*26 July 2022*

Guiding Principles for Authors and Associate Editors when Revising *Loss Data Analytics*

The first edition of *Loss Data Analytics* was completed in August 2018; some might say that the development was organic. We managed to get 19 authors from 8 countries in more or less in the same direction; the product is a free online book that was written in "R" and is hosted on Github.

In revising the book for the second edition, we now have the advantage of seeing the book in its entirety. We can use this base to promote cooperation and consistency among chapters. We did well with consistency of notation in the first edition but can now think strategically about how to develop topics within the book and how to make them accessible for different readerships.

A working group (dubbed the "Steering Committee") has developed the following sets of principles to guide authors and editors in developing a new and better version of *Loss Data Analytics*.

## I. From Inference to Analytics

One of our goals with the first edition was to emphasize the role of statistical inference and data within traditional loss modeling. The second edition expands on this theme by moving from statistical inference to *analytics*.

Specifically, we seek to have:

* More data-driven content (see General Principle II)
* Additional examples of exploratory data analysis (EDA), particularly regarding summarization and visualization
* In addition to model construction, in many instances we hope to promote issues associated with model selection
* Promote a greater use of simulation (e.g., mixed and compound distributions, also for numerical computations)

Because of its focus on the foundations, we do not anticipate that this text will feature extensive presentations of machine/statistical learning methods. Nonetheless, appropriate references in the "Further Resources" section will be helpful to for readers to develop this connection.

## II. Data, Data, Data

In contrast to traditional loss data models, we want to emphasize data demonstrations. In the new edition, we will compile a "data dictionary" that will provide a single source of different data sets.

* So for the examples and exercises, try *not* to emphasize examples that were developed for multiple choice (for professional exams) questions.
* Sometimes, small "toy" data sets, having just a few observations, can be useful to explain ideas.
* Strive to include analysis of data to illustrate the ideas.

## III. Data and R

We will use the "R" statistical software to illustrate loss data techniques. Consider using the work of Chapter 12 as a model to follow. Having said that, the book is about loss data analytics concepts, not about a specific software. The "R" scripts will be hidden (in the html version, in an appendix for the pdf version). In this way, the book will also be useful for Python or SAS users (for example). If an author thinks it useful, we can also include python or Excel scripts for secondary analyses. Authors can also look to incorporate examples done in the **Short Course** on *Loss Data Analytics*, available at: https://openacttexts.github.io/LDACourse1/ .

## IV. Types of Readers

We want to think beyond a traditional textbook where each concept is tightly integrated with prior portions of the book. Rather, think of this as a combination between this type of traditional textbook and a "wikipedia" type resource, where readers come in for selected sections.

With this as background, think about developing the book for the following different types of readers:

* A. *Beginning Students*. People with a year of calculus and one term of probability. No prior knowledge of mathematical statistics (but this course is a co-requisite), no knowledge of insurance/risk management. This book is their very first look at actuarial science.
* B. *Intermediate Students*. People with a background in calculus, probability, mathematical statistics (and possibly regression). At least one course in insurance/risk management.
* C. *Advanced Students*. People with more sophisticated mathematical background (e.g., type B about plus some stochastic processes) and good appreciation insurance/risk management.

The intent is that the book will be suitable for all types of readers. The main body of the text is targeted towards something in between types (A) and (B), with appendices to fill in the gaps. With appropriate signposting, the book will also be helpful for those that are closer to type (C).

## V. Goals

When a reader uses the book, the goal is *not* to teach them skills to pass a given actuarial exam. Hopefully, by learning this material, it provides students with the knowledge necessary to do so eventually. In the same way, we do *not* wish to get into all the nuances needed to immediately use the techniques in practice.

* For example, try to minimize applications of multivariate ideas such as GLMs. Our focus is on providing educational background so that learners can compete in the workforce and have productive careers.

Rather, the goal is to educate readers on the principles of loss data analytics that serve as a foundation for those who wish to pass professional exams and use these concepts in practice. With this guiding principle in mind, it is certainly reasonable for authors to include sample questions that have appeared on actuarial exams or complicated situations that are practice relevant. Moreover, it is helpful to align our table of contents with the professional associations (to the extent possible). This provides readers with additional incentives to learn the materials. But, their inclusion should be motivated by the basic education principle, not goals in and of themselves.

## VI. Definitions

The book relies on concepts from probability, statistics, insurance, and statistical software, among other foundation areas. It is important to define terms clearly - moreover, a definition should typically appear only once in a book. Otherwise, multiple definitions that are similar can sometimes provide contradictory information. Now that a sound first edition is available, when you are defining new terms, check out definitions that appear elsewhere include other chapters, appendices, and the glossary. To promote consistency, we also have an Appendix on Notation Conventions (but try to minimize use of mathematical symbols).

## VII. Technical Demonstrations of Results

We want to focus on the foundations and so short mathematical demonstrations of properties can be illuminating for some readers. However, we will use the "hide/show" feature so as not to distract from the main presentation of the ideas. So as not to intimidate early readers, try to minimize the amount of mathematical sophistication needed in the main body. For example, use "Proposition" instead of "Theorem" and use English words such as "is an element of" instead of a symbol. Naturally, this approach will make the work (slightly) longer but we can increase readership in this way. The Appendix on Notation Conventions is to maintain consistency among chapters, not promote the use of mathematical symbols.

## VIII. Actuarial versus General Topics

Because we are thinking of different types of readers, "signposting," or clearly labeling content, is important. In particular, we can expect some readers to have strong technical background in probability and statistics and wish to use the book for actuarial or insurance content. To achieve this goal, we want to clearly label sections or subsections as "actuarial" in nature.

## IX. Connections

Now that we have a first edition complete, we can think about enhancing connections among chapters. Chapter authors should look to see if a concept has already been introduced in prior chapters or the Appendices and take advantage of this content. Further, if an idea is presented that will amplified in later chapters, a short note to the reader foreshadowing additional explanation is very powerful. The goal is to minimize overlap and make each component as impactful as possible.

## X. Quizzes

End of the section quizzes were added to chapters, not by chapter authors, but by an independent team. Because the team was working across many chapters, they may not have had the same depth of content knowledge as the chapter authors. Because we wish to retain the quizzes, we ask chapter authors to review and revise the end of section quizzes as appropriate.

**Chapter Advice** Guidelines when Revising *Loss Data Analytics*

Please think about the following questions as you review each chapter:

* Does the structure make sense?
* What about the learning objectives?
* Ideas on the examples (too cookbook)?
* Does it fit in well with other chapters?

## Chapter 1 - Major Comment

How much do we want this book to be insurance-focused and how much analytics-focused?

* One route is to offer a little less insurance here and introduce only the terms necessary to understand the basic details of the examples in the text.
* We also want to move up an introduction to analytics using the current Chapter 13 as a base.
* Maybe call this “Introduction to Loss Data” and have the new chapter 2 provide the introduction to analytics.

We will combine these two introductory chapters or present shortened versions of each one to start the text.

