ACTEX Learning

Study Manual for Exam MAS-II

9th Edition

Jeffrey S. Pai, PhD, ASA, ACIA
Alisa Walch, FCAS, MA
Hong Li, PhD, FSA, ACIA
Ambrose Lo, PhD, FSA, CERA

A CAS Exam

ACTEX Learning

Study Manual for Exam MAS-II

9th Edition

Jeffrey S. Pai, PhD, ASA,
ACIA Alisa Walch, FCAS, MA
Hong Li, PhD, FSA, ACIA
Ambrose Lo, PhD, FSA, CERA

ACTEX Learning

Learn Today. Lead Tomorrow.

Actuarial & Financial Risk Resource Materials

Since 1972

Welcome to Actuarial University

Actuarial University is a reimagined platform built around a more simplified way to study. It combines all the products you use to study into one interactive learning center.

You can find integrated topics using this network icon.

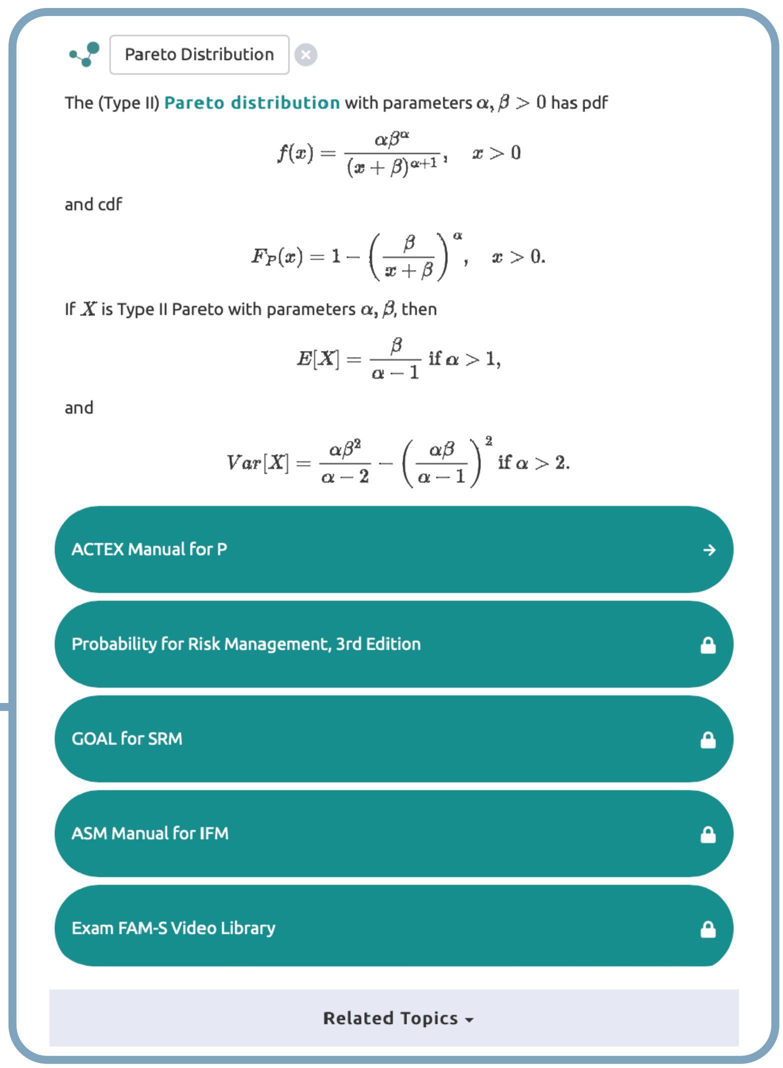
When this icon appears, it will be next to an important topic in the manual. Click the **link** in your digital manual, or search the underlined topic in your print manual.

1. Login to: www.actuarialuniversity.com

2. Locate the **Topic Search** on your exam dashboard and enter the word or phrase into the search field, selecting the best match.

3. A topic "**Hub**" will display a list of integrated products that offer more ways to study the material.

4. Here is an example of the topic **Pareto Distribution:**



Within the **Hub** there will be unlocked and locked products.

Unlocked Products are the products that you own.

ACTEX Manual for P →

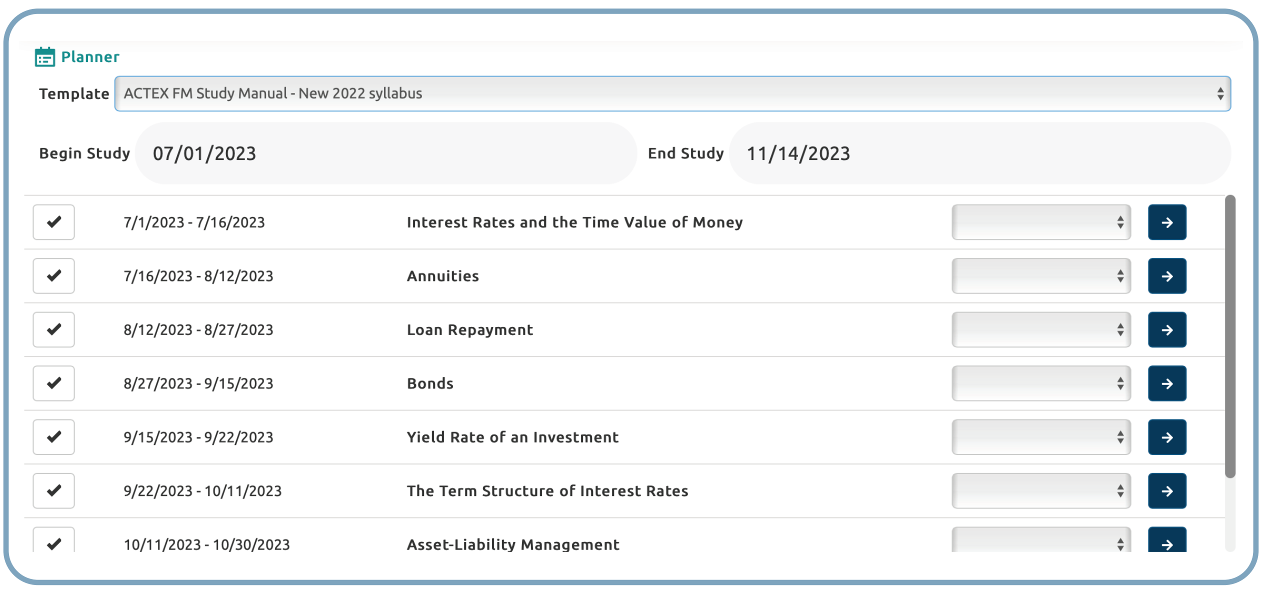
Locked Products are products that you do not own, and are available for purchase.

Probability for Risk Management, 3rd Edition

Many of Actuarial University's features are already unlocked with your study program, including:

Instructional Videos*	**Planner**
Topic Search	**Formula & Review Sheet**

Make your study session more efficient with our Planner!



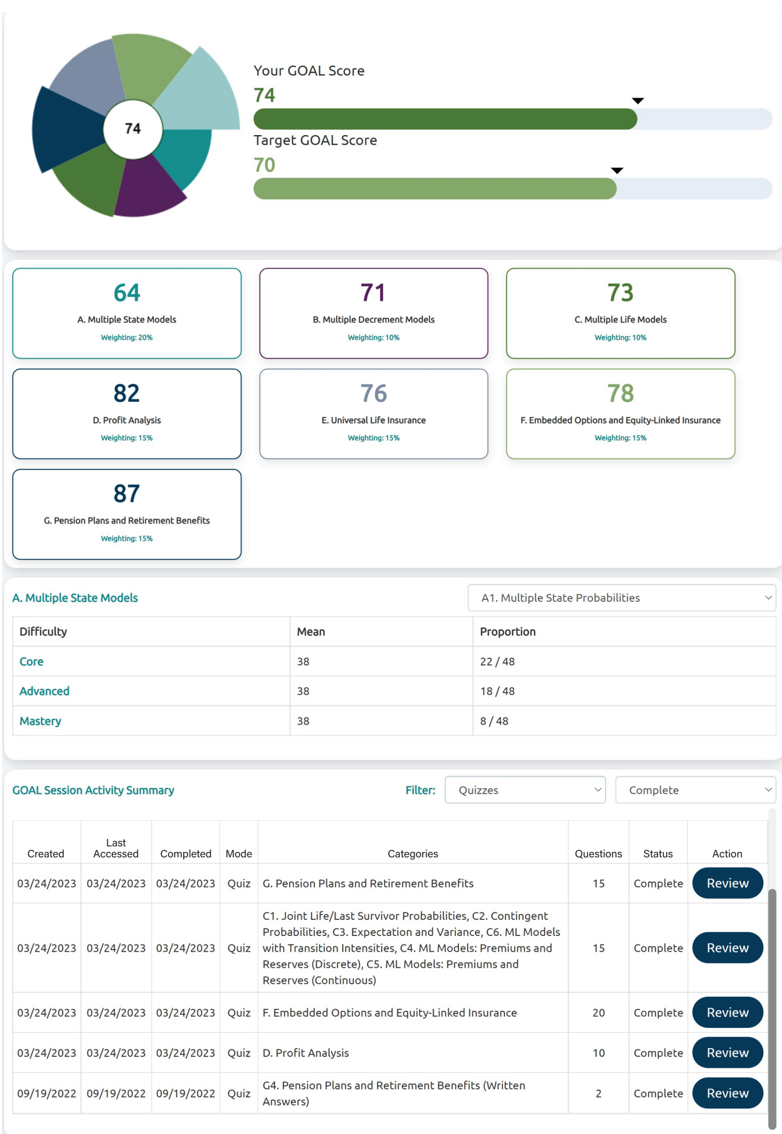
**Available standalone, or included with the Study Manual Program Video Bundle*

Track your exam readiness with GOAL Score!

