Unbiasied Estimation of the Economic Value of Pricing Strategies

D. Semenovich, J. Petterson

Abstract: The key assumption made in most real-world pricing systems is that the demand and cost functions are known exactly for each customer. While this makes the "optimal pricing" problem tractable, it also introduces substantial statistical difficulties. We demonstrate that under realistic assumptions on model uncertainty, the economic value estimates for "optimal" pricing strategies obtained via traditional methods can be overstated to a remarkable degree. We propose a new method for unbiased estimation of value of arbitrary pricing strategies, inspired by recent progress in the reinforcement learning.

Introduction

Tactical pricing of insurance products can often be effectively carried out adopting the so called "semi-myopic" customer model. Under this model customers have private willingness-to-pay, drawn from a distribution potentially dependent on their observed characteristics, and are taken to arrive at random. If customers' willingness-to-pay exceeds the proposed premium, they purchase the policy. The customers are not "strategic" in that they do not attempt to anticipate future price changes — customer arrival is taken to be independent of any pricing decisions.

A key assumption made in most real-world pricing systems is that the willingness-to-pay distributions (or, equivalently, demand functions), as well as the cost of providing cover, are known for each customer. While this makes the problem more tractable and permits to describe solutions qualitatively, it also introduces substantial statistical difficulties as we will show in this paper.

The prevalent approach (Murphy et al., 2000; Krikler et al., 2004) follows along these lines:

1. specify a demand model and a cost of cover model,
2. estimate their parameters using sales, exposure and claims cost data,
3. set up an optimal pricing problem using the above two models, with individual contract prices as decision variables,
4. use the objective/constraints values corresponding to the solution as estimates of the economic value of the resulting pricing strategy to the firm.

We demonstrate that the above framework is only adequate if both demand and risk cost model estimates have minimal prediction uncertainty. Once realistic assumptions are adopted, however, the economic value of resulting pricing strategies is overstated to a considerable degree (inflation by factors between 1.2 and 5 is not uncommon).

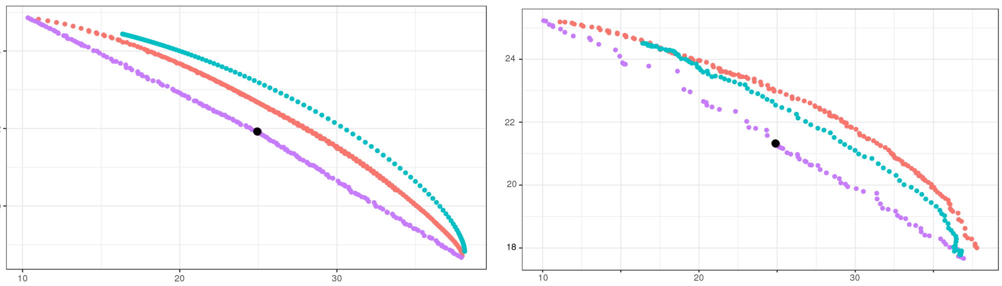

Figure 1: *Left:* Expected effects on conversions (x-axis) and margin (y-axis) as a result to $\pm 10\%$ per quote premium change, evaluated using a demand model. Black dot denotes the current portfolio position. Purple frontier represents operating points achievable with a base rate change. Red frontier indicates the effect of moving premiums towards a target loss ratio. Green frontier is a biased estimate derived using the traditional optimisation procedure. *Right:* Same premium changes as in the previous plot but evaluated using the proposed unbiased estimator. Note that the order of the frontiers is *reversed*, the simpler profitability based adjustment is now expected to outperform the "optimisation".

Traditional tests for goodness of fit, predictive accuracy and calibration used to validate risk cost and demand models are ultimately neither necessary nor sufficient to ensure correct estimation of the economic value. We propose a new family of unbiased evaluation metrics for pricing procedures, inspired by work in uplift modeling and reinforcement learning.

We compare results obtained using the traditional procedure and our proposed method in a real-world pricing scenario for a motor portfolio in Figure 1. Reversal in the order of frontiers suggests that the standard optimisation methods can result in strategies that underperform simple baselines in practice.

The sociological considerations that have allowed the current practice to become widely adopted despite its obvious shortcomings are left without comment.