As we reconsider this chapter, think about summarizing the respective application objectives. Those objectives are frequency modeling (Ch2), severity modeling (Ch3), aggregate loss computation (Ch5), ratemaking (Ch7), risk classification (Ch8), experience rating (Ch9, 12), insurance portfolio management (Ch10), loss reserving (Ch11), and dependence modeling (Ch14).

### Detailed Comments

1. (1.1.1) The connection to sports sounds odd, saying insurance is not as fun though more practical.
2. (1.1.3) Add non-life in “Most commercial and personal contracts are for a year”
3. (1.1.3) Provide some details or definitions of the terms in figure 1.1

## New Chapter 2

We will move Chapter 13 (and Section 14.1) up to become a new chapter that introduces analytics concepts. This chapter should emphasize “common methods” objectives. These objectives include the basic concepts of distributions (Ch2, 3), EDA/summary/visualization (a bit in Ch4), statistical inference including estimation (a bit in Ch4), model selection (a bit in Ch4, 13), simulation (Ch6), bootstrap (Ch6), Bayesian methods (a bit in Ch4, 9), and data preprocessing.

The following is a rough sketch of the new Chapter.

*2. General Process of Modeling in Data Analytics*

*2.1 Key Concepts*

Explains key concepts for modeling in the big data era, each with a short or a minimal explanation.

Key concepts are: Big Data, data driven, individualization, predictive (vs. generative. See Donoho’s 2017 seminal paper as well as Breiman’s 2001 one), “All models are wrong, but some are useful”, interpretability, non-parametric (vs. parametric), EDA (vs. CDA), robustness, computational statistics　(simulation, sampling, subsampling, resampling), etc.

Some items to be included already appear in the current Ch. 13, but others are not.

*2.2 Typical General Process*

Starts from similar contents to the current 13.2.1.

2.2.1 Types of Tasks

Explains types of tasks including regression vs. classification, supervised vs. unsupervised, prediction vs. explanation, etc.

2.2.2 Data Preparation

A shorter version of the current 13.1. Integrate with Chapter 1 Case.

2.2.3 EDA (as a stage in the modelling process)

Introduces very basic techniques of numerical summary and graphical summary.

The introduced summary techniques may include all the items in the current 4.1.1.

Integrate with Chapter 1 Case.

Perhaps mention in passing the concepts of PCA and clustering as in the current 13.3.1 (although we will not do so in this book).

2.2.4 Model Construction

A brief introduction to types of modeling methods focusing on basic MLE-based methods and extensions such as LM, GLM.

It partly covers the current contents of 13.3.2, but perspectives may be quite different.

Perhaps mention in passing the concepts of regularized GLM, tree-based methods, neural networks, etc. (although we will not do so in this book). May include some explanation of regularization if not in detail.

2.2.5 Model selection

A brief introduction to cross validation and information criteria.

*2.3 Example of modeling process*

Exploit the work done in Section 1.3 on the LGPIF.

## Chapter 2 - Major Comments

1. Consistent with our general principles, in this beginning chapter we will have to be very careful about the introduction of the basic definitions.

2. To streamline the presentation, we will not include the goodness of fit test here. Rather, to compare distributions, let's try to overlay plots of probability mass functions. This is similar to the approach used in Chapter 3.

3. We want to wrap up the chapter with a real data application of some type. What kind of problems do working actuaries run into while trying to model the frequency given the data available? Such a section would be valuable - not just for beginners.

### Detailed Comments

1. The level of detail given on the MLE estimation method may be overwhelming in the early chapters for new students, particularly the discussion of non-robustness and implementation of newton’s method. If the purpose of including MLE in the early chapters is to provide simple examples of application of MLE to loss data, perhaps retain the example to the Poisson, and move the binomial and negative binomial examples to an appendix (or hide these with an appropriate tab/link).
2. In section 2.6 it is shown that the negative binomial can be derived from a gamma mixture of Poissons. This same example is given in 4.4.1. Consider just keeping in one version (4.4.1?) and linking between the chapters.

## Chapter 3 - Major Comments

1. As with the prior chapter, we want to wrap up the chapter with a real data application of some type. What kind of problems do working actuaries run into while trying to model the severity given the data available? Such a section would be valuable - not just for beginners.

2. Move estimation of modified data to a later chapter.

### Detailed Comments

1. Examples in Sections 3.1.1, 3.1.2 and 3.1.3 refer to continuous distributions not yet introduced.
2. Maybe include a figure to illustrate links between different families of distributions.
3. Interesting figures to illustrate impact of parameters on the distributions.
4. Sub-section labelling for chapters 2 and 3 could be revised for more consistency. E.g., section 3.5 is ‘Maximum likelihood estimation’, whereas section 2.4 is ‘estimating frequency distributions’, but both just focus on MLE.
5. We need to introduce the basic concept of skewness somewhere. Chapter 3 may be the most likely place.
6. The Section 3.4.4 discussion of reinsurance needs to be coordinated with Chapter 10.

## Chapter 4 - Major Comments

1. Split off the Bayesian inference section into a separate chapter. This section could be revised to include a more data-driven introduction. If this new chapter was after the current Simulation and Resampling chapter, one could take advantage of that introduction to MCMC. The new chapter could also include a few ideas that would lead naturally into the Credibility chapter.

2. Re-organize the remaining topics into an estimation component and a model selection component. Currently, these two pieces are integrated in a way that makes it difficult for readers/instructors to read “selectively,” according to their needs. This suggestion is in accord with the traditional organization method - to present estimation first, then model selection. There was also sentiment on the committee to split these components into separate chapters.

### Detailed Comments

1. Discuss model selection for modified data (it currently does not exist).
2. Cross-validation is briefly introduced in Section 4.2.4 and elaborated upon in Chapter 6. See if you can introduce a data example in Chapter 4 to give the reader a better flavor as to how to make this happen.

## Chapter 5 - Major Comments

1. It would be helpful to add a real data application to the end of the chapter that shows how to build a probability model that describes the aggregate claims of an insurance company during a fixed period of time.
   1. Section 5.4.2 is about simulation. We will move this to the simulation chapter.
   2. The new section 5.4 will thus be shorter. It would be helpful to include “R” scripts to demonstrate the recursive formulas. Possibly do this in the context of a real data application.

### Detailed Comments

1. “Show Example” and “Hide Example” in all examples in this chapter should be replaced with “Show Solution” and “Hide Solution”.
2. The Follow-Up section in Example 5.2.1 should move to be part of the hidden solution of the example and not part of the text.
3. This sentence in the first paragraph of Section 5.3.2 is incomplete “Insurance on the aggregate loss S\_N, subject to a [deductible](https://openacttexts.github.io/Loss-Data-Analytics/C-AggLossModels.html) d, is called”
4. The title of Section 5.3.3 “Analytic Results” is not very informative.
   1. Should the subsection titles in Section 5.3. be more intuitive? E.g. 5.3.1. Moments and distribution of the aggregate loss, 5.3.2. Effect of Stop-loss insurance on aggregate losses, 5.3.3 Closed-form for aggregate loss distribution?
5. Consider keeping the proofs hidden to avoid distraction during reading.
6. In the formula of the expected value of the amount of the aggregate loss in excess of the deductible in the first paragraph of Section 5.3.2, the S is missing a subscript N.
7. Chapter 3 uses *X* to denote the loss incurred to the insured and *Y* denote the amount of paid claim by the insurer and differentiates between the two variables: The payment per loss and the payment per payment using the notation *YL* and *YP*. In Chapter 5, *XL and XP* are used instead of *YL* and *YP* with a note that says that this is done to reduce notation complexity. I recommend unifying the symbols across the book.
8. I think it would be good to include a paragraph in the Introduction section (5.1) about what is going to be covered in the rest of this chapter.
9. The learning outcomes that appear on other chapters/sections missing here.
10. Should there be a link for Technical Supplement 5.A.1/5.A.2 when mentioned?
11. Something missing in the second sentence in Section 5.3.2?
12. Chapter 5 quizzes need reconsideration. Some examples of these questions are given below.