Available for P, FM, FAM, FAM-L, FAM-S, ALTAM, ASTAM, MAS-I, MAS-II, & CAS 5

GOAL Score tracks your performance through GOAL Practice Sessions, Quizzes, and Exams, resulting in an aggregate weighted score that gauges your exam preparedness.

By measuring both your performance, and the consistency of your performance, **GOAL Score** produces a reliable number that will give you confidence in your preparation before you sit for your exam.



If your GOAL Score is a 70 or higher, you are well-prepared to sit for your exam!

See key areas where you can improve.

Detailed performance tracking.

Quickly return to previous sessions.

Contents

A Introduction to Credibility **1**

1 Limited Fluctuation Credibility **5**

1.1 Introduction 5
1.2 Full Credibility 7
1.3 Problem Set 22
1.4 Problem Set Solutions 29

2 Limited Fluctuation Credibility - Partial Credibility **37**

2.1 Problem Set 42
2.2 Problem Set Solutions 46

3 Bühlmann Credibility **51**

3.1 Problem Set 59
3.2 Problem Set Solutions 67

4 Bühlmann-Straub Credibility **77**

4.1 Problem Set 85
4.2 Problem Set Solutions 89

5 Bühlmann Credibility - Conjugate Pairs **93**

5.1 Problem Set 96
5.2 Problem Set Solutions 100

6 Bayesian Credibility - Introduction **105**

6.1 Problem Set 113
6.2 Problem Set Solutions 116

7 Bayesian Credibility - Conjugate Distribution **121**

7.1 Problem Set 138
7.2 Problem Set Solutions 142

8 Bayesian Credibility - Discrete Prior Distribution **145**

8.1 Problem Set 148
8.2 Problem Set Solutions 154

9 Non-parametric Empirical Bayes 163
9.1 Non-parametric Model in Bühlmann-Straub's Case . . . 163
9.2 Non-parametric Model in Bühlmann's Case . . . 167
9.3 Problem Set . . . 173
9.4 Problem Set Solutions . . . 178

B Linear Mixed Models 185

10 Introduction and Overview 189
10.1 Specification of LMMs . . . 189
10.2 General Matrix Specification . . . 190
10.3 Common Covariance Structures for Residuals . . . 191
10.4 Specification of the Marginal Model . . . 193
10.5 Marginal Model Implied by an LMM (aka Implied Marginal Model) . . . 193
10.6 Estimation . . . 194
10.7 Problem Set . . . 200
10.8 Problem Set Solutions . . . 212

11 Two-level Models for Clustered Data 219
11.1 Rat Pup Data . . . 219
11.2 Model Specification . . . 222
11.3 Hypothesis Tests and Model Selection . . . 227
11.4 Interpretation, Residual Diagnostics, and Prediction . . . 241
11.5 Summary . . . 248
11.6 Problem Set . . . 249
11.7 Problem Set Solutions . . . 258

12 Case Study: Two-Level Linear Mixed Models for Household Incomes 263
12.1 Introduction . . . 263
12.2 Data Description . . . 264
12.3 Exploratory Data Output . . . 265
12.4 Models and Fitting Results . . . 266
12.5 Hypothesis Tests and Analysis of Variance . . . 287
12.6 Model Diagnostics and Prediction . . . 292
12.7 Problem Set . . . 306
12.8 Problem Set Solutions . . . 308

13 Three-level Models for Clustered Data 313
13.1 Classroom Data . . . 314
13.2 Model Specification . . . 316
13.3 Hypothesis Tests and Model Selection . . . 320
13.4 Intraclass Correlation Coefficients . . . 326

13.5 Calculating Predicted Values 327
13.6 Summary 330
13.7 Problem Set 331
13.8 Problem Set Solutions 336

14 Models for Repeated-Measures Data: The Rat Brain Example 339
14.1 Rat Brain Data 339
14.2 Model Specification 341
14.3 Hypothesis Tests and Model Selection 343
14.4 Interpretation and Residual Diagnostics 346
14.5 Summary 349
14.6 Problem Set 350
14.7 Problem Set Solutions 354

15 Random Coefficient Models for Longitudinal Data: The Autism Example 359
15.1 Autism Data 359
15.2 Model Specification 362
15.3 Hypothesis Tests and Model Selection 363
15.4 Interpretation, Residual Diagnostics, and Predictions 367
15.5 Summary 373
15.6 Problem Set 374
15.7 Problem Set Solutions 376

16 Models for Clustered Longitudinal Data: The Dental Veneer Example 377
16.1 Veneer Data 378
16.2 Model Specification 378
16.3 Hypothesis Tests and Model Selection 379
16.4 Interpretation and Residual Diagnostics 382
16.5 Summary 386
16.6 Problem Set 387
16.7 Problem Set Solutions 391

17 Case Study: Using Linear Mixed Models for Clustered Longitudinal Data 393
17.1 Introduction 393
17.2 Data Description 394
17.3 Exploratory of data set 395
17.4 Models and Fitting Results 397
17.5 Hypothesis Tests and Analysis of Variance 416
17.6 Model Diagnostics and Prediction 418
17.7 Problem Set 457
17.8 Problem Set Solutions 459

18 Models for Data with Crossed Random Factors: The SAT Score Example 465

18.1 SAT Score Data . . . 466
18.2 Model Specification . . . 469
18.3 Hypothesis Tests and Model Selection . . . 470
18.4 Interpretation and Residual Diagnostics . . . 472
18.5 Summary . . . 477
18.6 Problem Set . . . 478
18.7 Problem Set Solutions . . . 481

19 Case Study: Using Linear Mixed Models with Crossed Random Factors 483
19.1 Introduction . . . 483
19.2 Data Description . . . 484
19.3 Exploratory Data Output . . . 485
19.4 Models and Fitting Results . . . 486
19.5 Hypothesis Tests and Analysis of Variance . . . 493
19.6 Model Diagnostics and Prediction . . . 494
19.7 Problem Set . . . 512
19.8 Problem Set Solutions . . . 514

C Statistical Learning 521

20 Assessing Model Accuracy 525
20.1 Measuring the Quality of Fit . . . 525
20.2 The Bias-Variance Trade-Off . . . 529
20.3 Classification Setting . . . 531

21 Model Validation and Selection 535
21.1 Assessing fit with plots of actual vs. predicted . . . 535
21.2 Measuring Lift . . . 536
21.3 Validation of Logistic Regression Models . . . 542
21.4 Assessing Model Accuracy with Classification Data . . . 544

22 Unsupervised Learning 549
22.1 Principal Components Analysis . . . 549
22.2 K-means Clustering . . . 555
22.3 Hierarchical Clustering . . . 559
22.4 Problem Set . . . 566
22.5 Problem Set Solutions . . . 575

23 Supervised Learning 581
23.1 K-Nearest Neighbors . . . 581
23.2 Regression Trees . . . 584
23.3 Classification Trees . . . 590

23.4 Advantages and Disadvantages of Trees . . . 596
23.5 Bagging, Random Forests, and Boosting . . . 597
23.6 Problem Set . . . 603
23.7 Problem Set Solutions . . . 614

24 Deep Learning Part I **621**
24.1 Single Layer Neural Networks . . . 621
24.2 Multilayer Neural Networks . . . 624
24.3 Convolutional Neural Networks . . . 626
24.4 Document Classification . . . 631

25 Deep Learning Part II **633**
25.1 Recurrent Neural Networks . . . 633
25.2 When to Use Deep Learning . . . 638
25.3 Fitting a Neural Network . . . 639
25.4 Interpolation and Double Descent . . . 643