Single period optimal pricing problem

We begin by reviewing a simple single period optimal pricing model. The aim is to maximise the total profit objective for a cohort of n policies subject to a constraint on the minimum retention level q, where for the i-th policy with risk characteristics $\mathbf{x}_i$ the proposed premium is denoted p_i, the demand is a random variable $D_i(p_i)$ indexed by premium, C_i is a random variable corresponding to the cost of claims and $R_i(p_i) = (p_i - C_i)D(p_i)$ a random variable corresponding to realised underwriting profit.

We can further assume a parametric form for the expectations with $\mathbb{E}[D_i(p_i)] = d(p_i, \mathbf{x}_i)$[1], $\mathbb{E}[C_i] = c(\mathbf{x}_i)$ and $\mathbb{E}[R_i(p_i)] = r(p_i, \mathbf{x}_i)$, all taken to be known exactly. If C_i and D_i are independent, this yields:

$$\begin{aligned} \underset{p_1,\dots,p_n}{\text{maximise}} \quad & \sum_{i=1}^{n} (p_i - c(\mathbf{x}_i))d(p_i, \mathbf{x}_i) = \sum_{i=1}^{n} r(p_i, \mathbf{x}_i) = r(\mathbf{p}) \\ \text{subject to} \quad & \sum_{i=1}^{n} d_i(p_i, \mathbf{x}_i) = q. \end{aligned} \tag{1}$$

Here the decision variables are premiums $p_i \geq 0$. We will refer to the solution of this problem as $\mathbf{p}^*$, with optimal underwriting profit given by $r(\mathbf{p})$.

[1] As we are dealing with demand levels for individual policies, d_i can also be interpreted as a probability.

In practice, we do not have access to the parametrised expectations of demand and cost random variables and instead we are working with their estimates $\hat{d}(p_i, \mathbf{x}_i)$ and $\hat{c}(\mathbf{x}_i)$ respectively. It is common practice to still use the optimisation problem of the same form as (1):

$$\begin{aligned} \underset{p_1,\dots,p_n}{\text{maximise}} \quad & \sum_{i=1}^{n} (p_i - \hat{c}(\mathbf{x}_i))\hat{d}(p_i, \mathbf{x}_i) = \sum_{i=1}^{n} \hat{r}(p_i, \mathbf{x}_i) = \hat{r}(\mathbf{p}) \\ \text{subject to} \quad & \sum_{i=1}^{n} \hat{d}_i(p_i, \mathbf{x}_i) = q. \end{aligned} \tag{2}$$

The solution to this surrogate problem is denoted as $\hat{\mathbf{p}}^*$ and the objective value as $\hat{r}(\hat{\mathbf{p}}^*)$. We will later show that under realistic assumptions on model error it obtains that:

$$r(\hat{\mathbf{p}}^*) < r(\mathbf{p}^*) < \hat{r}(\hat{\mathbf{p}}^*), \tag{3}$$

suggesting that the approximate optimisation procedure (2) not only leads to deterioration in performance relative to theoretically obtainable, but also provides a substantially biased estimate of the attainable value of the objective.

Before examining the properties of the naive estimate of the objective value $\hat{r}(\hat{\mathbf{p}}^*)$, however, we further observe that the problem (1) can be rewritten using policy *demand* as the decision variable, assuming one-to-one correspondence between premium and demand $p(d_i, \mathbf{x}_i) = d^{-1}(d_i, \mathbf{x}_i)$:

$$\begin{aligned} \underset{d_1,\dots,d_n}{\text{maximise}} \quad & \sum_{i=1}^{n} (p(d_i, \mathbf{x}_i) - c(\mathbf{x}_i))d_i = r(\mathbf{d}) \\ \text{subject to} \quad & \sum_{i=1}^{n} d_i \;= q, \end{aligned} \tag{4}$$

We can then formulate the Lagrangian:

$$L(d_1, \dots, d_n, \lambda) = \sum_{i=1}^{n} (p(d_i, \mathbf{x}_i) - c(\mathbf{x}_i))d_i + \lambda\Big(\sum_{i=1}^{n} d_i - q\Big)$$

and write the optimality conditions[2] as:

$$\begin{aligned} \frac{\partial L}{\partial d_i} &= 0, \quad 1 \le i \le n, \\ \frac{\partial L}{\partial \lambda} &= 0. \end{aligned}$$