* *Question:*Aggregate claims refer to
* *Question:*The two approaches in modelling aggregate losses use the
* *Question:*The individual risk model has *n* as the
* *Question:*The collective risk model has *N* as the
* *Question:*The frequency and severity distribution arise from the
* Question: A special case arises when the distribution of the loss amount from the 2-part framework has mean = b\_i (an insurance benefit) and variance = 0. This could represent
* *Question:*To apply the monte-carlo simulation to evaluate the distribution of aggregate losses

## Chapter 6 - Major Comments

For today's LDA, “simulation and sampling” is not simply an area with several useful techniques, but a fundamental approach that should be used universally for several purposes. Therefore, it is easier to understand, especially for beginners, if the techniques are introduced by the purpose for which they are used, rather than by each individual technique.

The purposes of simulation and sampling to be included in this book and the corresponding techniques are as follows.

* 1. To visualize complex distributions or to roughly compute their characteristic values by simulation --- random number generators, composition methods, the inverse transform method, crude Monte Carlo integration.
     + “Complex distributions” include compound distributions and mixture distributions, as well as distributions followed by statistics for which the distributions are not known.
  2. To accurately compute the characteristic values of complex distributions (except the Bayesian posterior distribution) by simulation --- variance reduction methods including importance sampling.
  3. To estimate the bias or standard error of an estimator or to obtain confidence intervals using resampling techniques --- jackknife, bootstrap.
  4. To tune or select a predictive model using resampling techniques --- cross-validation, bootstrap.

These roughly correspond to the following (new) suggested titles:

* 6.1 Definitions and Motivation for Simulation and Sampling
* 6.2 Random Number Generation from Distributions
* 6.3 Computing Characteristic Values by Simulation
* 6.4 Resampling Techniques for Statistical Inference
* 6.5 Resampling Techniques for Predictive Modeling

Although much of the work involves re-organization of current materials, there are important elements that are missing in the current version:

* + Ready-made random number generator functions, which names start from r in R. (They are used in the chapter as well as in other chapters, but no or little information is given.)
  + Basic examples of how to combine ready-made RNG to simulate complex distributions including compound distributions and mixture distributions.
  + Variance reduction methods. (6.1.3 deals with a thing related to variance reduction, but no general method is introduced.)

When re-organizing, bear in mind the following comments on the current section:

* Section 6.1 has problems from the learner’s point of view.
  + Need to introduce a new Section 6.1.3 on “Ready-made random number generator functions, which names start from r in R.”
  + In the old 6.1.3 (Simulation Precision), the importance of variance reduction is suggested, but no general prescriptions are given, which makes the subsection impractical.
  + The contents of the old 6.1.4 (Simulation and Statistical Inference) may be abrupt for beginners because the KS test is a bit too specific and rather unfamiliar to beginners.
* The contents of Section 6.2 are nice and adequate for introducing basic application examples of bootstrap for this book. But, as bootstrap has other different important usages including prediction error estimation and bagging, the title may be misleading.
* Section 6.3 compares and contrasts cross-validation to simulation techniques and bootstrap methods. But such a subject may be a bit too technical against beginners’ interests. Probably the section’s main subject should be the problem to be solved rather than the techniques.
* The importance sampling introduced in Section 6.4 is, or at least can be taken as, not standard. The accept-reject method in the section is not standard, either. Possibly, this section is just too short anyway.

For the new Section 6.3 on complex distributions:

* We can include the current Section 5.4.2 on aggregate claims distributions
* We should add a discussion of mixture distributions, following Sections 2.6, 3.3.5, and 3.3.6
* Statistics for which the distributions are not known is another possibility (or possibly in the new Section 6.4).

Additional comments.

1. I’d like to suggest that R code for basic bootstrap methods should be written without using the boot package (or any other specialized package for bootstrap techniques) at least initially so that the reader can more easily understand the basic procedures.

2. At least the current Example 6.2.2 doesn’t make sense in that it is not an example in which the bootstrapping technique is successful. A careful review of the examples and R code seems necessary. Comments for Example 6.2.2:

* Contrary to the author’s intention, theta\_3 is not a bias-corrected estimator, but only something to be used in the process of estimating the bias. A real bias-corrected estimator is “2\*theta\_1-mean(results$t)” instead of “mean(results$t)”, and in fact it will not work well in the current example.
* It seems that the example fails partly because the parameter, exp(mu), is too sensitive to the data.
* I feel the “second order correction” is not an appropriate name for theta\_2.
* If you take log(mu) instead of exp(mu), you get an apparently good example. Please try the following code.

sample\_x <- c(2.46,2.80,3.28,3.86,2.85,3.67,3.37,3.40,5.22,2.55,

2.79,4.50,3.37,2.88,1.44,2.56,2.00,2.07,2.19,1.77)

(theta\_1 <- log(mean(sample\_x)))

n <- length(sample\_x)

(theta\_2 <- mean(log(sample\_x))+var(log(sample\_x))/2)

#library(boot)

results <- boot(data=sample\_x,

statistic=function(y,indices) mean(log(y[indices])),

R=1000)

(theta\_3 <- 2\*theta\_1-mean(results$t))

plot(results)

print(results)

param <- function(x){

n <- length(x)

theta\_1 <- log(mean(x))

theta\_2 <- mean(log(x))+var(log(x))/2

results <- boot(data=x,

statistic=function(y,indices) log(mean(y[indices])),

R=999)

theta\_3 <- 2\*theta\_1-mean(results$t)

return(c(theta\_1,theta\_2,theta\_3))

}

set.seed(2074)

ns<- 200

est <- function(n){

call\_param <- function(i) param(rlnorm(n,0,1))

V <- Vectorize(call\_param)(1:ns)

apply(V,1,median)

}

VN=seq(15,100,by=5)

Est <- Vectorize(est)(VN)

matplot(VN,t(Est),type="l", col=2:4, lty=2:4, #ylim=c(-1,1),

xlab="sample size (n)", ylab="estimator")

abline(h=1/2,lty=1, col=1)

legend("topleft", c("raw estimator", "second order correction", "bootstrap"),

col=2:4,lty=2:4, bty="n")

## New Chapter 7

Add a **new** chapter on Bayesian methods. This will use elements from:

* Current Section 4.4 on “Bayesian Inference”
* Current Section 6.5 on MCMC

## Chapter 7 - Major Comments

1. Premium (written and earned), expenses (fixed and variable) & underwriting profit are provided as ‘insurance indicators’ in insurers’ yearly report in Malaysia. Add a description of these indicators based on industry data of any country, presumably using OECD publications.
2. Move the discussion of Section 7.2 on “Pricing Principles” to follow, or immediately precede, Section 7.6 on “Selecting a Premium.”
   1. Table 7.2 provides common properties of premium principles. An example using H(x) from this table would be nice.
   2. Nice explanation on personal automobile where one does pricing using expectations of losses, and complex health benefits where one does pricing based on entire distribution of losses. Any suggestion on which premium principle (from Table 7.1) is more suitable to be used for any of these two insurance products?