26 Lab: Deep Learning **647**
26.1 A Single Layer Network on the Hitters Data . . . 647
26.2 Multilayer Network on the MNIST Digit Data . . . 649
26.3 Convolutional Neural Networks . . . 652
26.4 Using Pretrained CNN Models . . . 654
26.5 IMDb Document Classification . . . 654
26.6 Recurrent Neural Networks . . . 657

D Time Series with Constant Variance **663**

27 Basic Concepts in Time Series Analysis **667**
27.1 Modeling Trends and Seasonal Variations: A General Approach . . . 667
27.2 Modeling Deterministic Trends and Seasonal Variations: A Regression Approach . . . 676
27.3 Serial Correlations . . . 693
27.4 Problem Set . . . 721
27.5 Problem Set Solutions . . . 735

28 ARIMA Models **743**
28.1 White Noise and Random Walks . . . 744
28.2 Autoregressive Models . . . 750
28.3 Moving Average Models . . . 767
28.4 ARMA Models . . . 777
28.5 Non-stationary Models . . . 789
28.6 Problem Set . . . 797
28.7 Problem Set Solutions . . . 824

E Practice Sets 837

Practice Set I Questions 839

Practice Set I Solutions 861

Practice Set II Questions 875

Practice Set II Solutions 895

Practice Set III Questions 907

Practice Set III Solutions 925

Index 937

Preface

This study guide covers the material in the Casualty Actuarial Society (CAS) Exam Modern Actuarial Statistics - II. To get the most current examination syllabus, please check out https://www.casact.org/exam/exam-mas-ii-modern-actuarial-statistics-ii.

This study guide consists of four parts, which cover the four topics listed in the syllabus. The topics, the textbook references, and the weights are shown below:

Part A. Introduction to Credibility	15-25%
Tse, Nonlife Actuarial Models 6.1 - 6.3, 7.1 - 7.4, 8.1 - 8.2, 9.1, 9.2	

Part B. Linear Mixed Models	10-20%
(i) West, Linear Mixed Models 1 - 8 (excluding 2.9.6), and Appendix B (ii) Notes on Shrinkage	

Part C. Statistical Learning	40-50%
James, *et al.*, An Introduction to Statistical Learning 2.2.3, 8, 10	

Part D. Time Series with Constant Variance	15-25%
Cowpertwait and Metcalfe, Introductory Time Series with R, Chapters 1-5 (excluding Sections 3.3 and 3.4), 6, and 7 (Sections 7.1, 7.2, and 7.3)	

Part E consists of three practice sets. These questions are exam style questions to practice for your next CAS MAS-II exam.

Please feel free to contact support@actexlearning.com if you see something that is incorrect or unclear.

Part A

Introduction to Credibility

This is the first part of the MAS-II syllabus. We will cover all the eight learning objectives from Nonlife Actuarial Models. There are 4 main topics in this part:

(a) Limited Fluctuation Credibility (Nonlife Actuarial Models, Chapter 6.1–6.3)

(b) Bühlmann and Bühlmann–Straub Credibility (Nonlife Actuarial Models, Chapter 7.1–7.4)

(c) Bayesian Credibility (Nonlife Actuarial Models, Chapter 8.1–8.2)

(d) Empirical implementation of Credibility (Nonlife Actuarial Models, Chapter 9.1–9.2)

Under each topic (chapter), there are examples and questions indicated by level of difficulty with "*" (core question), "**" (advanced question), and "***" (mastery).

Important concepts and formulas are summarized in each chapter followed by a problem set. We suggest you work on core questions first, and move on to advanced and mastery questions when you are ready to challenge yourself with more difficult questions. In this context, credibility is referring to how much we can trust and rely on the experience period data for future predictions. In general, the more data you have, the more credible it is. The techniques listed above are ways to combine the information from the data with preconceived models to achieve more accurate predictions.

Chapter 2

Limited Fluctuation Credibility - Partial Credibility

If the expected or observed number of claims is less than the standard for full credibility, then full credibility is not attained. In this case, a value of $Z < 1$ has to be determined and the updated prediction is $U = Z\,D + (1 - Z)\,M$.

For the claim frequency, we require that the probability of $(Z \times N)$ lying within the interval $(Z\mu_N - k\mu_N, Z\mu_N + k\mu_N)$ is $1 - \alpha$ for a given k:

$$\Pr(Z\mu_N - k\,\mu_N \le Z\,N \le Z\mu_N + k\,\mu_N) = 1 - \alpha$$
$$\implies \Pr\left(\frac{-k\,\mu_N}{Z\sigma_N} \le \frac{N - \mu_N}{\sigma_N} \le \frac{k\,\mu_N}{Z\sigma_N}\right) = 1 - \alpha$$

Recall that $c = \sigma_N^2/\mu_N$. Applying the normal approximation, we have

$$\frac{k\,\mu_N}{Z\sigma_N} = z_{1-\alpha/2} \implies \frac{k\,\mu_N}{Z\sqrt{\mu_N \cdot c}} = z_{1-\alpha/2} \implies Z = \left(\frac{k}{z_{1-\alpha/2}}\right)\sqrt{\frac{\mu_N}{c}} = \sqrt{\frac{\mu_N}{\lambda_F \cdot c}} = \sqrt{\frac{\mu_N}{n_0}}.$$

The partial credibility factors for claim severity, aggregate loss, and pure premium can be derived in a similar way. See Eq. (1.2.3) for the the value of c in the $(a, b, 0)$ class.

The **partial credibility factors** for claim frequency, claim severity, aggregate loss, and pure premium are summarized below.

Claim Frequency:

$$Z = \sqrt{\frac{\mu_N}{n_0}} = \sqrt{\frac{\mu_N}{\lambda_F \cdot c}} = \sqrt{\frac{\mu_N}{\left(\frac{z_{1-\alpha/2}}{k}\right)^2 \left(\frac{\sigma_N^2}{\mu_N}\right)}} \tag{2.0.1}$$

Claim Severity:

$$Z = \sqrt{\frac{N}{n_0}} = \sqrt{\frac{N}{\lambda_F\,C_X^2}} \tag{2.0.2}$$

Aggregate Loss and Pure Premium:

$$Z=\sqrt{\frac{\mu_N}{n_0}}=\sqrt{\frac{\mu_N}{\lambda_F(c+C_X^2)}}=\sqrt{\frac{\mu_N}{\lambda_F\left(\frac{\sigma_N^2}{\mu_N}+C_X^2\right)}} \tag{2.0.3}$$

It is also called the **square-root rule**. Within the square root, the denominator is the standard for full credibility of the corresponding risk measure, and the numerator, μ_N or N, is observed from data. μ_N is the expected number of claims coming from the data, and N is the observed number of claims. If μ_N can not be calculated from the data, then the observed number of claims can be used to calculate the partial credibility factor. Note that if the ratio is greater than 1, then full credibility is attained and $Z = 1$.

Example 2.1. *[a] A block of insurance policies had 896 claims this period with mean loss of 45 and variance of loss of 5,067. Full credibility is based on a coverage probability of 98% for a range of within 10% deviation from the true mean. The mean frequency of claims is 0.09 per policy and the block has 18,600 policies. Calculate Z for the claim frequency for the next period.

[a] Nonlife Actuarial Models, Example 6.11

Solution. The expected claim frequency for the block of policies is $\mu_N = (18,600)(0.09) = 1,674$. No frequency distribution is named, so a Poisson frequency distribution is assumed. Since $z_{1-\alpha/2} = \Phi^{-1}(0.99) = 2.326$ and $k = 10\%$, the full-credibility standard for claim frequency is $n_0 = \lambda_F = (2.326/0.1)^2 = 542$. Since $\mu_N > \lambda_F$, full credibility is attained for claim frequency and $Z = 1$.

Example 2.2. * (continued) A block of insurance policies had 896 claims this period with mean loss of 45 and variance of loss of 5,067. Full credibility is based on a coverage probability of 98% for a range of within 10% deviation from the true mean. The mean frequency of claims is 0.09 per policy and the block has 18,600 policies. Calculate Z for the claim severity for the next period.

Solution. Assuming a Poisson frequency distribution, the standard for full credibility for claim severity is

$$n_0=\lambda_F\,C_X^2=\left(\frac{2.326}{0.1}\right)^2\left(\frac{5,067}{45^2}\right)=1,357.$$

The block had $N = 896$ claims this period. Therefore, the partial credibility factor is

$$Z=\sqrt{\frac{N}{\lambda_F\,C_X^2}}=\sqrt{\frac{896}{1,357}}=0.813.$$

Example 2.3. ** Claim severity has mean 342 and standard deviation 408. An insurance company has 75,000 insurance policies. Using the classical credibility approach with coverage probability of 95% to within 5% of the aggregate loss, determine the credibility factor Z if the average claim per policy is 4%.

