Term Information

Effective Term	Spring 2023

General Information

Course Bulletin Listing/Subject Area	Mathematics
Fiscal Unit/Academic Org	Mathematics - D0671
College/Academic Group	Arts and Sciences
Level/Career	Graduate, Undergraduate
Course Number/Catalog	5588
Course Title	Practicum in Actuarial and Quantitative Risk Management
Transcript Abbreviation	Pract Quant Risk
Course Description	This course introduces students to the real world practice relative to advanced Actuarial Science topics and more specifically to the MAQRM program.
Semester Credit Hours/Units	Fixed: 3

Offering Information

Length Of Course	14 Week, 7 Week, 6 Week
Flexibly Scheduled Course	Never
Does any section of this course have a distance education component?	No
Grading Basis	Letter Grade
Repeatable	No
Course Components	Lecture
Grade Roster Component	Lecture
Credit Available by Exam	No
Admission Condition Course	No
Off Campus	Never
Campus of Offering	Columbus

Prerequisites and Exclusions

Prerequisites/Corequisites	2568, 3618, and 4530 or Stats 4201; or permission of department.
Exclusions	
Electronically Enforced	Yes

Cross-Listings

Cross-Listings

Subject/CIP Code

Subject/CIP Code	27.0101
Subsidy Level	Doctoral Course
Intended Rank	Junior, Senior, Masters

Requirement/Elective Designation

The course is an elective (for this or other units) or is a service course for other units

Course Details

Course goals or learning objectives/outcomes

- To introduce students to the real world practice and to improve their communication skills.

Content Topic List

- Option Pricing and Simulation
- Market Optimization and Elasticity Modeling
- Brownian Motion and Its Application in Quantitative Finance
- Pension and Retirement
- Data Analysis and Modeling

Sought Concurrence

No

Attachments

- 5588.pdf: Syllabus - second update

(Syllabus. Owner: Husen,William J)

Comments

- ***New: Updated syllabus including project grading. Updated course title.

Changed title, course syllabus/description and intended rank to reflect a course that advanced undergraduate actuarial science majors could take in addition to MAQRM students. With this, we would like to keep it as a 5000-level course. *(by Husen,William J on 09/29/2022 01:13 PM)*

- See feedback email sent to dept. 10/6/21 RLS *(by Steele,Rachel Lea on 10/06/2021 02:55 PM)*
- Hi Bill, A 5000-level course is intended for both undergraduate and graduate students. If this course is only for the MAQRM program, then it should be pitched at the 6000 level, right? *(by Vankeerbergen,Bernadette Chantal on 09/20/2021 12:09 PM)*

Workflow Information

Status	User(s)	Date/Time	Step
Submitted	Husen,William J	09/17/2021 11:24 AM	Submitted for Approval
Approved	Husen,William J	09/17/2021 11:29 AM	Unit Approval
Revision Requested	Vankeerbergen,Bernadette Chantal	09/20/2021 12:09 PM	College Approval
Submitted	Husen,William J	09/21/2021 08:56 AM	Submitted for Approval
Approved	Husen,William J	09/21/2021 08:57 AM	Unit Approval
Approved	Vankeerbergen,Bernadette Chantal	09/22/2021 02:09 PM	College Approval
Revision Requested	Steele,Rachel Lea	10/06/2021 02:56 PM	ASCCAO Approval
Submitted	Husen,William J	09/29/2022 01:13 PM	Submitted for Approval
Approved	Husen,William J	09/29/2022 01:13 PM	Unit Approval
Approved	Vankeerbergen,Bernadette Chantal	10/15/2022 12:39 PM	College Approval
Pending Approval	Cody,Emily Kathryn Jenkins,Mary Ellen Bigler Hanlin,Deborah Kay Hilty,Michael Vankeerbergen,Bernadette Chantal Steele,Rachel Lea	10/15/2022 12:39 PM	ASCCAO Approval

Math 5588 – Practicum in Actuarial and Quantitative Risk Management

- **COURSE OBJECTIVES**

  To introduce students to the real world practice and to improve their communication skills. There will be presentations by visiting professionals on topics drawn from their fields of expertise, and projects on real application problems will be assigned. Students will work as groups on these projects and give presentations.

- **CLASS FORMAT**

  Lecture and student presentation – 3 hours a week

- **PREREQUIAITE**

  The prerequiaite includes working knowledge in Interest Theory (Math 3618), Probability (Math 4530 or equivalent), Statistics (Stat 4201 or equivalent), and Linear Algebra (Math 2568 or equivalent).

- **TEXTBOOK**

  There is no text book for the course.

- **GROUP PROJECTS**

  There are several group projects with topics from actuarial and quantitative risk management. In each project, a group should discuss and study the case, create a model for the problem, decide on an analytical tool, gather and process data and information, analyze results, and prepare a presentation. Visiting professionals will be available to answer questions between classes when students working on each project. After completing the project, each group presents its result to the whole class, and visiting professionals give comments and feedbacks. The projects will be evaluated based on

  - Problem formulation and model selection
  - Selection and design of programing or computation methods
  - Data preparation and processing
  - Quality and accuracy of the result
  - Presentation

  Presentations should be in a format similar to PowerPoint with supporting technical detail or programming output.

- **INDIVIDUAL PRESENTATION**

  Each student will give a presentation on a topic of his/her choice. The topic should be from actuarial science, financial math, data science, business finance, economics, or any related field. Examples of a presentation include

  - exposition of an idea or theory,
  - introduction to a technical tool or platform,
  - modeling of a real problem and its solution,
  - *etc.*

  Each student should discuss his/her choice of topic with the instructor before starting working on the presentation, and should discuss the result with the instructor before presenting to the whole class. The presentation should be in a format similar to PowerPoint with supporting technical detail, computer output, or program demonstration. Individual presentations will be arranged in the later part of the semester.