### Detailed Comments

1. Referring to eqn (7.1), ‘loss’ is random and ‘expected cost’ is used to determine rates. Why do you use ‘expected cost’ instead of ‘expected loss’?
2. Under section 7.2.1 (pure premium method), Pure premium = (claim count/exposure) x (loss/claim count) = frequency x severity. To avoid confusion, this formula should be related to the premium formula in eqn (7.1).
3. Illustration on Table 7.5 ends with a nice note, mentioning that one might combine the cell ‘AOI=low & Terr=1’ with other cells, and this will be further explained in Chapter 8 (selecting variables) and Chapter 9 (credibility). Perhaps it should be mentioned that the process of partitioning or combining cells needs cautious judgments. Appropriate balance between homogeneity (by partitioning the cells) and volume of data (by combining the cells for statistical reliability) should be maintained.
4. Figure 7.1 (timeline of exposure) should be larger. I have to ‘zoom’ my computer screen to read the ‘calendar time -- 1 Jan 2019, 1 Jan 2020, 1 Jan 2021’ in the figure.
5. Figure 7.2 (timeline of claim development) also should be larger and clearer.
6. In the beginning of Section 7.5, it is mentioned that the readers should learn how to describe development of a claim over several payments and link that to various unpaid claim measures (including IBNR & case reserves). The link should be clear then, a graphical representation would be nice. Perhaps it can also be mentioned that further explanation is given in Chapter 11 (Loss Reserving).
7. In the beginning of Section 7.6, it is mentioned that the readers should learn how to define concentration curve & corresponding Gini statistic and to use concentration curve & Gini statistic for premium selection. However, in section 7.6.2 the discussion is on performance curve (not concentration curve).
8. Section 7.4 on “Heterogeneous Risks” could also well go into the Risk Classification chapter 8. The Chapters 7 and 8 authors should coordinate to ensure appropriate synergies/cross-referencing. One way of handling heterogeneity is via regression, a topic utilized in Chapter 7.
   * In Eqn (7.1) --> Premium = loss + expense + UW Profit. Section 8.1 also mentions this formula, which is slightly in different terms --> Gross premium = expected loss + expected expenses + profit. To avoid confusion, the formula should be consistent throughout.

## Chapter 8 - Major Comments

This is a great chapter but the approach is not quite aligned with the rest of the book. We would like to see a chapter that deals more with classification and less with Poisson regression. The rest of the book does not presume knowledge nor use knowledge of regression and so its inclusion is a bit out of place here.

**A. How to deal with GLM**

1. We recommend retaining GLM/Poisson regression but with few technical details. When GLM is utilized, we recommend a much more data-driven approach. Once readers of this book understand the basics of risk classification, it is expected that they will be able to write GLM code in R without much knowledge of GLM theory.

The theory of Poisson regression and more general GLM theory can be learned if needed from many other existing textbooks (or later in a new chapter or appendix to be added in this book), it is not necessary to fully describe GLM in the middle of this book.

2. A few comments on *why* actuaries use GLM would be nice. A short historical perspective on minimum bias methods culminating in their relationship with GLMs given by Mildenhall (1999) would be helpful. One source is Frees, Derrig, Meyers (2014), “[Predictive modeling applications in actuarial science](https://books.google.com/books?hl=en&lr=&id=QHrsAwAAQBAJ&oi=fnd&pg=PR13&dq=info:bptBrUIaIPgJ:scholar.google.com&ots=RjKgAKbx0B&sig=9_LfPec-M6opNZjLuIeEWVXPibw)”, page 7-8.

**B. How to Include Classification Ideas**

For a resource that presents ideas of classification, consider "The Discriminating (Pricing) Actuary" (<https://www.tandfonline.com/doi/full/10.1080/10920277.2021.1951296?src>=). This is a survey article on discrimination that we can think of as on classification but with some ethics thrown in. For chapter 8, the ethics are not necessary (although a few comments here and there might be of interest to readers). For classification, you might consider excerpts from:

1.1 How Insurers Discriminate

1.2 Prohibitions

2.1 What is Actuarial Fairness?

2.2. Characteristics of Sensitive Variables (Currently in Appendix TS8.B)

3.1. Adverse Selection, Moral Hazard, and Incentives (Some of this is in the current chapter 8)

3.3. Price Discrimination

4. Actuarial aspects of Rate Regulation (currently in Appendix TS 7.A)

5.5. Big Data (we might use this in Chapter 13)

6.1. Strategies for Mitigating Proxy Discrimination

### Detailed Comments

1. Thus, ignoring frictional expenses associated with the administrative expenses and ~~the~~ profit

2. lifestyle is one word

3. For example, consider a health insurance ~~industry~~ where smoking status is an important risk factor for mortality and morbidity.

4. Eventually leading to a collapse of business -> eventually leading to business collapse

5. There should be a link to Technical Supplement TS 8.B

6. The link directly above 8.2 says “Show Quiz Solution” but the entire question and solution are hidden.

7. Quiz questions 2, 3, and 5 say which of the following, but then there are no options following.

8. as it is easier to work with in estimating process, similar to (8.12). -> as it is easier to estimate, similar to (8.12)

9. This ~~clearly~~ shows that the Poisson rate parameter λ varies across different tariff cells, with the same log linear form used in a Poisson regression framework.

10. The policyholder age variable only applies to type A (automobile) vehicle, and there is are no ~~policy~~ policies in the first age band.

11. The missing relativity may be estimated by some interpolation or the professional judgement of the actuary. (I think it would be extrapolation, not interpolation)

12. The adverse selection example is health insurance, but then the actual example discusses life insurance and mortality. I would just change it to a life insurance company.

13. Section 7.4 on “Heterogeneous Risks” could also well go into the Risk Classification chapter 8. The Chapters 7 and 8 authors should coordinate to ensure appropriate synergies/cross-referencing.

## Chapter 9 - Major Comments

This is a great chapter that was originally written as a stand-alone presentation. By relying on results from other chapters, this chapter can be shortened and improved. Moreover, readers would benefit from seeing more data-driven examples.