Solution. For 95% coverage probability, $\alpha = 1 - 0.95 = 0.05$ and thus $z_{1-\alpha/2} = z_{0.975} = \Phi^{-1}(0.975) = 1.96$. We are given $\mu_X = 342$, $\sigma_X = 408$. The frequency distribution isn't mentioned, so we assume a Poisson distribution. For $k = 0.05$, the standard for full credibility for aggregate loss is

$$n_0 = \lambda_F(1 + C_X^2) = \left(\frac{1.96}{0.05}\right)^2 \left(1 + \frac{(408)^2}{(342)^2}\right) = 3,724.$$

The expected number of claims is $\mu_N = (75,000)(4\%) = 3,000$ and the credibility factor is

$$Z = \sqrt{\frac{\mu_N}{\lambda_F(1 + C_X^2)}} = \sqrt{\frac{3,000}{3,724}} = 0.898.$$

Example 2.4.** (continued) The premium currently used to reflect a pure premium per policy is $M = 14$. The pure premium experienced during the past year has an average of $D = 20$ per policy. Using the limited fluctuation credibility approach, what pure premium per policy should be reflected in the new rates?

Solution. The credibility factor is $Z = 0.898$ from the previous example. For $M = 14$ and $D = 20$, the new rate is $U = Z\,D + (1 - Z)M = (0.898)(20) + (1 - 0.898)(14) = 19.388$.

Example 2.5.*** You're reviewing some work from a former co-worker and notice that they didn't write down one of their underlying assumptions. Since this person no longer works for the company, you can't just ask them for it, but instead have to retrace their steps to figure it out. Here is the information that you know:

(i) The number of claims per policy follows a negative binomial distribution with $r = 3$ and $\beta = 0.5$.

(ii) Claim sizes follow an exponential distribution with mean 1,000.

(iii) The number of claims and claim sizes are independent.

(iv) 1,500 policies were observed and generated 2,000 claims totalling \$3 million in losses.

(v) Using limited fluctuation credibility and $k = 0.02$, the predicted aggregate losses for the 1,500 policies was \$2.475 million.

(vi) The expected aggregate losses were used for the manual rate.

What was the coverage probability used for this analysis?

Solution. Use equation (1.1.1) to find Z, where $U = 2,475,000$, $D = 3,000,000$, and M equals the expected aggregate losses from the 1,500 policies. Using the given distributions, we have

$$M = 1500\mu_N\mu_X = 1500(3)(0.5)(1000) = 2,250,000$$

and

$$U = Z\,D + (1-Z)\,M \implies U = Z(D-M) + M$$
$$\implies 2,475,000 = Z(3,000,000 - 2,250,000) + 2,250,000$$
$$\implies Z = 0.3.$$

Because aggregate losses were being predicted, then equation (2.0.3) must have been used to calculate Z. We're given the distribution for the frequency per policy and can create the distribution for the frequency for 1,500 policies (by multiplying r by 1,500), but the results will be the same regardless of which frequency distribution is used. For consistent notation, we'll use the frequency distribution for all 1,500 policies. From the MAS-II Tables, we get

$$C_X^2 = \frac{\sigma_X^2}{\mu_X^2} = \frac{E(X^2) - \mu_X^2}{\mu_X^2} = \frac{2\cdot\theta^2 - \theta^2}{\theta^2} = \frac{\theta^2}{\theta^2} = 1.$$

For equation (2.0.3), we need to use the expected number of claims from the policies in the numerator. **Note** : An easy mistake to make here would be to use the observed number of claims instead of the expected number of claims for the credibility factor. From Eq. (1.2.3) we have $c = 1+\beta$. Putting this together, we use equation (2.0.3) to find $z_{1-\alpha/2}$:

$$Z = \sqrt{\frac{\mu_N}{\left(\frac{z_{1-\alpha/2}}{k}\right)^2 (c + C_X^2)}}$$

$$\implies 0.3 = \sqrt{\frac{2,250}{\left(\frac{z_{1-\alpha/2}}{0.02}\right)^2 (1 + 0.5 + 1)}} \implies z_{1-\alpha/2} = 2$$

Looking at the Normal table in the MAS-II Tables, we see that $1-\alpha/2 = 0.9772$, which implies $\alpha = 4.56\%$ and the coverage probability is 95.44%.

Recap: Partial credibility factors:

Example 2.1: Calculate Z for the claim frequency: $Z = \sqrt{\dfrac{\mu_N}{\lambda_F}}$

Example 2.2: Calculate Z for the claim severity: $Z = \sqrt{\dfrac{N}{\lambda_F\, C_X^2}}$

Example 2.3: Calculate Z for the aggregate loss: $Z = \sqrt{\dfrac{\mu_N}{\lambda_F + \lambda_F\, C_X^2}}$

Example 2.4: Calculate the updated prediction using Z.

Example 2.5: Calculate the coverage probability (negative binomial + exponential case).

Summary

The main topic of this chapter is to calculate the Updated prediction U:

$$U = Z\,D + (1 - Z)\,M,$$

where D is based on the recent claim experience Data, M is based on a rate specified in the Manual, and the weight Z is the credibility factor.

When the sample size is large enough, we assign full credibility $Z = 1$, and thus $U = D$ (i.e., the updated predication = the observed data). If not, we apply the square-root rule for partial credibility. Table **2.1** summarizes the formulas of partial-credibility factors for the four risk measures.

Things to remember:

(i) The standard for full credibility for aggregate loss and that for pure premium are the same.

(ii) The standard for full credibility for aggregate loss is the sum of the standard for claim frequency and the standard for claim severity.

(iii) The partial credibility is also called the **square-root rule**. Within the square root, the denominator is the standard for full credibility of the corresponding risk measure.

Table **2.1**: Summary of standards for partial credibility factor Z if Poisson frequency

Loss measures	Partial-credibility factor Z
Claim frequency	$Z = \sqrt{\dfrac{\mu_N}{\lambda_F}}$
Claim severity	$Z = \sqrt{\dfrac{N}{\lambda_F\, C_X^2}}$
Aggregate loss or Pure premium	$Z = \sqrt{\dfrac{\mu_N}{\lambda_F(1 + C_X^2)}}$

Recall: $\lambda_F = \left(\dfrac{z_{1-\alpha/2}}{k}\right)^2$, $C_X = \dfrac{\sigma_X}{\mu_X}$ and $1 + C_X^2 = \dfrac{\mathrm{E}(X^2)}{\mu_X^2}$.

2.1 Problem Set

Question 2.1. * (4B 1985 Spring #30) The 1984 pure premium underlying the rate equals 1000. The loss experience is such that the actual pure premium for that year equals 1200 and the number of claims equals 600.

If 5400 claims are needed for full credibility and the square root rule for partial credibility is used, estimate the pure premium underlying the rate in 1985.

(Assume no change in the pure premium due to inflation.)

Question 2.2. ** (4B 1986 Spring #35) You are in the process of revising rates.

(i) The premiums currently being used reflect a pure premium per insured of 100. The pure premium experienced during the two year period used in the rate review averaged 130 per insured.

(ii) The average frequency during the two year review period was 250 claims per year.

Using a full credibility standard of 2500 claims and assigning partial credibility according to the limited fluctuation credibility approach, what pure premium per insured should be reflected in the new rates?

(Assume that there is no inflation.)

Question 2.3. * (4B 1991 Spring #23 revised) Claim counts for a group follow a Poisson distribution. The standard for full credibility is 12,000 expected claims. We observed 6,600 claims and a total loss of 12,300,000 for a group of insureds.

If our prior estimate of the total loss is 13,200,000, determine the classical credibility estimate of the total loss for the group of insureds.

Question 2.4. * (4B 1992 Spring #6 revised) You are given the following information for a group of insureds:

Prior estimate of expected total losses	20,000,000
Observed total losses	25,000,000
Observed number of claims	10,008
Required number of claims for full credibility	17,792

Using the methods of classical credibility, determine the estimate for the group's expected total losses based upon the latest observation.

Question 2.5. ** (C 2000 Spring #26) You are given:

(i) Claim counts follow a Poisson distribution.

(ii) Claim sizes follow a lognormal distribution with coefficient of variation 3.

(iii) Claim sizes and claim counts are independent.

(iv) The number of claims in the first year was 1000.

(v) The aggregate loss in the first year was 6.75 million.

(vi) The manual premium for the first year was 5.00 million.

(vii) The exposure in the second year is identical to the exposure in the first year.

(viii) The full credibility standard is to be within 5% of the expected aggregate loss 95% of the time.

Determine the limited fluctuation credibility pure premium (in millions) for the second year.

Question 2.6. *** (Exam C 2001 Fall #15) You are given the following information about a general liability book of business comprised of 2500 insureds:

(i) $X_i = \sum_{j=1}^{N_i} Y_{ij}$ is a random variable representing the annual loss of the ith insured.