- **EXPECTED WORKLOAD**

  Students are expected to spend 6 or more hours each week to work on the class projects or individual presentation.

- **PEER EVALUATION**

  There will be a midterm and a final peer evaluation. In these evaluation, each student gives him/herself and each member of his/her group a score from A to E based on

  - Attendance of group work
  - Participation and contribution to each group project

  A peer evaluation form is attached.

- **GRADE**

  The course grade is based on student's participation, performance, projects, and the class presentationa. The distribution is

  - Attendance, 10%,
  - Group projects, 50%
  - Individual presentation, 20%
  - Peer evaluation, 20%

  A letter grade will be determined by the total percent a student receives for the semester as 90-100% A, 80-89% B, 65-79% C, 50-64% D.

- **SAMPLE SCHEDULE**
  - Week 1 Introduction and Technical Preparation
  - Week 2 *Splines and Smoothing, Bootstrap* (Liberty Mutual)
  - Week 3 *Rate Making* (Nationwide P& C)
  - Week 4 & 5 *Brownian Motion and Its Application in Quantitative Finance*
  - Week 6 *Pension and Retirement* (Aon)
  - Week 7 *Market Optimization and Elasticity Modeling* (Cincinnati Insurance)
  - Week 8 *Annuities* (Nationwide Financial)
  - Week 9 *Life Insurance* (Nationwide Financial)
  - Week 10 *Inflation and the Impacts to Pricing Auto Insurance* (Grange)
  - Week 11 & 12 Individual Presentations
  - Week 13 & 14 *Data Analysis and Modeling Projects*

- **ACADEMIC MISCONDUCT**
  It is the responsibility of the Committee on Academic Misconduct to investigate or establish procedures for the investigation of all reported cases of student academic misconduct. The term "academic misconduct" includes all forms of student academic misconduct wherever committed; illustrated by, but not limited to, cases of plagiarism and dishonest practices in connection with examinations. Instructors shall report all instances of alleged academic misconduct to the committee (Faculty Rule 3335-5-48.7). For additional information, see the Code of Student Conduct at `http://studentlife.osu.edu/csc/` .

- **DISABILITY SERVICES STATEMENT**
  Students with disabilities that have been certified by Student Life Disabilities Services (SLDS) will be appropriately accommodated and should inform the instructor as soon as possible of their needs. SLDS contact information: `slds@osu.edu`; 614-292-3307; 098 Baker Hall, 113 W. 12th Avenue.

Evaluation

Self Evaluation

Your Name _____ Grade: _____

Comments

Peer Evaluation

Team Member Name _____ Grade: _____

Comments

Team Member Name _____ Grade: _____

Comments

SAMPLE LECTURES AND PROJECTS

- **Property and Casualty Insurance Pricing Considerations**
  Lecture Slides and Project Assignment

Property and Casualty Insurance Pricing Considerations

January 28, 2020

Bob Weishaar

1

Introduction

Bob Weishaar

- Ph.D. Mathematics and Masters of Applied Statistics, OSU
- Fellow of the Casualty Actuarial Society (FCAS)
- VP, Small Business Analytics, Cincinnati Insurance Company
- Pricing, research, and consulting roles at Farmers, Nationwide, LexisNexis, State Auto, and Motorists

The views expressed in this presentation are those of the speaker and not any third party.

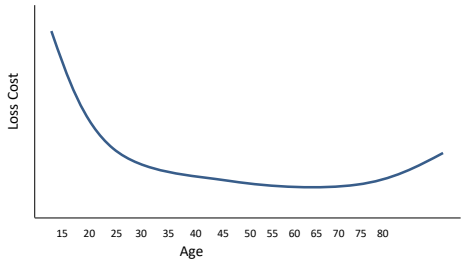
2

Agenda

- Introduction to personal lines auto pricing considerations
- Integrated Pricing
- Price elasticity modeling

3

How do average insurance losses vary by driver age?



4

What are some other driver attributes that might matter?

- How important are prior accidents and violations?
- Is a person's credit score related to whether they are likely to have an accident?
- Should you pay more or less for selecting higher "limits" of insurance?
- What about *prior* limits of insurance?
- Annual miles driven?
- Work schedule?
- Number of Facebook friends?

5

What are some vehicle attributes that might matter?

- Type of vehicle
- Age of vehicle
- Number of previous owners of car
- Vehicle safety features … making cost higher or lower?

6

What geographical data matters?

- Population density (city vs country)
- Crime statistics
- Weather patterns
- Traffic patterns: What would a roundabout do to frequency and severity?

7

Cell phone data

- Could you calculate annual miles driven from a cell phone?
- How would you know if the cell phone owner was driving?
  - Side of car entered
  - "Finger print"
- What else might matter?
  - Braking, acceleration, and cornering
    - How does this compare to the use of age?
    - How about prior accidents and violations?
  - Speed
  - Time of day
  - Location
  - Distracted driving

8

Simplified auto pricing rating algorithm

- Policy premium = Sum (Vehicle premiums)
- Vehicle premium = Sum (Coverage premiums)
- Coverages: Bodily Injury (BI), Property Damage, Comprehensive, Collision, Uninsured Motorists, etc.
- BI premium =
  - Base rate (e.g., $341)
  - * territory factor (e.g., 1.7 for urban versus 1.0 for rural)
  - * age factor (e.g., 3.2 for 18 year old versus 1.0 for 50)
  - * gender factor (e.g., 0.95 for female)
  - * credit factor (e.g., 0.70 for good credit)
  - * increased BI limit factor (e.g., 1.5 for $300,000)
  - * prior BI limit factor (e.g., 0.8 for $300,000)

