double count variable effect)
  - Ignores interdependencies or interactions between variables(reduces predictive power)
- Minimum Bias Procedures
  - Doesn’t capture the influence of variable on results
  - Lacks statistical framework to access quality of modelling work
  - No credible range for parameters estimate
- Generalised Linear Models
  - Allows for interactions and correlations between variables
  - Statistical framework allows explicit assumptions to be made
  - The method of solving GLMs is technically more efficient
  - Provides statistical diagnostics which aid in selecting only significant factors and validating model assumptions

Better Pricing makes an Impact

TOWERS WATSON

Better pricing makes a big impact

- Better pricing helps an insurance company in each and every aspect
  - Maintaining the Profitability levels
  - Customer Value Management
  - Brand value
  - Appropriate Risk Selection
  - Increasing Volume and Growth



Thank you

Reference Papers:

- A Practitioner's Guide to Generalized Linear Models(Authored by Duncan Anderson)
- A Systematic Relationship between Minimum Bias and Generalised Linear Models(Authored by Stephen Mildenhall)