(ii) $N_1, N_2, \ldots, N_{2500}$ are independent and identically distributed random variables following a negative binomial distribution with parameters $r = 2$ and $\beta = 0.2$

(iii) $Y_{i1}, \ldots, Y_{iN_i}$ are independent and identically distributed random variables following a Pareto distribution with $\alpha = 3$ and $\theta = 1000$.

(iv) The full credibility standard is to be within 5% of the expected aggregate losses 90% of the time.

Using classical credibility theory, determine the partial credibility of the annual loss experience for this book of business.

Question 2.7. ** (C 2003 Fall #35) You are given:

(i) X_{partial} = pure premium calculated from partially credible data

(ii) $\mu = \mathrm{E}\left[X_{\text{partial}}\right]$

(iii) Fluctuations are limited to $\pm k$ of the mean with probability P

(iv) Z = credibility factor

Which of the following is equal to P?

A. $\Pr\left[\mu - k\mu \leq X_{\text{partial}} \leq \mu + k\mu\right]$

B. $\Pr\left[Z\mu - k \leq ZX_{\text{partial}} \leq Z\mu + k\right]$

C. $\Pr\left[Z\mu - \mu \leq ZX_{\text{partial}} \leq Z\mu + \mu\right]$

D. $\Pr\left[1 - k \leq ZX_{\text{partial}} + (1 - Z)\mu \leq 1 + k\right]$

E. $\Pr\left[\mu - k\mu \leq ZX_{\text{partial}} + (1 - Z)\mu \leq \mu + k\mu\right]$

Question 2.8. * You are given the following information for limited-fluctuation credibility:

(i) Claim frequency follows a Poisson distribution.

(ii) Claim frequency and claim severity are independent.

(iii) The full credibility standard for claim frequency is 400.

(iv) In calculating partial credibility factors, you use $\mu_N = N = 100$ from the observed data.

(v) The credibility factors for the claim frequency and the claim severity are the same.

Determine the credibility factor for the aggregate loss.

Question 2.9. *** For a block of policies with 900 observed claims, the claim frequency per policy follows a binomial distribution with $m = 4$ and $q = 0.2$. Claim severity follows a Pareto distribution with unknown parameters. All credibility standards have the same coverage probability and range parameter, k. The credibility factor for frequency is 0.6511, while the credibility factor for aggregate loss is 0.2987.

What is the coefficient of variation for the severity distribution?

Question 2.10. ** You are provided with the following claims frequency data about the portfolio of a large commercial auto insurance:

(i) The claims frequency of each vehicle is independently and identically distributed.

(ii) The overall average claims frequency is $M = 0.1$ per earned car year.

(iii) The limited-fluctuation credibility standard is 1,250 claims.

(iv) In its prior experience, X claims were observed in 2,000 earned car years.

(v) The updated prediction of claims frequency is 0.16 per earned car year using limited-fluctuation credibility.

Determine X, the number of observed claims from the 2,000 earned car years in its prior experience.

MAS-II past exam questions

Question 2.11. (MAS-II 2018 Fall #6) You are given the following information:

(i) A block of insurance policies had 1,384 claims this period.

(ii) The claims had a mean loss of 55 and variance of loss of 6,010.

(iii) The mean frequency of these claims is 0.085 per policy.

(iv) The block has 21,000 policies.

(v) Full credibility is based on a coverage probability of 98% for a range of within 5% deviation from the true mean.

You calculate the partial-credibility factor for severity, Z_x, and the partial-credibility factor for pure premium, Z_p , using the limited-fluctuation credibility method.

Calculate the absolute difference between Z_x and Z_p.

Question 2.12. (MAS-II 2019 Spring #5) You are provided with the following claims frequency data about the portfolio of a large commercial auto insurer:

(i) The claims frequency of each vehicle is independently and identically distributed.

(ii) The overall average claims frequency is 0.2 claims per earned car year.

(iii) The variance of the hypothetical means is 0.3.

(iv) The expected value of the process variance is 1200.

(v) The limited-fluctuation credibility standard is 1083 claims.

You are asked to change the credibility methodology from limited-fluctuation Bühlmann-Straub for a policyholder with 200 claims in 1800 earned car years in its prior loss experience.

Calculate the percentage change in the estimate of claims frequency for this policyholder due to the change in methodology.

Question 2.13. (MAS-II 2019 Spring #6) An insurance company is currently using a limited-fluctuation credibility approach for a line of business with the following assumptions:

(i) The claim frequency follows a Poisson distribution.

(ii) The mean of the claim frequency is large enough to justify the normal approximation to the Poisson.

(iii) The square root rule is used to determine partial credibility.

(iv) The standard for full credibility is the number of claims at which there is a 99% probability that the observed aggregate loss is within 5% of the mean.

You are given the following information about a block of 10,000 policies:

(i) The mean claim frequency is 0.12.

(ii) The mean claim severity is 100.

(iii) The variance of claim severity is 14,400.

Calculate the credibility for this block of policies using the partial credibility method for aggregate loss.

2.2 Problem Set Solutions

Question 2.1. We are given that 5400 claims are required for full credibility but the actual number of claims was only 600. Thus the credibility factor is $Z = \sqrt{600/5400} = 1/3$. The observed rate is 1200 so the new rate will be

$$U = ZD + (1 - Z)M = (1/3)(1200) + (2/3)(1000) = 1066.67.$$

Question 2.2. The full credibility standard is given as 2500. The data is based on a total of $N = (2)(250) = 500$ claims, so its credibility is $Z = \sqrt{500/2500} = 0.4472$. For $M = 100$ and $D = 130$, the new rate is $U = Z\,D + (1 - Z)M = (0.4472)(130) + (1 - 0.4472)(100) = 113.42$.

Question 2.3. The credibility factor for claim frequency is

$$Z = \sqrt{\frac{\mu_N}{n_0}} = \sqrt{\frac{\mu_N}{\lambda_F}} = \sqrt{\frac{6,600}{12,000}} = 0.74162.$$

The estimated total loss is

$$U = Z\,D + (1 - Z)\,M = (0.74162)(12,300,000) + (1 - 0.74162)(13,200,000) = 12,532,542.$$

Question 2.4. The credibility factor is

$$Z = \sqrt{\frac{\mu_N}{\lambda_F}} = \sqrt{\frac{10,008}{17,792}} = 0.75.$$

The estimate for the group's expected total losses is

$$U = Z\,D + (1 - Z)\,M = (0.75)(25,000,000) + (1 - 0.75)(20,000,000) = 23,750,000.$$

Question 2.5. We have $\alpha = 1 - 0.95 = 0.05$ and thus $z_{1-\alpha/2} = z_{0.975} = 1.96$. The minimum number of claims needed for the aggregate loss to be fully credible is

$$n_0 = \lambda_F(1 + C_X^2) = \left(\frac{1.96}{0.05}\right)^2 (1 + 3^2) = 15366.4,$$

$$Z = \sqrt{\frac{\mu_N}{\lambda_F(1 + C_X^2)}} = \sqrt{\frac{1000}{15366.4}} = 0.255,$$

$$U = Z\,D + (1 - Z)\,M = (0.255)(6.75) + (0.745)(5) = 5.45.$$

Question 2.6. We have $k = 0.05$ and $P = 0.9$. Find y such that $\Phi(y) = (1 + P)/2 = 0.95$. Hence, $y = 1.645$. Using the Pareto formula in Table 1.2, we have $1 + C_x^2 = 2(\alpha - 1)/(\alpha - 2)$. For a compound negative binomial with Pareto severity, the expected number of claims needed for full credibility is

$$n_0 = \left(\frac{y}{k}\right)^2 (1 + \beta + C_X^2) = \left(\frac{y}{k}\right)^2 \left(\beta + \frac{2(\alpha - 1)}{\alpha - 2}\right) = (1082.41)\left[0.2 + \frac{2(3-1)}{3-2}\right] = 4546.122.$$

The expected number of claims is $\mathrm{E}[N] = r\beta = 0.4$ per insured and $n = (2500)\mathrm{E}(N) = 1000$ for the book of 2500 insureds. Hence, the partial credibility of the annual loss experience for this book of business is

$$Z = \sqrt{\frac{n}{n_0}} = \sqrt{\frac{1000}{4546.122}} = 0.47.$$

Question 2.7. We are to limit the fluctuation in the term ZX_{partial}. The mean is $\mathrm{E}[ZX_{\text{partial}}] = Z\mu$ and the fluctuations are limited to $\pm k\mu$ of the mean. Thus

$$\Pr[Z\mu - k\mu \le ZX_{\text{partial}} \le Z\mu + k\mu] = P \implies \Pr[\mu - k\mu \le ZX_{\text{partial}} + (1 - Z)\mu \le \mu + k\mu] = P.$$