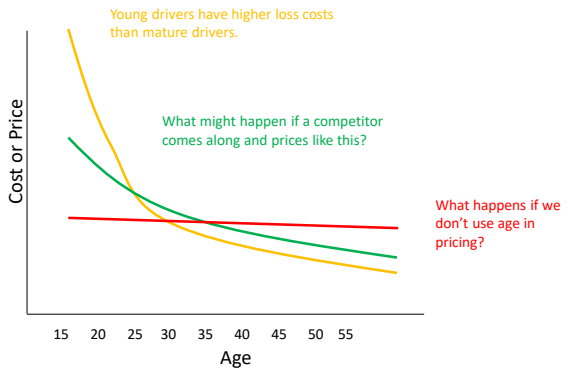
9

Cost-based pricing is a big piece of the puzzle

- First, the overall rate level:
  - If you collect $100 of premium:
    - How much can we pay out in losses? Expenses? Should total be < or > $100?
    - We get investment income
    - We need a profit provision with a risk margin
  - We're pricing ahead of time, so also need to consider:
    - Loss development (do we really know how much our products have cost us?)
    - Loss and premium trends (inflation and other macroeconomic trends)
- Then, segmentation to understand relativities:
  - Age, geography, credit score, etc.
  - Uses predictive modeling (usually regression)

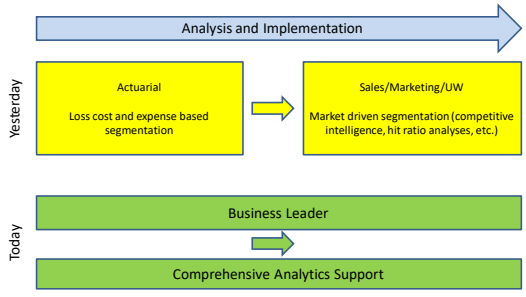
10

Cost isn't the only factors to consider when pricing



11

The Actuary's evolving role in decision making



12

Metrics that influence pricing decisions

- Costs
  - Losses
  - Expenses

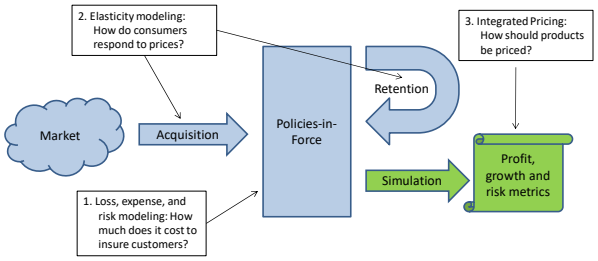
Historical domain of the actuary

- Retention rate: acceptance of renewal offers
- Quote volumes
- Hit (or close/conversion) rate: acceptance of new business offers
- Transition rates
  - Deterministic: Driver and vehicle aging
  - Random: Increasing limits, buying a new car
- Competitive position
- Price sensitivity

13

What is Integrated Pricing?

- Use of the pricing lever to meet financial objectives
- Profitable acquisition and retention throughout the consumer lifecycle

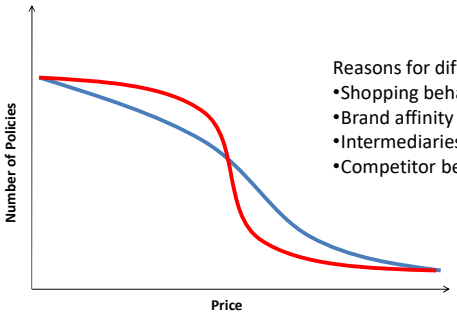


14

Price elasticity: how demand varies with price

15

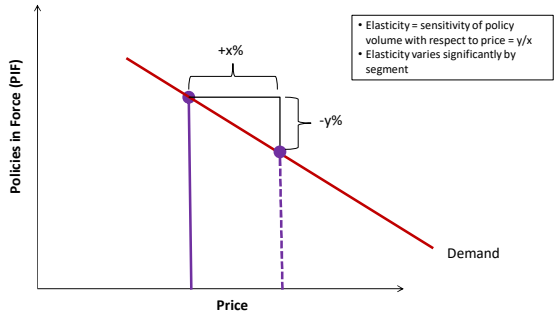
Price elasticity varies significantly by segment



Reasons for differences
- Shopping behavior
- Brand affinity
- Intermediaries
- Competitor behavior

16

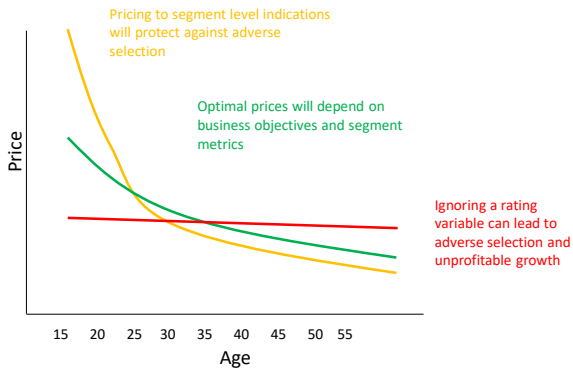
Demand curves are derived by elasticity modeling



- Elasticity = sensitivity of policy volume with respect to price = y/x
- Elasticity varies significantly by segment