Note that:

$$\frac{\partial L}{\partial d_i} = \frac{\partial r}{\partial d_i} + \lambda$$

and therefore if the portfolio is priced optimally, marginal profit with respect to demand for each policy is constant:

$$\frac{\partial r}{\partial d_i} = -\lambda. \tag{5}$$

This condition is intuitive — should $\frac{\partial r}{\partial d_i} \neq \frac{\partial r}{\partial d_j}$ for some i and j, we can reallocate de-

[2] The solution does not need to be unique in general, however, for monotone demand functions from certain parametric families e.g. logistic and probit, the optimisation problem is convex which would mean that the solution is unique or solutions form a convex set.

mand between contracts i and j in such a way as to increase total profit.

Effects of model uncertainty

We now demonstate that the surrogate optimisation problem (2) is subject to a facet of the phenomenon that often causes overparametrised statistical models to "overfit" in sample.

The effect of model uncertainty can be studied more easily if instead of (2) we consider a *local linearisation* (i.e. first order Taylor expansion) of the demand parametrised problem (4) around demand vector $\mathbf{d}^{(0)}$ instead:

$$\begin{aligned}
\underset{w_1,\dots,w_n}{\text{maximise}} \quad & \sum_{i=1}^{n} \Big(r(d_i^{(0)}, \mathbf{x}_i) + \frac{\partial r}{\partial d_i} w_i \Big) = r(\mathbf{w}) \\
\text{subject to} \quad & \sum_{i=1}^{n} (d_i^{(0)} + w_i) = q, \\
& -1 \le w_i \le 1.
\end{aligned} \tag{6}$$

Omitting the constant term $r(\mathbf{d}_o)$ from the objective and observing that $\sum_{i=1}^{n} d_i^{(0)} = q$, we can simplify the above as:

$$\begin{aligned}
\underset{w_1,\dots,w_n}{\text{maximise}} \quad & \sum_{i=1}^{n} \frac{\partial r}{\partial d_i} w_i = r(\mathbf{w}) \\
\text{subject to} \quad & \sum_{i=1}^{n} w_i = 0, \\
& -1 \le w_i \le 1.
\end{aligned} \tag{7}$$

It is intuitive that the solution $\mathbf{w}^*$ is attained if we set $w_i^* = 1$ for those policies i where $\frac{\partial r}{\partial d_i}$ is larger than M, the median entry of $(\frac{\partial r}{\partial d_1}, \dots, \frac{\partial r}{\partial d_n})$, and $w_i^* = -1$ where it is smaller. The objective value corresponsing to $\mathbf{w}^*$ is then given by $\sum_{i=1}^{n} |\frac{\partial r}{\partial d_i} - M|$. It represents improvement to profit r attainable by perturbing demand by no more than one unit for each contract relative to the initial demand vector $\mathbf{d}^{(0)}$.

Notice that if we substitute a noisy estimate of marginal profit $\hat{\frac{\partial r}{\partial d}} = \frac{\partial r}{\partial d} + \epsilon$, our view of expected profit improvements can generally only go up. This means that any model uncertainty will result in statistically biased estimates of expected profit.

Now we attempt to quantify this bias. This will require further assumptions:

$$\begin{aligned}
\epsilon &\sim \mathcal{N}(0, \sigma_a), \\
\frac{\partial R}{\partial d} &\sim \mathcal{N}(0, \sigma_b).
\end{aligned}$$

What is the degradation in true performance as we increase the noise σ_a?

For conciseness we will refer to ϵ as a, $\frac{\partial R}{\partial d}$ as b and the corresponding probability density functions as $p_{\sigma_a}(a)$ and $p_{\sigma_b}(b)$ respectively.