Answer: E

Question 2.8. The standard for full credibility for claim frequency is $n_0 = 400$. The credibility factor for claim frequency is (Eq. 2.0.1)

$$Z = \sqrt{\frac{\mu_N}{n_0}} = \sqrt{\frac{\mu_N}{\lambda_F}} = \sqrt{\frac{100}{400}} = 0.5.$$

The credibility factor for claim severity is also 0.5. Hence, the credibility factor for claim severity can be obtained as (Eq. 2.0.2):

$$Z = \sqrt{\frac{N}{\lambda_F\, C_X^2}} \implies 0.5 = \sqrt{\frac{100}{\lambda_F\, C_X^2}} \implies \lambda_F\, C_X^2 = 400$$

The credibility factor for the aggregate loss is (Eq. 2.0.3)

$$Z = \sqrt{\frac{\mu_N}{\lambda_F + \lambda_F\, C_X^2}} = \sqrt{\frac{100}{400 + 400}} = 0.3536.$$

Question 2.9. From the credibility factor for claim frequency and using equation (2.0.1), we can solve for the claim frequency full credibility standard. The problem gives us the number of observed claims, but does not give us enough information to determine the number of expected claims from the block of policies. So, we'll use the number of observed claims with equation (2.0.1).

$$Z_{freq} = \sqrt{\frac{N}{n_{0,freq}}} \implies 0.6511 = \sqrt{\frac{900}{n_{0,freq}}} \implies n_{0,freq} = 2,123$$

We can use the claim frequency full credibility standard to get λ_F, using equation (1.2.4).

$$n_{0,freq} = \lambda_F \left(\frac{\sigma_N^2}{\mu_N}\right) \implies 2,123 = \lambda_F \left(\frac{4(0.2)(0.8)}{4(0.2)}\right) \implies \lambda_F = 2,653.75$$

Now, we have all of the information we need to use equation (2.0.3) and solve for the coefficient of variation for the severity distribution.

$$Z_{agg} = \sqrt{\frac{N}{n_{0,freq} + \lambda_F\, C_X^2}} \implies 0.2987 = \sqrt{\frac{900}{2,123 + 2,653.75 \cdot C_X^2}} \implies C_X = \sqrt{3}$$

Question 2.10. We are given that the limited-fluctuation credibility standard is $n_0 = 1,250$ claims. The expected number of claims from the sample is equal to $\mu_N = (2,000)(0.1) = 200$. Hence the partial credibility factor is (Sec. 2)

$$Z = \sqrt{\mu_N/n_0} = \sqrt{200/1,250} = 0.4.$$

The observed claim frequency is $D = X/2,000$ and the updated prediction of claims frequency per earned car year is $U = 0.16$. Using the limited-fluctuation credibility, we have

$$U = Z\,D + (1-Z)\,M \implies 0.16 = (0.4)(X/2000) + (1-0.4)(0.1) \implies X = 500.$$

Hence, in its prior experience, 500 claims were observed in 2,000 earned car years.

Question 2.11. Given $\alpha = 1 - 98\% = 2\%$, we have $z_{1-\alpha/2} = \Phi^{-1}(0.99) = 2.326$. Given $k = 5\%$, the full-credibility standard for claim frequency is

$$\lambda_F = \left(\frac{2.326}{0.05}\right)^2 = 2,164.11.$$

The estimated coefficient of variation for claim severity is

$$C_X = \frac{\sigma_X}{\mu_X} = \frac{\sqrt{6,010}}{55}.$$

Calculate

$$\lambda_F\, C_X^2 = (2,164.11)(6,010/55^2) = 4,299.6.$$

The block had $N = 1,384$ claims and hence the partial credibility factor for severity is

$$Z_x = \sqrt{\frac{N}{\lambda_F\, C_X^2}} = \sqrt{\frac{1,384}{4,299.6}} = 0.567.$$

The mean frequency is 0.085 and hence $\mu_N = (21,000)(0.085) = 1,785$. The partial credibility factor for pure premium is

$$Z_p = \sqrt{\frac{\mu_N}{\lambda_F + \lambda_F\, C_X^2}} = \sqrt{\frac{1,785}{2,164.11 + 4,299.6}} = 0.526.$$

Therefore, $|Z_x - Z_p| = |0.567 - 0.526| = 0.041$.

Question 2.12. We are given that the observed claim frequency is $D = 200/1,800 = 1/9$, the overall average claims frequency is $M = 0.2$ per earned car year, and the limited-fluctuation credibility standard is 1,083 claims. The expected number of claims from the samples is equal to $\lambda_N = 1800(0.2) = 360$. Hence the partial credibility factor is (Sec. 2)

$$Z = \sqrt{\frac{\lambda_N}{\lambda_F}} = \sqrt{360/1,083} = 0.57655.$$

The updated prediction is

$$U = Z\,D + (1 - Z)\,M = (0.57655)(1/9) + (1 - 0.57655)(0.2) = 0.14875.$$

We are given that the variance of the hypothetical means is $\hat{\sigma}^2_{\text{HM}} = 0.3$, the expected value of the process variance is $\hat{\mu}_{\text{PV}} = 1,200$, and hence

$$\hat{k} = \frac{\hat{\mu}_{\text{PV}}}{\hat{\sigma}^2_{\text{HM}}} = \frac{1,200}{0.3} = 4,000.$$

Using Bühlmann-Straub credibility (Sec. 4), we have

$$\hat{Z}_i = \frac{m_i}{m_i + \hat{k}} = \frac{1,800}{1,800 + 4,000} = 9/29,$$

$\bar{X}_i = D = 1/9$, $\bar{X} = M = 0.2$, and

$$P = \hat{Z}_i \bar{X}_i + (1 - \hat{Z}_i)\bar{X} = (9/29)(1/9) + (20/29)(0.2) = 5/29.$$

The percentage change is $(P - U)/U = (5/29 - 0.14875)/0.14875 = 15.9\%$.

Question 2.13. We are given $\alpha = 1 - 0.99 = 0.01$, $k = 0.05$, and hence (Eq. 1.2.5)

$$\lambda_F = \left(\frac{z_{1-\alpha/2}}{k}\right)^2 = \left(\frac{z_{.995}}{k}\right)^2 = \left(\frac{2.576}{0.05}\right)^2 = 2,655. \qquad (2.2.1)$$

The standard for full credibility for aggregate loss is $\lambda_F(1 + C_X^2)$ where $C_X = \sigma_X/\mu_X$ is the coefficient of variation of the claim severity (p. 15). Therefore,

$$C_X^2 = \sigma_X^2/\mu_X^2 = 14,400/(100)^2 = 1.44$$
$$\lambda_F(1 + C_X^2) = (2,655)(1 + 1.44) = 6,478.$$

We are given that the mean claim frequency is 0.12 and hence we expect to observe $\mu_N = (0.12)(10,000) = 1,200$ claims from a block of 10,000 policies. The partial credibility factor for aggregate loss is (Chapter 2)

$$Z = \sqrt{\frac{\mu_N}{\lambda_F(1 + C_X^2)}} = \sqrt{\frac{1,200}{6,478}} = 0.43.$$

Chapter 3

Bühlmann Credibility

Learning objective:

(i) Understand the basic framework of Bühlmann credibility

(ii) Calculate different variance components for Bühlmann credibility

(iii) Calculate Bühlmann credibility factor and estimates for frequency, severity, and aggregate loss

Consider a risk group with loss measure denoted by X, which may be **claim frequency**, **claim severity**, **aggregate loss**, or **pure premium**. The risk measure X is characterized by a parameter Θ. Denote the conditional mean of X given $\Theta = \theta$ by

$$\mu_X(\theta) = \mathrm{E}(X|\Theta = \theta)$$

and the conditional variance by

$$\sigma_X^2(\theta) = \mathrm{Var}(X|\Theta = \theta).$$

Here Θ is a random variable and its distribution is called the **prior distribution** in the Bayesian statistical inference. Therefore, the conditional mean and the conditional variance of X become random variables in Θ. Denote the **hypothetical mean** by

$$\mu_X(\Theta) = \mathrm{E}(X|\Theta) \tag{3.0.1}$$

and the **process variance** by

$$\sigma_X^2(\Theta) = \mathrm{Var}(X|\Theta). \tag{3.0.2}$$

The expected value of the hypothetical means (also called the **unconditional mean** or overall mean) of X is[i]

$$\mathrm{E}(X) = \mathrm{E}[\mathrm{E}(X|\Theta)] = \mathrm{E}[\mu_X(\Theta)]. \tag{3.0.3}$$

[i] $\mathrm{E}[\mathrm{E}(X|\Theta)] = \int_\Theta \int_X x f(x|\theta)\mathrm{d}x\, \pi(\theta)\mathrm{d}\theta = \int_X x \int_\Theta f(x,\theta)\mathrm{d}\theta \mathrm{d}x = \int_X x\, f(x)\mathrm{d}x = \mathrm{E}(X)$