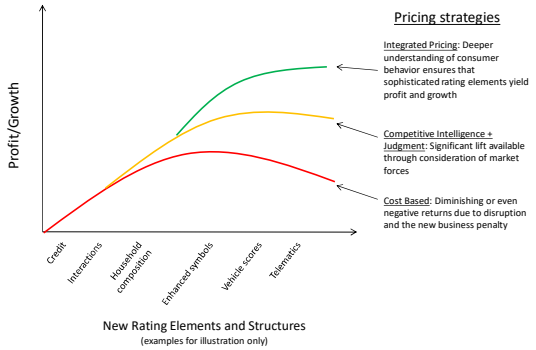
17

Optimal segment level pricing is a balancing act



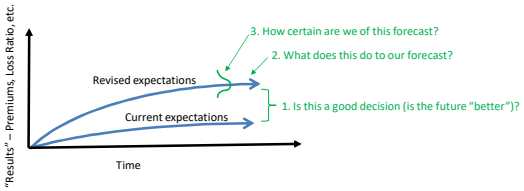
18

Techniques have evolved to derive profitable pricing strategies



19

Good decision making vs accurate forecasting



20

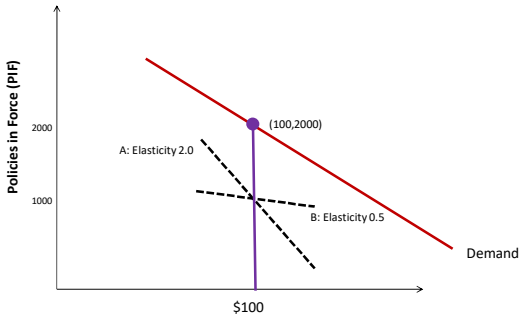
Notes on "Price Optimization"

- Price Optimization is a broad expression applied to a set of pricing strategies in the insurance industry
- Several states have expressed concerns about some of the techniques being employed
- Many of the examples in this presentation and corresponding problem set are used only to illustrate mathematical and economics concepts
- Besides being bound by local laws and regulations, Actuaries are bound by codes of professional conduct
- Global regulatory frameworks vary considerably

Due to the negative connotation of the phrase "price optimization," these techniques are now often referred to by other terms such as "integrated pricing"

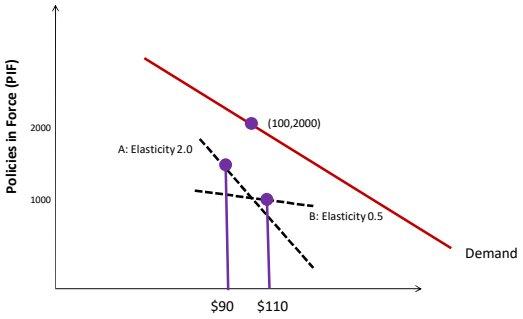
21

Integrated Pricing can shift the demand curve



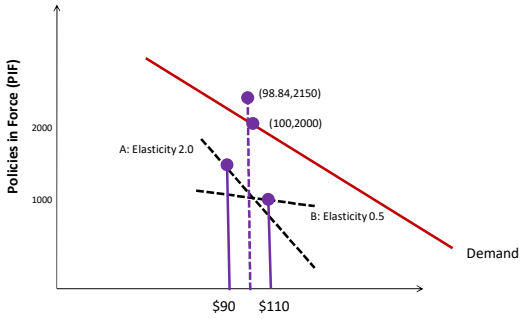
22

Integrated Pricing can shift the demand curve



23

Integrated Pricing can shift the demand curve



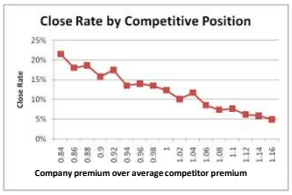
24

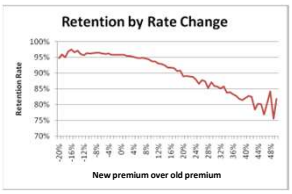
Integrated Pricing can shift the demand curve

25

Elasticity modeling

Elasticity is the sensitivity of policy volume with respect to price





New business responds to differences in competitive position

Renewal business responds primarily to rate change

= Percent change in PIF divided by percent change in price

26

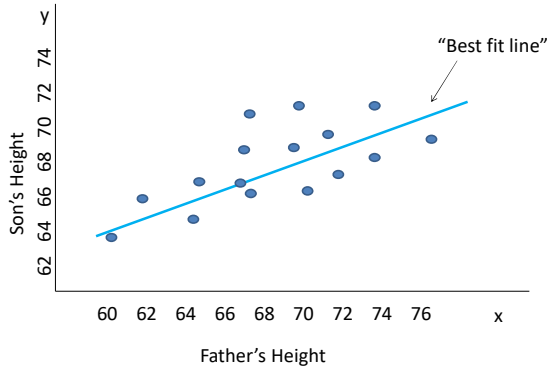
Elasticity modeling terminology

- Strike rate: Percent of offers accepted
  - Close rate for new business
  - Retention for renewals
- Premium differential: Relevant price comparison
  - Competitive position for new business
  - Typically premium change for renewals
- Conceptually: A regression model where elasticity is the coefficient of the premium differential.