* Section 9.2 on “Limited Fluctuation Credibility” is probably not a good way to start the chapter. Although one can make a case that it was first historically, it is not a method that is used in major parts of the world (e.g., Europe). So, consider making this the last section of the chapter so that it can be skipped by readers not interested in this approach.
* Section 9.5 on "Bayesian Inference and Bühlmann Credibility" needs to be tightly integrated with the new Bayes chapter. As it stands, Section 9.5 emphasizes situations with closed-form predictors. Do we want to restrict ourselves in this way? The second edition of Loss Data Analytics will have an entire chapter devoted to Bayesian Inference. We should be able to incorporate many of the ideas of Bayesian inference in this chapter.
* Section 9.6 is on "Estimating Credibility Parameters". As part of a new chapter on random effects models, we intend to introduce prediction of random effects. This addition will allow readers to see many of the mathematical underpinnings of credibility models as well as prediction of random effects, the topic of this section. A challenge for all authors will be to integrate these ideas across chapters into a single coherent presentation.
* The current Sections 9.5 and 9.6 will be greatly reduced in size. As additions, we hope to see a major data-driven example with supporting *R* code.

### Detailed Comments

1. Sections 9.1 and 9.2 present some classic approaches to credibility. These sections would benefit from using the "show/hide" features. As an example of relying on other results, Section 9.2.2 should link to Chapter 5 (and shorten the presentation).
2. Section 9.3 introduces the Buhlmann model.
   1. As another example of relying on other results, we need to integrate 9.3 with the iterated expectations presentation in Section 16.2. For sign-posting, we should identify it as a (statistical) random effects model.
   2. A nice section. Personally, I would not make such a big deal about those classic terms such as "Expected Value of the Process Variance" and "Variance of the Hypothetical Means" but I am happy to let the author decide.
3. Section 9.4 introduces "Bühlmann-Straub Credibility." This is a reasonable topic at least from a historical development perspective.
   1. For current applications, I think the main reason it is useful because it provides one example as to how to modify credibility formulas for non-iid situations. References to other extensions would be helpful.
   2. This section in particular could be shortened using a "show/hide" approach.
4. Need to tie this chapter to Chapter 12 on "Experience Rating using Bonus-Malus".
5. One person on the committee reported that:

It is a very beautiful presentation. But the way of describing Buhlmann’s approach is rather uncomfortable to me. I would like to see more respect for original ideas, such as what Buhlmann himself wrote in his co-authored book.

## Chapter 10 - Major Comments

* Do what you can to promote a unified theme among the different topics in Chapter 10.
  + Maybe a specific example can be used to give the motivation for this chapter.
  + For example, the 1998 rain example in 10.2 can be taken to section 10.1 and extended to cover other topics in the chapter?
* A data-driven example would be good in Section 10.2.1 (e.g. fitting some (heavy-tailed) data with Weibull, Gamma and Pareto distns).
* Try to emphasize the links to Chapter 5 on Aggregate Models. For example, should the mean and variance of compound Poisson, compound Binomial and compound negative binomial (referring back to Ch5) be derived for both the insurer and reinsurer after the application of proportional/non-proportional reinsurance?
* The Section 10.3.1 introduction to coherent risk measures should be retained. However, it would be improved with a further discussion of “why” it is important to know about this type of risk measure. One thought would be to give examples that demonstrate situations where coherence axioms fail.

### Detailed Comments

* Leave “the conference room” out in 10.1?
* Use heavy-tailed distn instead of heavy-tail distn.
* Space between “variable possesses” in 10.2.1
* In Section 10.2.2, consider comparing tails based on hazard rates
* In 10.3 should the variance of return, downside semi-variance of return also be introduced?
  + Also maybe the relation between the investor’s utility function and the risk measures can be discussed.
* In 10.3.1, there is a part saying “Special Case. The Standard Deviation Principle (10.1) is a Coherent Risk Measure”. But that part is, I feel, highly misleading.
  + The Standard Deviation Principle is not a Coherent Risk Measure in the normal sense. So, it’s at least misleading even though I can imagine what is intended to be shown there.
  + More importantly in my opinion, the reader should learn that Coherence holds only when all four conditions are satisfied for “any pair” of random variables X and Y. So, I believe that that the example is not appropriate (if not simply wrong).
* Terminology
  + Please reconsider the terminology adopted here.
  + At least in my preference, TVaR appeared in the current version should be called CTE as in Hardy (2006), <https://www.soa.org/globalassets/assets/files/edu/c-25-07.pdf>, and CVaR should be called TVaR as in the *Loss Models*, especially when you actually refer to Hardy (2006) and the *Loss Models*.
  + This suggested terminology is consistent with, e.g., Denuit et al., Actuarial Theory for Dependent Risks, one of the most comprehensive texts in the field, in which CVaR is differently defined.
* Quota Share proposition
  + In 10.4.1.1, this proposition is not well explained. In the current version, the LHS and RHS of the inequality are mismatched (as LHS relates to the reinsurer and RHS relates to the insurer), “K” suddenly appears without definition, and the last equality seems to be inappropriate.
  + An alternative is as follows:
    - The inequality: Var((1-c)X) ≤ Var(X-g(X))
    - The second equality: E[g(X)] = cE[X]
    - The last equality: c^2=Q/Var(X)
  + In addition, mathematically speaking, to change the first equality to Var(Y\_insurer) = Q ≤ Var(X) makes it more rigorous.
  + It is because, not practically but theoretically, a stop-loss “retention”, in which case g(X) is of max(X-a, 0), can be more desirable to the reinsurer!
* Surplus Share Proportional Treaty
  + In the last sentence of 10.4.3.1, “if X is the loss” must be “if X is the amount insured”. Right?
* I found another misleading sentence in the last paragraph of 10.3.3:
  + “Third and perhaps foremost, TVaR is a coherent risk measure and thus is able to more accurately capture the diversification effects of insurance portfolio. Herein, we do not intend to provide the proof of the coherent feature for TVaR, which is considered to be challenging technically.”
  + The first sentence is misleading. Here, TVaR is the conditional tail expectation, and, generally, it’s not called a coherent risk measure. If you read the previous paragraph very carefully, you could positively interpret that this sentence is written only for the case in which it is assumed every random variable is continuous, and then the sentence would be OK. But, readers are not so careful.
  + CVaR in the same section, which is the average value at risk, is coherent. So, if you would provide “the proof of the coherent feature for…”, it would be for CVaR. Then, it’s true that the proof would be rather technical, but I don’t think it is “considered to be challenging technically” especially when you assume every random variable is continuous.

## Chapter 11 - Major Comments

Even though incomplete, this is still a superb chapter. The first three sections are well-done, providing ample motivation for the “why” part as well as a detailed discussion of the data.

* For actuaries with a practical bent, we recommend including in Section 12.3 a short treatment of the Bornhuetter-Ferguson method.
* Section 11.4.1 on GLMs and Loss Reserving will be valuable when completed. However, except for Chapters 8 and Section 12.5, we don't have GLMs in other parts of the text. So, the treatment will need to be gentle.
* Section 11.4.3 on bootstrapping will be able to take advantage of the introduction to bootstrapping that is in Section 6.2. Will need these two pieces to reinforce one another.