Denote the **expected value of the process variance** (EPV) by

$$\mu_{\mathrm{PV}} = \mathrm{E}[\mathrm{Var}(X|\Theta)] = \mathrm{E}[\sigma_X^2(\Theta)] \tag{3.0.4}$$

and the **variance of the hypothetical means** (VHM) by

$$\sigma^2_{\mathrm{HM}} = \mathrm{Var}[\mathrm{E}(X|\Theta)] = \mathrm{Var}[\mu_X(\Theta)]. \tag{3.0.5}$$

The **unconditional variance**, or **total variance**, of X is[ii]

$$\mathrm{Var}(X) = \mathrm{E}[\mathrm{Var}(X|\Theta)] + \mathrm{Var}[\mathrm{E}(X|\Theta)] = \mu_{\mathrm{PV}} + \sigma^2_{\mathrm{HM}}. \tag{3.0.6}$$

Define k as the ratio of EPV to VHM, i.e.

$$k = \frac{\mathrm{EPV}}{\mathrm{VHM}} = \frac{\mu_{\mathrm{PV}}}{\sigma^2_{\mathrm{HM}}}, \tag{3.0.7}$$

which is the key (Bühlmann) parameter in formulating the Bühlmann credibility factor. The credibility factor is defined as

$$Z = \frac{n}{n+k}, \tag{3.0.8}$$

where n represents the number of observations. The updated prediction is calculated the same way as with Limited Fluctuation Credibility

$$U = Z\,D + (1-Z)\,M,$$

where D is calculated from the sample (sample mean) and M is calculate from population (overall mean or unconditional mean).

Note that the credibility factor is a function of n, EPV, and VHM. Increasing the sample size n will increase the credibility factor, and hence assign more weight on the observed data in updating our revised prediction for future losses. A small EPV, or large VHM will give rise to a small k. The risk groups will be more distinguishable in the mean when k is smaller, in which case we may put more weight on the data.

Example 3.1. * You are given the following:

(i) X is a random variable with mean Θ_1 and variance Θ_2.

(ii) Θ_1 is a random variable with mean 2 and variance 4.

(iii) Θ_2 is a random variable with mean 8 and variance 32.

Determine the value of the Bühlmann credibility factor Z, after $n = 8$ observations of X.

Solution. Let Θ denote the vector random variable representing Θ_1 and Θ_2, i.e. $\Theta = (\Theta_1, \Theta_2)$. The hypothetical mean is $\mathrm{E}(X|\Theta) = \Theta_1$, the process variance is $\mathrm{Var}(X|\Theta) = \Theta_2$. The expected value of the process variance (EPV) is $\mathrm{E}[\mathrm{Var}(X|\Theta)] = \mathrm{E}(\Theta_2) = 8$, the variance of the hypothetical

[ii] $\mathrm{E}[\mathrm{Var}(X|\Theta)] + \mathrm{Var}[\mathrm{E}(X|\Theta)] = \mathrm{E}[\mathrm{E}(X^2|\Theta) - \mathrm{E}(X|\Theta)^2] + \mathrm{E}[\mathrm{E}(X|\Theta)^2] - \{\mathrm{E}[\mathrm{E}(X|\Theta)]\}^2 = \mathrm{E}(X^2) - [\mathrm{E}(X)]^2 = \mathrm{Var}(X)$

means (VHM) is $\mathrm{Var}[\mathrm{E}(X|\Theta)] = \mathrm{Var}(\Theta_1) = 4$, and thus

$$k = \frac{\mathrm{EPV}}{\mathrm{VHM}} = \frac{\mathrm{E}[\mathrm{Var}(X|\Theta)]}{\mathrm{Var}[\mathrm{E}(X|\Theta)]} = \frac{8}{4} = 2.$$

Therefore, the credibility factor is

$$Z = \frac{n}{n+k} = \frac{8}{8+2} = 0.8.$$

Aggregate Losses

Denote S as aggregate losses

$$S = X_1 + X_2 + \cdots + X_N.$$

The hypothetical mean of S is

$$\mu_S(\Theta) = \mathrm{E}(S|\Theta) = \mathrm{E}(N|\Theta)\mathrm{E}(X|\Theta), \tag{3.0.9}$$

and the process variance of S is

$$\sigma_S^2(\Theta) = \mathrm{Var}(S|\Theta) = \mu_N(\Theta)\sigma_X^2(\Theta) + \sigma_N^2(\Theta)\mu_X^2(\Theta). \tag{3.0.10}$$

Therefore, the **expected value of the process variance** (EPV) of the aggregate losses is

$$\mu_{\mathrm{PV}} = \mathrm{E}[\mathrm{Var}(S|\Theta)] = \mathrm{E}[\sigma_S^2(\Theta)] \tag{3.0.11}$$

and the **variance of the hypothetical means** (VHM) by

$$\sigma_{\mathrm{HM}}^2 = \mathrm{Var}[\mathrm{E}(S|\Theta)] = \mathrm{Var}[\mu_S(\Theta)]. \tag{3.0.12}$$

If N is Poisson distributed with parameter Λ, then $\mu_N(\Theta) = \sigma_N^2(\Theta) = \Lambda$ and

$$\sigma_S^2(\Theta) = \Lambda[\sigma_X^2(\Theta) + \mu_X^2(\Theta)] = \Lambda\,\mathrm{E}(X^2|\Theta). \tag{3.0.13}$$

Assume that $\{X_1, \cdots, X_n, X_{n+1}\}$ are independently and identically distributed (iid) given the parameter θ. Bühlmann (1967) estimated the future loss measure X_{n+1} based on a linear predictor of the past observations $\boldsymbol{X} = \{X_1, \cdots, X_n\}$. The predictor minimizes the mean squared error[iii] in predicting X_{n+1} over the joint distribution of Θ, $\boldsymbol{X}$, and X_{n+1}. **Bühlmann's approach** is also called the **greatest accuracy approach** or the **least squares approach**.

The predictor of X_{n+1} is given by

$$\hat{X}_{n+1} = Z\,\bar{X} + (1-Z)\,\mu_X, \tag{3.0.14}$$

[iii] Nonlife Actuarial Models, p. 203–206

where $\hat{X}_{n+1}$ is called the **Bühlmann premium**,

$$Z = \frac{n}{n+k} \tag{3.0.15}$$

is called the **Bühlmann credibility factor** or simply the Bühlmann credibility, $\bar{X} = \sum X_i/n$ is the sample mean, $\mu_X = \mathrm{E}(X)$ is the unconditional mean, and k is defined in equation (3.0.7). In general, n represents the number of observations, which is:

(i) the number of periods, over which the number of claims is aggregated, for predicting **claim frequency** N.

(ii) the number of claims, for predicting **claim severity** X .

(iii) the number of periods of claim experience for predicting **aggregate loss** S.

The following question calcualte the value of Bühlmann's k for aggregate losses when the prior distribution is continuous.

Example 3.2. ** (4B 1999 Spring #13) You are given the following:

(i) The number of claims follows a distribution with mean λ and variance 2λ.

(ii) Calim sizes follow a distribution with mean μ and variance $2\mu^2$.

(iii) The number of claims and claim sizes are independent.

(iv) λ and μ have a prior probability distribution with joint density function $f(\lambda, \mu) = 1$, $0 < \lambda < 1$, $0 < \mu < 1$.

Determine the value of Bühlmann's k for aggregate losses.

Solution. We are given $\Theta = (\lambda, \mu)$. It's easy to show that the prior parameters λ and μ are independent and uniformly distributed. Hence

$$\begin{aligned} \mathrm{E}(\lambda\mu) &= \mathrm{E}(\lambda)\mathrm{E}(\mu) = (1/2)(1/2) = 1/4, \\ \mathrm{E}[(\lambda\mu)^2] &= \mathrm{E}(\lambda^2)\mathrm{E}(\mu^2) = (1/3)(1/3) = 1/9, \\ \mathrm{Var}(\lambda\mu)] &= 1/9 - (1/4)^2 = 7/144. \end{aligned}$$

Using Equations (3.0.10) - (3.0.12), the hypothetical mean of S is

$$\mu_S(\Theta) = \mathrm{E}(S|\Theta) = \mathrm{E}(N|\lambda)\mathrm{E}(X|\mu) = \lambda\mu,$$

and the process variance of S is

$$\sigma_S^2(\Theta) = \mu_N(\Theta)\sigma_X^2(\Theta) + \sigma_N^2(\Theta)\mu_X^2(\Theta) = (\lambda)(2\mu^2) + (2\lambda)(\mu^2) = 4\lambda\mu^2.$$

Therefore, the **expected value of the process variance** (EPV) of the aggregate losses is

$$\mu_{\mathrm{PV}} = \mathrm{E}[\sigma_S^2(\Theta)] = \mathrm{E}(4\lambda\mu^2) = 4(1/2)(1/3) = 2/3$$

and the **variance of the hypothetical means** (VHM) by

$$\sigma_{\mathrm{HM}}^2 = \mathrm{Var}[\mu_S(\Theta)] = \mathrm{Var}(\lambda\mu) = 7/144.$$

Therefore, $k = \mu_{\mathrm{PV}}/\sigma_{\mathrm{HM}}^2 = (2/3)/(7/144) = 13.71.$

In the next three examples you will see that the calculation is a bit tedious when the prior distribution (risk groups) is discrete.