  Strike rate=(close rate predictors)+(prem diff)*(elasticity predictors)

27

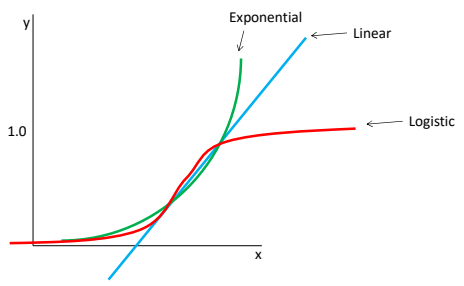
Regression modeling is often used to predict metric values



28

Regression modeling and link functions

Link functions determine the range of acceptable predictions



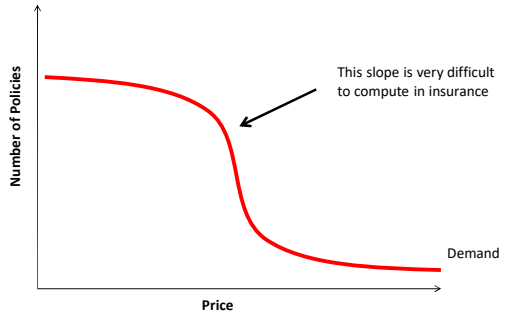
29

Examples of regression models in insurance

- Loss
  - Frequency: Log link with Poisson error
  - Incidence: Logit link with binomial error
  - Severity: Log link with gamma error
  - Pure premium: Log link with Tweedie error
- Strike rate: Logit link with binomial error
- Marketing response: Logit link with binomial error
- Elasticity: Nonlinear form
  - Need strike rate between 0 and 1
  - Need coefficient of premium differential to be negative (higher price → lower volume)

30

Economics of Insurance Product Pricing



31

Elasticity modeling is a difficult business problem

- Target is unknown
  - Contrasted with loss or strike rate modeling
- Cannot offer same customer two different prices at same point in time. Alternatives:
  - Random price tests: only way to get truly unbiased estimates of elasticity (common outside United States)
  - Other methods introduce bias
    - Over time – competitor prices and marketing programs change
    - Between segments – Segment behavior might cause differences
- You intend to change one of the predictors (price), so an accurate coefficient is important. In many models, only the prediction is important.
  - Suppose: Strike rate = a*(age) + b*(years licensed) + c*(price)
    - If a and b are "wrong", not a big deal. (Should still try to get this right, of course)
    - If c is wrong, results could be disastrous. But price is generally a function of the other attributes!

32

Conclusion

- Actuaries play critical functional roles in insurance companies:
  - State rate indications
  - Reserve analyses
  - Etc.
- Product Management centralizes P&L (profit & loss) responsibility.
  - Often filled by Actuaries
  - Increasingly rely on actuaries for broader analytics support
- Even without sophisticated models, a commitment to the portfolio simulation mindset improves decision making

33

Using elasticity to meet business objectives

Find the price change that maximizes total profit…

Price change	0%
Price	100
Costs	80
Volume	1000
Elasticity	2
Total profit	20000

34

P&C Pricing Considerations

January 28, 2020

Robert_Weishaar@cinfin.com

Definitions:

- Profit = (Price-Costs)*(Policy Count)
- Margin (which is profit percentage or ratio) = (Price-Costs)/Price
- Costs = losses + expenses
- Strike rate: Percent of offers accepted, called close (or hit or conversion) rate for new business and retention for renewals. We'll take strike rate to be a function of price in these problems.
- Elasticity: -(% change in strike rate) / (% change in price) = $-(\frac{\Delta S}{S})/(\frac{\Delta P}{P})$
- Elasticity can also be written as $-(\frac{S'}{S})/(\frac{P'}{P})$ which simplifies to $-\frac{S'P}{S}$
- Be careful of negative signs! In the math, elasticity is a positive number that measures the strength of an inverse relationship. In discussions, the negative direction might be emphasized by actually saying it is "negative."

Note: Unless otherwise directed, treat elasticity in a "simple" manner. For example, if price is currently $100, volume 1000, and elasticity 2, then a 10% drop in price to $90 increases the volume by 20% (=2*10%) to 1200. On the other hand, if you drop the price first to $95 and then $90 and use an elasticity of 2 in each step, you'd get a new volume of ~1216. In reality, elasticity is a *point* estimate and doesn't apply to every jump in price.

Problems (Hint: Use Excel Solver add-in when finding numerical solutions.)

0. Suppose there are two segments A and B with the following characteristics:

	A	B	Cummulative
Price	200	200	
Profit per policy	30	30	600,000
Policy Volume	10,000	10,000	20,000
Elasticity	4	2	

   a. Find prices that maximize profit and hold cumulative volume constant. ($191.67 & $216.67 → $641,667)
   b. Find prices that maximize volume and hold cumulative profit constant. ($188.21 & $213.20 → 21,038)

1. Find the one-year profit maximizing margin as a function of elasticity only given that:
   - Elasticity, cost, and quote volume are constant
   - Profit = (price-cost) * quote volume * strike rate

   (This is a calculus problem, and you'll need to use the product rule. Profit is a function of price, and you want to maximize profit by changing price. Strike rate is also a function of price. In the end you should have margin as a simple function of elasticity … all other constants should drop out.)

2. Suppose your book consists of blue and red cars with the following metrics:

	Blue cars	Red cars
Loss	300	350
Expenses	60	50
Elasticity	2	4
Quote volume	2000	1000
Close rate	40%	30%
Price	500	500

   a. Find prices that maximize profit dollars
   b. Do these prices agree with your derivation in #1? Why or why not?
   c. Maximize profit subject to the constraint that no volume is lost.
   d. Maximize volume subject to the constraint that no profit is lost.

3. Your competitive intelligence team has notified you that competitors have started using sod type in pricing home insurance. Houses with fescue lawns have losses that are 20% higher than those with Kentucky blue grass. Competitors begin charging 5% more for fescue and 5% less for blue grass. You study the segments and find that the fescue owners are more elastic. Given these values:

	fescue	BG
quotes	1000	1000
close rate at competitive position of 1	10%	10%
Previous prices (you and competition)	$300	$300
Actual losses and expenses	$300	$250
New competitor price	$315	$285
Elasticity	3.0	1.5

   a. What happens to profit and volume if you do nothing in reaction to your competitors' changes? (For simplicity, use $300 in the denominator for all price change calculations.)
   b. Should you implement a sod factor in your rating plan?
   c. What factors should you use if you want to maximize volume while maintaining original profit levels?
   d. Assuming regulators are fine with sod type as a rating variable, could you support the factors you selected in c?