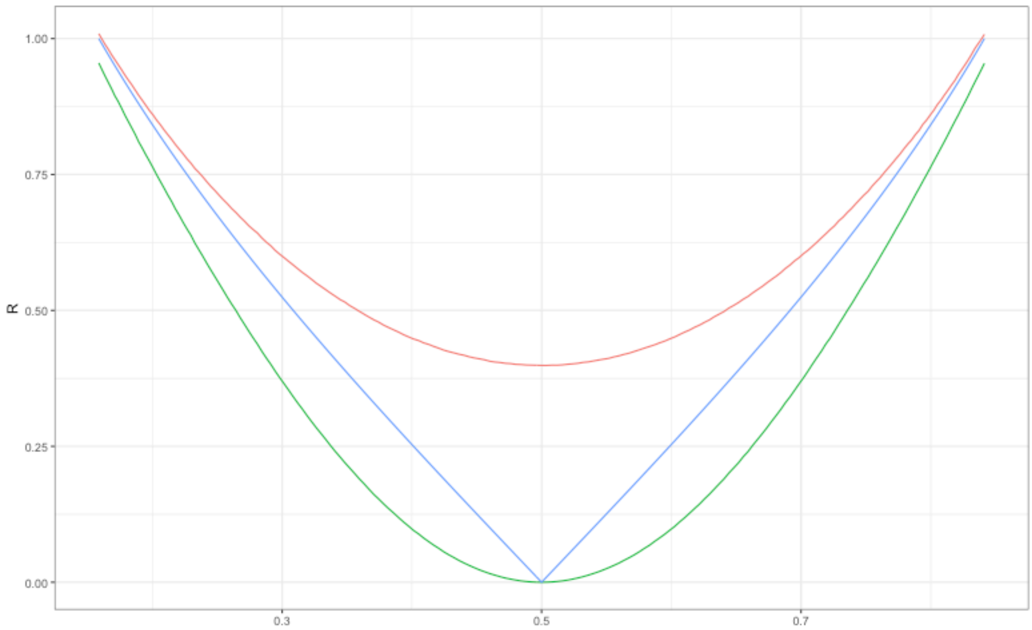

Figure 2: A numerical example showing the bias inherent in the traditional "optimal" pricing procedures. The x axis corresponds to the quantiles of the true marginal profit of a policy and the y axis to the profit either achieved or estimated. The area under the blue line represents the total profit improvement realisable if the true marginal profit with respect to demand is known. The area under the green line shows the profit attained if the noisy estimate of marginal profit is used to guide pricing decisions. Finally the area under the red line is the *biased estimate* of profit that would be achieved. The gap between red and green lines corresponds total bias in traditional optimal pricing.

True Estimator, True Metric

Decision and measure are based on the true marginal profit $\frac{\partial R}{\partial d}$. Note that w^* here is a step function over true marginal profit taking values of $\{-1, 1\}$, as characterised in the previous section.

$$\begin{aligned} \mathbb{E}R(w^*) &= \int_{-\infty}^{\infty} p_{\sigma_b}(b)\,\mathrm{sign}(b) b\, db \\ &= -\int_{-\infty}^{0} p_{\sigma_b}(b) b\, db + \int_{0}^{\infty} p(b) b\, db \\ &= \frac{2\sigma_b}{\sqrt{2\pi}}. \end{aligned} \tag{8}$$

Noisy Estimator, True Metric

Decision is based on a noisy estimator $\frac{\partial R}{\partial d} + \epsilon$, but we measure the profit using the true metric ($\frac{\partial R}{\partial d}$). Here $\hat{w}^*$ is a step function over taking values of $\{-1, 1\}$ over the noisy values of marginal profit $\frac{\partial R}{\partial d} + \epsilon$.

$$
\begin{aligned}
\mathbb{E}R(\hat{w}^*) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p_{\sigma_a}(a)p_{\sigma_b}(b)\,\mathrm{sign}(a+b)b\,db\,da \\
&= \int_{-\infty}^{\infty} p_{\sigma_a}(a)\left(\int_{-\infty}^{\infty} p_{\sigma_b}(b)\,\mathrm{sign}(a+b)b\,db\right)\,da \\
&= -A_1 + A_2,
\end{aligned}
$$

where

$$
\begin{aligned}
A_1 &= \int_{-\infty}^{\infty} p_{\sigma_a}(a)\left(\int_{-\infty}^{-a} p_{\sigma_b}(b)b\,db\right)\,da \\
&= -\frac{\sigma_b}{\sqrt{2\pi}}\int_{-\infty}^{\infty} p_{\sigma_a}(a)e^{-\frac{a^2}{2{\sigma_b}^2}}\,da \\
&= -\frac{1}{\sqrt{2\pi}}\frac{{\sigma_b}^2}{\sqrt{{\sigma_a}^2+{\sigma_b}^2}}
\end{aligned}
$$