### Detailed Comments

* Section 11.3.2 states "Since the chain-ladder method is a purely deterministic and intuitively natural algorithm to complete a run-off triangle, we are not able to determine how reliable that point estimator is or to model the variation of the future payments. To answer such questions an underlying stochastic model that reproduces the chain-ladder reserve estimates is needed." This statement will confuse students whom have been introduced to machine learning techniques for predictions that are simply algorithms, often with no probabilistic basis. Better to re-phrase slightly.
* The chapter contains several small data sets. We should be able to make them readily available via a OneDrive link.
* Section 11.3.2. It may help students to refer to Appendix Section 16.2 on iterated expectations.
* Line 15 of Section 11.3.2. For the variance, would it be cleaner to condition on the entire history (not just the current value of the process) such as D\_I? This is what is done below in the Section when developing the variance estimator.
  1. **Data**

Run-off triangles

* A run-off triangle with incremental payments is displayed in Figure 11.6, the data are taken from Wüthrich and Merz (2008). Figure 11.7 shows the same information, but in cumulative format. There are several typos in the Figures. As examples, for accident year 2004 (i=1), the cumulative payments in years 4 until 9 (j=4,5,6,7,8,9) are:

10978 11041 11106 11121 11132 11148

so that the incremental payments in years 5 until 9 (j=5,6,7,8,9) are:

65 65 16 11 16

But the incremental payments in the same years (from Figure 11.6) are:

621 658 149 111 158

Loss reserve notation

* Loss reserve for accident year *i* is

A picture containing text, clock

Description automatically generated

To illustrate this, Figure 11.6 (incremental payments) can include a column on loss reserve for accident year *i*:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| accident year (i) | payment delay (j) | | | | | | | | | | | reserve | |
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |  | |
| 1 | 5,947.0 | 3,721.2 | 895.7 | 207.8 | 206.7 | 621.2 | 658.1 | 148.5 | 111.3 | 158.1 | - | |
| 2 | 6,346.8 | 3,246.4 | 723.2 | 151.8 | 678.2 | 366.0 | 527.3 | 111.9 | 116.3 | X29 | X29 | |
| 3 | 6,269.1 | 2,976.2 | 8470.5 | 262.8 | 152.7 | 654.4 | 535.5 | 892.4 | X38 | X39 | X38+X39 | |
| 4 | 5,863.0 | 2,683.2 | 722.5 | 190.7 | 133.0 | 883.4 | 433.3 | X47 | X48 | X49 | X47+X48+X49 | |
| 5 | 3,778.9 | 2,745.2 | 653.9 | 273.4 | 230.3 | 105.2 | X56 | X57 | X58 | X59 | X56+...+X59 | |
| 6 | 6,184.8 | 2,828.3 | 572.8 | 244.0 | 105.0 | X65 | X66 | X67 | X68 | X69 | X65+...+X69 | |
| 7 | 5,600.2 | 2,893.2 | 563.1 | 225.5 | X74 | X75 | X76 | X77 | X78 | X79 | X74+...+X79 | |
| 8 | 5,288.1 | 2,440.1 | 528.0 | X83 | X84 | X85 | X86 | X87 | X88 | X89 | X83+...+X89 | |
| 9 | 5,290.8 | 2,357.9 | X92 | X93 | X94 | X95 | X96 | X97 | X98 | X99 | X92+...+X99 | |
| 10 | 5,675.0 | X101 | X102 | X103 | X104 | X105 | X106 | X107 | X108 | X109 | X101+...+X109 | |

Similarly, Figure 11.7 (cumulative payments) can include a column on loss reserve for accident year *i*:

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| accident year (i) | payment delay (j) | | | | | | | | | | reserve |
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 1 | 5,947 | 9,668 | 10,564 | 10,772 | 10,978 | 11,041 | 11,106 | 11,121 | 11,132 | 11,148 | - |
| 2 | 6,347 | 9,593 | 10,316 | 10,468 | 10,536 | 10,573 | 10,625 | 10,637 | 10,648 | C29 | C29-C28 |
| 3 | 6,269 | 9,245 | 10,092 | 10,355 | 10,508 | 10,573 | 10,627 | 10,636 | C38 | C39 | C39-C37 |
| 4 | 5,863 | 8,546 | 9,269 | 9,459 | 9,592 | 9,681 | 9,724 | C47 | C48 | C49 | C49-C46 |
| 5 | 5,779 | 8,524 | 9,178 | 9,451 | 9,682 | 9,787 | C56 | C57 | C58 | C59 | C59-C55 |
| 6 | 6,185 | 9,013 | 9,586 | 9,831 | 9,936 | C65 | C66 | C67 | C68 | C69 | C69-C64 |
| 7 | 5,600 | 8,493 | 9,057 | 9,282 | C74 | C75 | C76 | C77 | C78 | C79 | C79-C73 |
| 8 | 5,288 | 7,728 | 8,256 | C83 | C84 | C85 | C86 | C87 | C88 | C89 | C89-C82 |
| 9 | 5,291 | 7,649 | C92 | C93 | C94 | C95 | C96 | C97 | C98 | C99 | C99-C91 |
| 10 | 5,676 | C101 | C102 | C103 | C104 | C105 | C106 | C107 | C108 | C109 | C109-C100 |

* 1. **Chain-Ladder**
* The development factor *fj* is:

A picture containing schematic

Description automatically generated

It should be mentioned that the formula *fj* is for *j*=1,…,*J*-1 (in this example 🡪 *j*=1,…,8)

* There is a typo in the calculation of *f0*:

Diagram

Description automatically generated

Using the values in Figure 11.7, *f0* = (9668+….+7649)/(5947+….+5291) = 1.5515, not 1.4925

* The variance parameter is:

Text

Description automatically generated

It should be mentioned that the formula *j2* is for *j*=1,…,*J-2* (in this example 🡪 *j*=1,…,7)

To find *J-12*, in this example *82*, the formula suggested by Mack’s model is:

* Figure 11.7 (cumulative payments) can also include the rows for development factor *fj* and variance parameter *j2*

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| accident year (i) | payment delay (j) | | | | | | | | | | reserve |
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 1 | 5,947 | 9,668 | 10,564 | 10,772 | 10,978 | 11,041 | 11,106 | 11,121 | 11,132 | 11,148 | - |
| 2 | 6,347 | 9,593 | 10,316 | 10,468 | 10,536 | 10,573 | 10,625 | 10,637 | 10,648 | C29 | C29-C28 |
| 3 | 6,269 | 9,245 | 10,092 | 10,355 | 10,508 | 10,573 | 10,627 | 10,636 | C38 | C39 | C39-C37 |
| 4 | 5,863 | 8,546 | 9,269 | 9,459 | 9,592 | 9,681 | 9,724 | C47 | C48 | C49 | C49-C46 |
| 5 | 5,779 | 8,524 | 9,178 | 9,451 | 9,682 | 9,787 | C56 | C57 | C58 | C59 | C59-C55 |
| 6 | 6,185 | 9,013 | 9,586 | 9,831 | 9,936 | C65 | C66 | C67 | C68 | C69 | C69-C64 |
| 7 | 5,600 | 8,493 | 9,057 | 9,282 | C74 | C75 | C76 | C77 | C78 | C79 | C79-C73 |
| 8 | 5,288 | 7,728 | 8,256 | C83 | C84 | C85 | C86 | C87 | C88 | C89 | C89-C82 |
| 9 | 5,291 | 7,649 | C92 | C93 | C94 | C95 | C96 | C97 | C98 | C99 | C99-C91 |
| 10 | 5,676 | C101 | C102 | C103 | C104 | C105 | C106 | C107 | C108 | C109 | C109-C100 |
| Developmt factor fj | f0 | f1 | f2 | f3 | f4 | f5 | f6 | f7 | f8 | - |  |
| Variance parameter j2 | 02 | 12 | 22 | 32 | 42 | 52 | 62 | 72 | 82 | - |  |

* Since the examples on the calculation of development factor fj are shown, examples on the calculation of variance parameter j2 can also be shown. As examples, using the values from Figure 11.14:

Table

Description automatically generated

* R coding for the MSEP per occurrence (MSEPCij) and the MSEP for total reserve (MSEPsum(Cij)) are shown, nice work here.