4. You are devising a new rating plan that rates only on how long a policy has been in effect. You are given the following information:

Loss and expense	400
NB Elasticity (sensitivity to change from current price)	5
Renewal Elasticity (sensitivity to rate change from prior term)	0.5
Price	500
Quotes	1000
Close rate (at current prices)	20%
Retention rate (when no rate change)	90%

   Your new prices should be of the form: price=xy^t where t is the number of terms the policy has been in effect and x and y are constants that are fixed from the beginning (you can't change them each year). For example, x=100 and y=0.9 means new business pays $100 and each year the customer gets a 10% discount. y=1.2 means they pay 20% more each year. Let t=0 for new business. You will start writing this business in a new company. Note that for new business, elasticity measures sensitivity to a change from the current price of $500. For renewals, customers are sensitive to the change from the price paid in the previous term.
      a. Determine x and y so that profits are maximized in the third year of the new company's operation.
      b. Determine x and y so that volume is maximized in the third year of operations given that profit is limited to 15% of premium.

      NOTE: In the "third year" of operations, the company will have new business, business that is one year old, and business that is two years old. You want to consider the whole book of business.

5. Recessive Corporation is a small auto insurance company with 100 annual car policies, each priced at 120 dollars. With no change to price, the company expects to lose 20 of these policies to competitors resulting in an 80% policy retention. A pricing actuary has advised you that the renewal elasticity for this business is 2.0.
   a. Calculate the expected policy retention if Recessive decides to decrease their price to 60 dollars, treating elasticity in the simple manner described in the Note above.
   b. Now suppose the strike rate for a given price (x) is represented by the function, $Y = \frac{1}{1+e^{ax+b}}$. Using the information from Recessive, calculate the a and b values for this function.
   c. What is the expected policy retention after the same decrease to premium using the new logistic form?
   d. Identify an issue which may arise if Recessive decides to rely on the answer in part (a).
   e. To check that the simple approach is reasonable for small price changes, calculate the percent impact to volume if price increases to $121.20 (a 1% increase).

   Hint: Apply the quotient rule to find Y'. Don't forget elasticity = $-\frac{Y'x}{Y}$

SAMPLE LECTURES AND PROJECTS

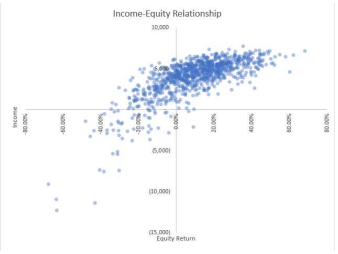
- **Splines, Smoothing, and Bootstrap**
  Lecture Slides and Project Assignment

Introduction to Splines and Smoothing

1

A representative problem

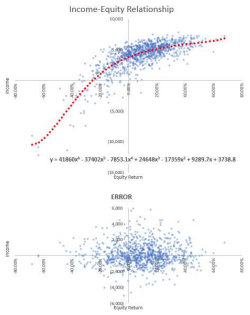
- I would like to be able to represent the average relationship between Equity Return and Income.
- Many ways to do this.
- What are my choices?
- What is best?



2

Polynomial Regression

- Common approach is to fit a polynomial regression.
- Several problems can be material:
  - Curve really isn't a polynomial, so you wind up with systematic misfits.
  - Curve can have kinks and twists driven by noise in the data.
- Extrapolation of high-order polynomials can lead to injury or death.



3

Smoothing Splines

- The main idea behind smoothing splines is to fit low-order piecewise regressions to the data.
  - Piecewise means that you don't have to have a single model to fit the whole curve – so you minimize systematic misfits.
  - Low-Order means that are less likely to fit to noise in the data and extrapolation is generally safer.
- Basic process is:
  - **Define the boundaries between the pieces ("knots").** In general want wider pieces where the data is thin or regular; narrower pieces where the data is heavy or irregular.
  - **Choose the polynomial function ("basis function").** The higher the polynomial the smoother the fit. May allow fewer pieces.
  - **Fit the piecewise regressions.**
  - **Can enhance this method by choosing something more sophisticated than polynomial basis functions.** B-splines are a popular choice that is more numerically stable than ordinary polynomials.

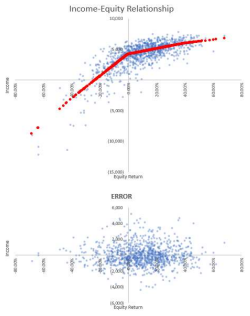
4

Linear Spline

- We will choose knot points $\xi = \{-80\%, -40\%, 0\%, 40\%\}$ to define the linear segments.
- We will fit a linear regression for each segment, but we also want the endpoints to link up (continuity).
- A specific functional form does this:

$$y = \beta_0 + \beta_1 x + \sum_{k=1}^{4} b_k (x - \xi_k)_+$$

- Less parameters, similar fit, safer extrapolation!



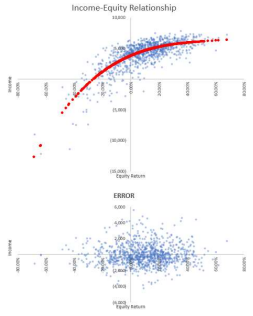
5

Cubic Spline

- Continuing the previous example, we don't like the "sharp angles"
- We can correct this by using a higher order polynomial AND also requiring the that first derivative be continuous (eliminates the sharp angles).
- In general, for a polynomial of order D and K knots…

$$y = \beta_0 + \beta_1 x + \ldots \beta_D x^D + \sum_{k=1}^{K} b_k (x - \xi_k)_+^D$$

- Only one additional parameter needed for each increase in order!