$$
\begin{aligned}
A_2 &= \int_{-\infty}^{\infty} p_{\sigma_a}(a)\left(\int_{-a}^{\infty} p_{\sigma_b}(b)b\,db\right)\,da \\
&= \frac{\sigma_b}{\sqrt{2\pi}}\int_{-\infty}^{\infty} p_{\sigma_a}(a)e^{-\frac{a}{2{\sigma_b}^2}}\,da \\
&= \frac{1}{\sqrt{2\pi}}\frac{{\sigma_b}^2}{\sqrt{{\sigma_a}^2+{\sigma_b}^2}}
\end{aligned}
$$

and so

$$
\mathbb{E}R(\hat{w}^*) = \frac{2}{\sqrt{2\pi}}\frac{{\sigma_b}^2}{\sqrt{{\sigma_a}^2+{\sigma_b}^2}}. \tag{9}
$$

Noisy Estimator, Noisy Metric

Decision and profit estimates are both based on the noisy estimator $\frac{\partial R}{\partial d} + \epsilon$.

$$
\begin{aligned}
\mathbb{E}\hat{R}(\hat{w}^*) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p_{\sigma_a}(a)p_{\sigma_b}(b)\,\mathrm{sign}(a+b)(a+b)\,db\,da \\
&= \int_{-\infty}^{\infty} p_{\sigma_a}(a)\left(\int_{-\infty}^{\infty} p_{\sigma_b}(b)\,\mathrm{sign}(a+b)(a+b)\,db\right)\,da \\
&= -B_1 + B_2,
\end{aligned}
$$

where

$$
\begin{aligned}
B_1 &= \int_{-\infty}^{\infty} p_{\sigma_a}(a) \left(\int_{-\infty}^{-a} p_{\sigma_b}(b)(a+b)\, db \right) da \\
&= \frac{1}{2} \int_{-\infty}^{\infty} p_{\sigma_a}(a) a \,\mathrm{erfc}\left(\frac{a}{\sqrt{2{\sigma_b}^2}} \right) da - \frac{\sigma_b}{\sqrt{2\pi}} \int_{-\infty}^{\infty} p_{\sigma_a}(a) e^{-\frac{a^2}{2{\sigma_b}^2}} da \\
&= -\frac{1}{\sqrt{2\pi}} \frac{{\sigma_a}^2}{\sqrt{{\sigma_a}^2 + {\sigma_a}^2}} - \frac{1}{\sqrt{2\pi}} \frac{\sigma_b^2}{\sqrt{{\sigma_a}^2 + {\sigma_b}^2}} \\
&= -\frac{\sqrt{{\sigma_a}^2 + {\sigma_a}^2}}{\sqrt{2\pi}}
\end{aligned}
$$

$$
\begin{aligned}
B_2 &= \int_{-\infty}^{\infty} p_{\sigma_a}(a) \left(\int_{-a}^{\infty} p_{\sigma_b}(b)(a+b)\, db \right) da \\
&= \int_{-\infty}^{\infty} p_{\sigma_a}(a) \left(\int_{-a}^{\infty} p_{\sigma_b}(b) a\, db \right) da + \int_{-\infty}^{\infty} p_{\sigma_a}(a) \left(\int_{-a}^{\infty} p_{\sigma_b}(b) b\, db \right) da \\
&= \frac{1}{2} \int_{-\infty}^{\infty} p_{\sigma_a}(a) a \,\mathrm{erf}\left(\frac{a}{\sqrt{2{\sigma_b}^2}} + 1 \right) + \frac{\sigma_b}{\sqrt{2\pi}} \int_{-\infty}^{\infty} p_{\sigma_a}(a) e^{-\frac{a}{2{\sigma_b}^2}} da \\
&= \frac{\sqrt{{\sigma_a}^2 + {\sigma_a}^2}}{\sqrt{2\pi}}
\end{aligned}
$$

and so

$$
\mathbb{E}\hat{R}(\hat{w}^*) = \frac{2}{\sqrt{2\pi}} \sqrt{{\sigma_a}^2 + {\sigma_b}^2}. \tag{10}
$$