**11.4 GLMs & Bootstrap**

Model Specification

* Under this section, three GLMs widely used for loss reserving are discussed – Poisson, overdispersed Poisson & Gamma regressions. The author mentions that typical GLMs work with incremental run-off triangles. The GLM treats the occurrence year (i) and payment delay (j) as factor variables. It would be nice if the data for this format is shown. It can easily be done using R. As an example:

Text

Description automatically generated

:

Table

Description automatically generated

Model Estimation and Prediction

* Under this section, a point estimate in the upper triangle is:



and a cell in the lower triangle is predicted as:



Again, it would be nice if an example of R coding is shown. As an example, we can use over-dispersed Poisson using R:

Table

Description automatically generated

Text

Description automatically generated

To predict the ‘lower triangular data’:

Text

Description automatically generated

Table

Description automatically generated

:

## Chapter 12 – Major Comments

This is a terrific topic and provides background in an important area not available in other introductions to Loss Models. However, consistent with other chapters, I suggest re-organizing the material to make it accessible to a broader readership.

1. In the comments below, there are several requests to develop the “why” part. Especially readers who do not work in jurisdictions that utilize BMS, additional motivation is needed to motivate the study of this topic.
2. Markov chains, transition probabilities, and concepts of stationarity are not employed extensively in the rest of the book. So, we need to segregate introduction of these ideas from BMS applications.
   1. As it stands, the technical level of this chapter is much higher than other parts of the textbook. We should make an effort to organize the material so that readers can about BMS without wading through all the technical material.
3. We are not sure we understand the aim of Section 12.5. It seems to be quite disconnected from the other subsections in this chapter, and needs a more detailed introduction.

### Detailed Comments

**General:**

* Hunger for bonus is mentioned only in the introduction section. It would be good the analyse the effect of this on the BMSs technically.
* I think sticking to one terminology for the subsection titles would be good (either BMS or NCD).
* Currently, this chapter deals only with closed populations. But I think that open populations may be of more interest in practice. In any case, I think, perhaps optimistically, a simulation of changes in distribution is not so hard for students, even if the theory of, say, Markov chain is taken as advanced.

**Section 12.1:**

* Make sure year=policy year. E.g …if they do not have any claims during the policy term.
* Section 12.1 is well done. I think it would benefit with a comparison with other experience rating schemes, using Chapter 9 as a starting point. Some of the thoughts from Section 12.5.1 on Premium Rating may be brought in here or cross-referenced. That is, you could take some of the Section 12.5.1 ideas and say them in simpler language here and the retained the more precise (yet technical) approach of Section 12.5.1.

Section 12.2.1:

* I’m not sure what “country’s economic level” means here.
* Refers to Table 12.1 twice.
* In Table 12.1 I don’t think class is necessarily the same as claim-free years in other systems. Might be confusing. E.g. in Table 12.2 this might not be the case..
* The Section 12.2 description of NCD System in Several Countries is well-done. However, I think starting the section with broader statements about the use of BMS/NCD throughout the world would be helpful. A quick internet search shows a 2010 paper of https://link.springer.com/article/10.1057/gpp.2010.37 suggest that BMS is used in most Asian countries. This suggests starting Section 12.2 with a statement like, "bonus-malus is used world-wide including important economies such as Malaysia, Switzerland, France, Brazil, and ... These economies represent" something like millions of people, or X% of world GDP... Is this still true?
* Here is another approach: This subsection presents BMSs for three countries with different characteristics, which is good. I suggest to give all these three systems under one subsection, with a shorter explanation for Malaysian system. Also explain why only these three systems are presented at the beginning of the subsection and refer the reader to a source for other BMSs around the world (e.g. Lemaire, 1995).

**Section 12.2.2:**

* My other comment is that although Switzerland is given as one of the examples in 12.2, it has not been used in the later examples. I think as one of the most complicated BMSs, it would be interesting to see how it differs from “simpler” systems. E.g. in Section 12.4.4 I believe it will take much longer to reach stationarity under this system. Related to this, how can we compare different BMSs? E.g. BMS for Switzerland seems “tougher” than the BMS for Brazil. Is it possible to compare these formally?
* For the Switzerland NCD I found the following sentence confusing: “These levels are equivalent to … and the following discounts: …”
* Table 12.3 Table heading in both columns Loading/Discount?
* I think it would be clearer to mention the entry level explicitly (i.e. entry level is class 12, which corresponds to 100% prem level …)

**Section 12.3.1:**

* The Markov model could be discussed in more detail
  + An option could be to include a basic introduction to these concepts in chapter 12 using simple worked examples prior to the more detailed and complicated BMS examples. For example, simple examples might use 2x2 transition matrices for illustration of the concepts of transition probabilities and stationary probabilities.
  + One approach is to have Section 12.3.1 provide definitions of Markov Chains and transition probabilities (much as it is now). Then, make Section 12.3.2 applications of 12.3.1 probability ideas to BMS. This is not a lot different than it is now, just a little extra sign-posting.
  + Another approach is to introduce a Technical Supplement that provides the definitions for Markov chains. In this way, we can also easily refer to its application in Mack's model for Loss Reserves, Section 11.3.2, as well as the MCMC in the new Bayesian chapter.
  + Might also dig into, e.g. assumptions, conditions for unique stationary distn, etc. Is the current theoretical explanation in this chapter already a bit too much, or too advanced? Or is it still not enough? Think about the main body as material every reader should know, use Technical Supplements for extra fun stuff.
* I would rephrase the following sentence: “All probabilities must also be non-negative (since they are probabilities)”
* Explain what Poisson with lambda=0.10 means.

**Section 12.4:**

The structure of Section 12.4 is useful. However, it would be really helpful to start the section saying **why** it is important to have stationary transition probabilities. After all, in an economy with an evolving population, the demographics may be such that a stationary probabilities are not all that useful. Just a little motivation to get the reader started.

**Section 12.5.2**: What is the relation between the given equation and the number of claims, N? Define N\_i, i.e. discrete r.v? I think the model should be given explicitly here.

**Section 12.5.4:** The final equation should be explained further (e.g. what is f(.) here, how is this different than no heterogeneity case, etc).