6

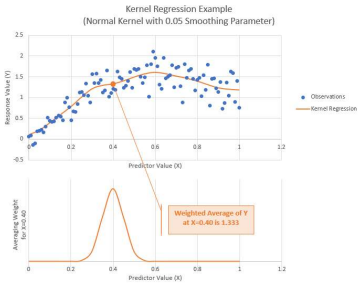
Kernel Regression

- The big issue with splines is that knot and polynomial choice are highly judgmental, with a lot of interaction between these choices. Results may not be very stable.
- Kernel Regression is a technique that achieves similar results, with fewer choices, and good stability.
- Basic process is:
  - We are going to take a local average of all the points that are nearby.
  - Choose a "weighting kernel" that shows me how to compute the locally weighted average.
  - Compute the weighted average anywhere I want to know the smoothed value.
  - Can extend this method by "adjusting" the local points with polynomials, then taking the average. Used to provide better tail extrapolations.

7

Kernel Mean

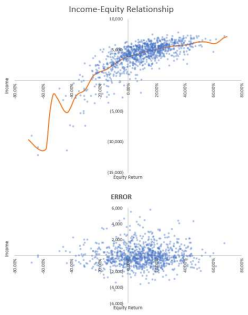
- The mean is a 0^{th} order polynomial.
- The Normal Kernel is a common choice for computing the weighted average. Data structure can drive selection of other kernels.
- The main modeling choice is the kernel standard deviation "scale" or "smoothing" parameter.
- Choice is judgmental. Generally want to see a few bumps in the fitted curve but not too many.



8

Kernel Mean

- Using a kernel standard deviation of 0.0333 we see the functional form that is indicated by the data itself. (No basis function choice needed.)
- But we also see that the mean gets very noisy in the tails of the data. There is little data here and the local mean behaves erratically.
- Would like to expand the averaging window, but this leads to flattening of the data and losing information about the data relationship.



9

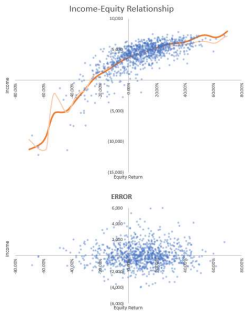
Kernel Mean VBA Function

```
Function KernelMean(EvaluationPoint, DataColumnX, DataColumnY, ScaleConstant, DistributionName)
    'Computes kernel mean at EvaluationPoint using the symetric "Normal" kernel or
    'the assymetric "Lognormal" kernel, based upon the sample DataColumn.  Degree of
    'smoothing is controlled by ScaleConstant.
    '
    'May insert other kernels within the SELECT structure.

    For i = 1 To DataColumnX.Rows.Count
        Select Case DistributionName
            Case "Normal": wt = Application.NormDist(EvaluationPoint, DataColumnX(i), ScaleConstant, False)
            Case "Lognormal": wt = Application.NormDist(Log(EvaluationPoint), Log(DataColumnX(i)), ScaleConstant, False) / DataColumnX(i)
        End Select

        SumWt = SumWt + wt
        KernelMean = KernelMean + Application.Max(wt, 0) * DataColumnY(i)

    Next i
    |
    KernelMean = KernelMean / SumWt
End Function
```

10

Linear Kernel Regression

- For this example, we are using a kernel standard deviation of 0.050 – a wider window so that more data goes into tail estimates.
- But we also use a linear trend adjustment to "centralize" points before we compute the average.
- Gives us the best of both worlds:
  - Get a smooth relationship.
  - Only need to choose two parameters, standard deviation and trend factor.
  - No knot selection.
  - Data tells us the functional form rather than trying to guess a form.



11

Introduction to Bootstrap Estimation

12

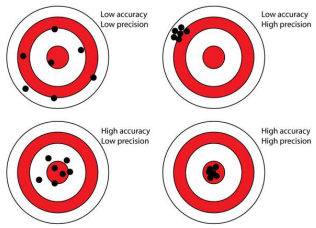
Stochastic Projection

- What do we mean when we say stochastic projection?
  - Stochastic Integration
    Estimating a statistical measure such as expected value, a percentile or CTE.
  - Solve Heat-Equation PDE
    Estimating risk-neutral values.
  - Real-World Scenario Testing
    Evaluating what would "really happen" in individual scenarios.
  - Prescribed Calculation
    Prescribed calculations performed on prescribed scenarios.
- "Nested Stochastic" situations combine multiple techniques.

13

Simulation Performance

- Simulations processes typically highlight precision. (Is my answer this quarter similar to last quarter?)
- Close attention to specifics of valuation problem can improve precision.
- Accuracy matters over the long term. (Both good and bad valuations eventually converge to the same value!)
- Inaccuracy usually the result of poor modeling techniques.



14

Bootstrap Estimation

- You can't control what you can't measure!
- When you run a simulation and compute a metric, how do you know if you are close to the "true" answer?
- The brute force approach is to…
  - Run a really large simulation and see what result you get.
  - Run multiple simulations and see how "repeatable" your answers are.
  - Both are hard to do.
- Can you do something easy and get the same information?

15

Bootstrap Estimation

Concrete Example

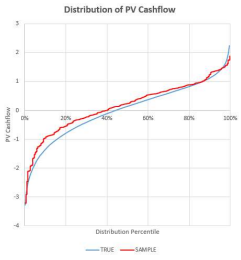
- You run 250 scenarios and measure the 95CTE of PV cashflow.
  - Is the value you measured close to the true fully-converged value?
  - How many scenarios would you have to run to be close enough?
- The "information" from your simulation consists of 250 PV cashflow values.