This provides the following decomposition:

$$
\mathbb{E}\hat{R}(\hat{w}^*) = \frac{2}{\sqrt{2\pi}} \frac{\overbrace{{\sigma_a}^2}^{\text{noise on metric estimation}} + {\sigma_b}^2}{\sqrt{\underbrace{{\sigma_a}^2}_{\text{noise on decision}} + {\sigma_b}^2}}. \tag{11}
$$

We observe that when $\sigma_a = 0$ we recover (8), adding noise to the decision criterion reduces the expected value of profit R and adding noise to the evaluation metric increases it, yielding:

$$
\mathbb{E}R(\hat{w}^*) \leq \mathbb{E}R(w^*) \leq \mathbb{E}\hat{R}(\hat{w}^*).
$$

Unbiased estimation

We can construct an unbiased estimator of expected profit if we conduct validation "out of sample".

Assume we have history of sales and claims data in the form $S = \{(\mathbf{x}_i, p_i, d_i, c_i, \psi_i)\}_{i=1}^N$, where ψ_i is the propensity estimate of charging premium p_i for risk $\mathbf{x}_i$. In the ideal scenario these propensities are based on active randomisation with known probabilities.

This history has not been used directly to parametrise either demand or claims cost models (and so we can assume individual realisations to be independent of prediction error). Once we obtain a vector $\hat{\mathbf{p}}$ of proposed prices for each policy using a procedure such as (2), an unbiased estimate of profit is given by the so called *inverse probability weighted estimator* (Horvitz and Thompson, 1952; Dudik et al., 2014):

$$\hat{r}_{\mathrm{IPW}}(\hat{\mathbf{p}}^*) = \frac{1}{N}\sum_{i=1}^{N}(p_i - c_i)d_i\frac{\mathbb{I}(p_i = \hat{p}_i^*)}{\psi_i}.$$

The IPW estimate can be somewhat noisy on small samples. The variance is magnified by the ratio of at least $\frac{1}{\underset{i}{\operatorname{argmax}}\ \psi_i}$:

$$\operatorname{Var}[\hat{r}_{\mathrm{IPW}}(\hat{\mathbf{p}}^*)] \quad = \quad \frac{1}{N}\sum_{i=1}^{N}\left((p_i - c_i)d_i\frac{\mathbb{I}(p_i = \hat{p}_i^*)}{\psi_i}\right)^2 - \hat{r}_{\mathrm{IPW}}(\hat{\mathbf{p}}^*)^2.$$

To be used successfully, it is essential that the randomisation of p_i is carried out over the same small set of values in relation to some reference price p_i^0 as that used in the optimisation procedure to derive $\hat{p}_i^*$. In some cases it may also be necessary to substitute c_i with model based value $\hat{c}(\mathbf{x}_i)$.

We note that repalcing $\mathbb{I}(p_i = \hat{p}_i^*)$ with a kernel $\kappa(p_i, \hat{p}_i^*)$ satisfying certain properties may substantially reduce this variance while the resulting estimator remains unbiased under only mild assumptions. This will be explored in future work.

Using IPWE for the evaluation of pricing decisions is conceptually equivalent to out of sample testing of predictive models.

Conclusion

We have highlighted a substantial statistical issue with the standard approaches to optimal pricing of insurance contracts and have propose an alternative method of evaluation. Several questions that have not been addressed in this note are:

- extending IPS to deal with continuous action space (fine grained price adjustments);
- finding optimal pricing strategies in the reinforcement learning framework.

References

Dudik M., Erhan D., Langford J., and Li, L., Doubly Robust Policy Evaluation and Optimization, *Statistical Science*, 2014, 29, pp. 485– 511.

Horvitz, D. G., Thompson, D. J., A generalization of sampling without replacement from a finite universe, *Journal of the American Statistical Association*, 1952, 47, pp. 663–685.

Krikler S., Dolberger D., and Eckel J., Method and tools for insurance price and revenue optimisation, *Journal of Financial Services Marketing*, 2004, 9, pp. 68–79.

Murphy K. P., Brockman M. J., and Lee P. K. W., Using generalized linear models to build dynamic pricing systems for personal lines insurance, *Casualty Actuarial Society Forum*, Winter 2000, pp. 107–140.