**Section 12.5.5:**

* What does optimal relativity measure? Why do we need this?
* “Lambda hat is the constant expected…”: There is no lambda hat in the first equation. Should this be lambda bar?
* Re-write to employ the show/hide the math development.
* Issues of convergence rates in Section 12.4.5/6 are interesting but might be better placed in technical supplement. Regardless of where this Section winds up, we need more motivation as to why this is important to learn.

**Section 12.5.6**

* “Also, since the obtained form of stationary probabilities are rather complex, in this section we choose not to include any R codes for the determination of optimal relativities.” I would still prefer to give the code with show/hide feature.
* How can one interpret constrained and unconstrained optimal relativities? What is the difference between them?
* The final paragraph (“Note that the obtained values of optimal relativities…”) needs more explanation.

## Chapter 14 - Major Comments

* Section 14.1 is to be moved to the (currently numbered) Chapter 13, possibly as a Technical Supplement to that Chapter.
* We recommend splitting Section 14.2 into two parts. The first part will retain discussion of Pearson and Spearman correlations as well as Kendall’s tau. The other portion, currently Section 14.2.3-14.2.6, will be placed in a new section at the back of the chapter (after the copula material).
  + Considering introducing a measure of tail dependency in the first part. This is currently done in Section 14.5.4.3; however, it is another measure of variable association albeit one where we are interested in the relationship at extreme values. The definition of upper and lower tail dependency in terms of copulas could then still be included later in the chapter.
* Most of the material is the rest of the chapter is useful yet it requires serious re-organization (see the following detailed comments). The one new piece that needs to be added is on the **simulation of dependent distributions**.

### Detailed Comments

**Section 14.3. Introduction to Copulas**

The main recommendation is to provide more mathematical underpinnings for the introduction to copulas. Here is one plan of attack:

* A. Start with the definition of a copula.
  + Use Frank as an example remind readers of what are the conditions needed for a (multivariate) function to be a distribution function. Comment that Frank was chosen (and often used) because it is a simple function that is easy to interpret.
  + Move up Figure 14.6 to help students visualize a copula.
  + Include some problems (or maybe simple written examples using two or more different copulas (e.g., the independence copula and the Frank copula with a specific parameter value) may be useful. Attached are some tutorial questions/solutions that have been used in class for example of possible question style.
* B. Give Sklar's theorem (in both directions). Provide short proofs, using the hide/show features.
  + Even at this stage, we can emphasize the advantage of modeling separately marginals and dependence structure with the copula.
* Consider mentioning the idea of a copula density here although the development can wait until later. Maybe density of Frank's copula can be an exercise.

**Section 14.4 Application Using Copulas**

This section is well-done. The simulation of copulas is built into 14.4.3. Because simulation is so important to actuarial science, we need to give this procedure again in a later section, in greater generality.

**Section 14.5 Types of Copulas**

* We should remove "Properties" and make it a separate subsection, replete with more mathematical details (with hide/show).
* We should mention all the other types of Archimedean copulas, possibly in a reference section.
* For the rest of the chapter, it would be nice to compare the choices of copulas in some fashion. One approach is to revert to the example in Section 4.4 and calculation something using the fitted Frank copula. For example, we might calculate the expected level of expenses given that losses will exceed a given amount. Then, repeat this calculation for the other copulas described in Section 14.5.

**(New) Section 14.6 Properties of Copulas**

* In the (current) Section 14.5.4.1 Bounds on Association, a discussion of co- and anti-monotonicity is needed. A few references on the importance of this topic (possibly in the related References section) would also be valuable.
* In the (current) Section 14.5.4.3 Tail Dependency, we can refer to earlier work on the introduction of these measures, without regard to the distribution.

**14.6 Why is Dependence Modeling Important?**

* Maybe split this section into two parts. The first part would retain the current portfolio example. The second would be a wide-ranging (qualitative) discussion of potential copula applications.
* Could discuss dangers of application of copulas without understanding their limitations and properties. (e.g., GFC and Gaussian copula.)
  + Mention that there are many other ways of conducting dependence modeling – copulas work well for many actuarial applications.
* Reinforce simulation with copulas idea.
* Mention how copulas can be used for regression applications (Emphasize the diagnostic purposes in subsection 14.4.3 for the marginals in the example).

**Other Comments**

Given the title of the chapter, I was expecting dependence concepts at the beginning of the chapter, such as Fréchet Class, bounds, comonotonicity, antimotonicity. So, you might considering changing the title to something like "Quantifying Dependence".

14.4.3: Transforming the data to check against a normal distribution may be a distraction from the primary purpose of modelling dependence. Perhaps hide this as a link?

**- Small typos:**

o For the discrete case (p.458)

o Here γ is **a/the** dependence parameter

o …this statistic depends…(omit the “s” in statistic, page 461)

o …Clayton copula is parameterized by … is defined by… (page 467)

## Appendices - Major Comments

• The objective of these supplemental chapters need to be clearly specified. As it stands, Chapter 15 Appendix A is more of a crash course in statistical inference than an appendix. The same is true for Chapters 16 and 17. For example, for Chapter 15 we might say something like "Chapter 15 provides an outline of topics that one would encounter in an introductory course on statistical inference. This chapter exists to develop a set of notations and definitions that are consistent with loss data analytics. Moreover, this chapter helps readers to appreciate the different points of emphasis between general statistical inference and those concepts employed in loss data analytics."

• We need to critically evaluate the supporting information provided in the Appendices and how relevant it is to the content of the book. Authors should ask themselves whether (1) it is all necessary and essential, (2) appropriately cited in the chapters of the book, and (3) if this complementary material that expands the reader’s understanding beyond what he/she read in the book chapters, or does it provide requisite knowledge that is needed to understand the content of the book chapters.

• There is some redundancy in these appendices. For example, maximum likelihood estimation is introduced in Section 15.2.2 of Appendix A. The likelihood function and log likelihood function are introduced in this section and introduced again in more depth in Appendix C. More cross-referencing and/or elimination of unnecessary material may be warranted.

## Detailed Comments

• Each Appendix has the title Chapter X, Appendix Y followed by the topic. I think it should either be named a Chapter or an Appendix but not both.

• Section 15.1 introduces statistical inference as the process of making conclusions on the characteristics of the population using a random sample. The Wisconsin Property Fund example provided in this section uses the entire population of individual claims from the 2010 experience (1,377 claims) as a random sample that is representative of the population of the 2010, 2011 claims and so forth. I was not very comfortable with this setup especially if this Appendix is intended for readers with no requisite knowledge about sampling techniques. This example only presents descriptive statistics of the 2010 experience (which is in fact the 2010 population) with no reference to inference.

• Section 15.1.1 starts with a statement that in my opinion is not quite accurate "In statistics, a sampling error occurs when the sampling frame, the list from which the sample is drawn, is not an adequate approximation of the population of interest." Again a reader with no requisite background about sampling, what a sampling frame is, and what is sampling error may misunderstand this statement and believe that sampling error only occurs as a result of poor choice of the sampling frame which is not the case.