16

Bootstrap Estimation

Some facts that will be important…

- Sample distribution asymptotically approaches true distribution.
- In a sample, the "range" of distribution is much more accurately represented than "level" of distribution.
- In a sample, the center of distribution is more smoothly represented than the tails.



17

Bootstrap Estimation

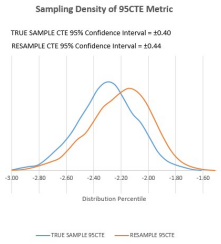
Method

1. From the observed 250 PV cashflow, grab a random sample (with replacement) of n PV Cashflows. This is a statistically valid proxy for running an entirely new simulation of n scenarios – but without the work!
2. Compute your metric of interest (95%-ile of PV Cashflow).
3. Repeat this process many times to estimate the standard error (sampling range) of your metric.
4. The asymptotic distribution of your sample metric is Normal, and you can use the "square-root rule" to estimate required sample size.

18

Bootstrap Estimation

- The "true" sampling distribution (hard to compute) and the "resample" sampling distribution have the same shape.
- The resample location is wrong because the resampled distribution doesn't reflect the mean level very well.
- But the 95% confidence range (range of results) is very accurately reproduced.
- Bootstrapping can tell you how precise your answer is, but it cannot tell you how accurate.
- A small sample may not provide a good representation for tail metrics, but this tends to result in wide confidence ranges – you are getting the right feedback!



19

Project

20

Part 1

When computing the value of equity derivatives, a common technique is to generate "risk-neutral scenarios" – simulations of equity paths – and use these paths to compute the expected payoff of the derivative. This technique is called Monte Carlo valuation.

For this project, we are going to use a Monte Carlo valuation to duplicate the Black-Scholes formula results. (There are multiple ways to solve the same math problem! So you can check your answers using closed form formulas.) Assume the following values.

1. In spreadsheet or other mathematical modeling package, simulate risk-neutral equity scenarios.
   - Start each scenario with S_0=1 and use $S_t = S_{t-1} + rS_{t-1} \cdot dt + \sigma S_{t-1} \cdot Z \cdot \sqrt{dt}$ where Z is a random draw from a standard Normal distribution. (Hint: you may need to use the Inverse CDF method to generate the Z draws.)
   - Generate 250 random scenarios, in monthly time increments, out to one year.
   - Store the results in an array with 250 rows (250 scenarios) and 13 columns (t=0,1,…12).
   - How do you know that you did this simulation properly? (Hint: what are the assumptions of the Black-Scholes model?) What statistical properties of this data matrix can you evaluate. Show those tests.

21

Part 1 (continued)

2. For values of S at t=1year compute $e^{-r} \cdot \max(0, X - S)$. This is the present value of the put payoff at the maturity of the option.
   - Compute the average value of $e^{-rT} \cdot \max(0, X - S)$ across all 250 scenarios. This is a Monte Carlo estimate of the put option value.
   - Compare this result to the Black-Scholes formula for put price. Are they close? How do you explain what you see?
3. Use the Bootstrapping technique to resample your 250 present value put payoffs.
   - Generate twenty new samples of size n=250 and compute the average present value put payoff for each new sample.
   - What is the standard error of your result in part 2?
   - If I wanted to be very sure that my Monte Carlo answer was accurate to within ±0.001, how many equity scenarios would I have to simulate?
   - Run a sample of this size and show that your new answer is within this tolerance. (Be careful – this is a trick question!)

22

Part 2

In practice it is usually very time consuming to run a Monte Carlo simulation, so we want to reuse our calculation results wherever we can.

For this project, we are going to reuse the simulation results from the previous problem. We will develop a "pricing formula" for puts with **six months to maturity** by comparing the payoff in month 12 to the equity level at month 6. (The time to maturity between month 6 and month 12 is six months.)

<u>Kernel Approach</u>

1. Reusing the simulation results from Part 1, compute the six-month present value of the put payoff at maturity $e^{-r \cdot 6\ months} \cdot \max(0, X - S_{12})$. This present values the 12 month put payoff to a date six months earlier.
2. Make a graph that plots $e^{-r \cdot 6\ months} \cdot \max(0, X - S_{12})$ versus the equity level in month 6, S_6. Add to this graph a plot of the Black-Scholes price of a put with six months left to maturity, given the equity levels at S_6. How do you interpret what you see?
3. Fit a kernel regression through the cloud of points in the previous graph. What smoothing constant makes sense. Compare this kernel regression to Black-Scholes price. How would you describe the quality of fit? When is the kernel safe to use?
4. Use the bootstrap approach to estimate the standard error of the kernel regression as three different levels of S_6=(0.8, 1.0, 1.2). (Hint: You need to resample both S_6 and S_{12} from the original simulation.) Do these results confirm your observations in 3?

23

Part 3

- <u>OLS Approach</u>
- Repeat the previous analysis, substituting ordinary least-squares regression (cubic polynomial) for the kernel regression.
- Compare and contrast with the results you saw using the kernel approach. (Hint: you may want to plot the error=[regression answer]-[Black-Scholes formula answer] vs S_6 for both the kernel regression and the ordinary least-squares regression.)
  - Which method gives you more accurate valuations?
  - Which method gives you more precise valuations?
  - Do your answers depend upon the level of S_6?
- For many practical applications, I care about the first derivative of the option value curve. How do the Kernel and OLS methods compare? How could I adjust each method to provide better estimates of the first derivative?

24