Example 3.3. **[a] An insurance company sells workers compensation policies, each of which belongs to one of three possible risk groups. The risk groups have claim frequency N that are Poisson distributed with parameter λ and claim severity X that are gamma distributed with parameters α and θ. Claim frequency and claim severity are independently distributed given a risk group, and the aggregate loss is S. The data of the risk groups are given in the table below. Suppose the claim experience last year was 26 claims with an average claim size of 12.
Calculate the updated prediction of the **claim frequency**.

Risk Group	Relative frequency	Distribution of N: Poisson($\Lambda = \lambda$)	Distribution of X: gamma (α,θ)
1	0.2	$\lambda = 20$	$\alpha = 5,\ \theta = 2$
2	0.4	$\lambda = 30$	$\alpha = 4,\ \theta = 3$
3	0.4	$\lambda = 40$	$\alpha = 3,\ \theta = 2$

[a] Nonlife Actuarial Models, Example 7.5

Solution. From the Table above, calculate

$$\begin{aligned}
\mathrm{E}(\Lambda) &= (0.2)(20) + (0.4)(30) + (0.4)(40) = 32,\\
\mathrm{E}(\Lambda^2) &= (0.2)(20)^2 + (0.4)(30)^2 + (0.4)(40)^2 = 1,080,\\
\mathrm{Var}(\Lambda) &= \mathrm{E}(\Lambda^2) - [\mathrm{E}(\Lambda)]^2 = 1,080 - (32)^2 = 56.
\end{aligned}$$

Since N is Poisson distributed. The hypothetical mean is $\mathrm{E}(N|\Lambda) = \Lambda$ and the process variance is $\mathrm{Var}(N|\Lambda) = \Lambda$. The expected value of the process variance is $\mu_{\mathrm{PV}} = \mathrm{E}[\mathrm{Var}(N|\Lambda)] = \mathrm{E}(\Lambda) = 32$, the variance of the hypothetical means is $\sigma^2_{\mathrm{HM}} = \mathrm{Var}[\mathrm{E}(N|\Lambda)] = \mathrm{Var}(\Lambda) = 56$, and thus the Bühlmann's k is

$$k = \frac{\mu_{\mathrm{PV}}}{\sigma^2_{\mathrm{HM}}} = \frac{32}{56} = 4/7.$$

We have one year of experience, i.e. $n = 1$, and thus the Bühlmann's credibility factor is

$$Z = \frac{n}{n+k} = \frac{1}{1+4/7} = 7/11.$$

The claim experience last year was 26 claims, i.e. $D = 26$, the overall mean of N is

$$\mu_N = \mathrm{E}[\mathrm{E}(N|\Lambda)] = \mathrm{E}(\Lambda) = 32,$$

i.e. $M = 32$, and thus the updated prediction of the claim frequency is

$$Z\,D + (1-Z)\,M = (7/11)(26) + (4/11)(32) = 28.18.$$

Example 3.4. ** (continued) Calculate the updated prediction of the average **claim size**.

Solution. Denote Γ the vector random variable representing α and θ. The number of expected claims is

$$(0.2)(20) + (0.4)(30) + (0.4)(40) = 4 + 12 + 16 = 32.$$

The relative frequency in the given table is the probability a randomly selected individual came from a given risk group. We need the probability a randomly selected claim comes from a given risk group, which is dependent on the expected number of claims from that risk group. The probabilities of a randomly selected claim coming from the three risk groups are

$$\begin{aligned}
&\Pr(\alpha = 5, \theta = 2) = 0.2(20)/32 = 4/32 = 0.125,\\
&\Pr(\alpha = 4, \theta = 3) = 0.4(30)/32 = 12/32 = 0.375,\\
&\Pr(\alpha = 3, \theta = 2) = 0.4(40)/32 = 16/32 = 0.5,
\end{aligned}$$

respectively. The hypothetical mean is $\mu_X(\Gamma) = \mathrm{E}(X|\Gamma) = \alpha\theta$ and the process variance is $\sigma_X^2(\Gamma) = \mathrm{Var}(X|\Gamma) = \alpha\theta^2$. The conditional means and variances of the severity distributions are organized in the following table:

Group	Prob.	λ	Col 2 $\times$ Col 3	Prob. of X	α	θ	$\mathrm{E}(X\|\Gamma) = \alpha\theta$	$\mathrm{Var}(X\|\Gamma) = \alpha\theta^2$
1	0.2	20	4	0.125	5	2	10	20
2	0.4	30	12	0.375	4	3	12	36
3	0.4	40	6	0.500	3	2	6	12

Calculate

$$\begin{aligned}
\mathrm{E}[X] &= \mathrm{E}[\mathrm{E}(X|\Gamma)] = (0.125)(10) + (0.375)(12) + (0.5)(6) = 8.75,\\
\mathrm{E}\{[\mu_X(\Gamma)]^2\} &= \mathrm{E}[\mathrm{E}(X|\Gamma)^2] = \mathrm{E}[(\alpha\theta)^2] = (0.125)(10)^2 + (0.375)(12)^2 + (0.5)(6)^2 = 84.5.
\end{aligned}$$

The expected value of the process variance is

$$\mu_{\mathrm{PV}} = \mathrm{E}[\mathrm{Var}(X|\Gamma)] = (0.125)(20) + (0.375)(36) + (0.5)(12) = 22,$$

the variance of the hypothetical means is

$$\begin{aligned}
\sigma_{\mathrm{HM}}^2 &= \mathrm{Var}[\mu_X(\Gamma)] = \mathrm{E}\{[\mu_X(\Gamma)]^2\} - \{\mathrm{E}[\mu_X(\Gamma)]\}^2 = \mathrm{E}[\mathrm{E}(X|\Gamma)^2] - \{\mathrm{E}[\mathrm{E}(X|\Gamma)]\}^2\\
&= 84.5 - (8.75)^2 = 7.9375,
\end{aligned}$$

and thus the Bühlmann's k is

$$k = \frac{\mu_{\mathrm{PV}}}{\sigma_{\mathrm{HM}}^2} = \frac{22}{7.9375} = 2.7717.$$

The sample size for claim severity is 26 and thus the credibility factor is

$$Z = \frac{n}{n+k} = \frac{26}{26 + 2.7717} = 0.904.$$

The average claim size is $\bar{X} = 12$, i.e. $D = 12$, and overall mean of X is $\mathrm{E}(X) = 8.75$, i.e. $M = 8.75$, and thus the updated prediction of the claim severity is

$$Z\,D + (1-Z)\,M = (0.904)(12) + (1-0.904)(8.75) = 11.69.$$

Example 3.5. *** (continued) Calculate the updated prediction of the **aggregate loss**.

Solution. Let Θ denote the vector random variable representing Λ and $\Gamma = (\alpha, \theta)$. Since claim frequency and claim severity are independently distributed, the hypothetical mean of S is

$$\mu_S(\Theta) = \mathrm{E}(S|\Theta) = \mathrm{E}(N|\Lambda)\mathrm{E}(X|\Gamma) = \Lambda\alpha\theta,$$

and the process variance of S is

$$\sigma_S^2(\Theta) = \mathrm{Var}(S|\Theta) = \Lambda[\sigma_X^2(\Gamma) + \mu_X^2(\Gamma)] = \Lambda(\alpha\theta^2 + \alpha^2\theta^2) = \Lambda\alpha\theta^2(1+\alpha).$$

The conditional means and variances of S are organized in the following table:

Group	Prob.	(Λ,α,θ)	$\mathrm{E}(S\|\Gamma) = \mu_S(\Theta) = \Lambda\alpha\theta$	$\mathrm{Var}(S\|\Gamma) = \sigma_S^2(\Theta) = \Lambda\alpha\theta^2(1+\alpha)$
1	0.2	(20,5,2)	200	2,400
2	0.4	(30,4,3)	360	5,400
3	0.4	(40,3,2)	240	1,920

Calculate

$$\mathrm{E}(S) = \mathrm{E}[\mathrm{E}(S|\Theta)] = \mathrm{E}(\Lambda\alpha\theta) = (0.2)(200) + (0.4)(360) + (0.4)(240) = 280,$$
$$\mathrm{E}\{[\mu_S(\Theta)]^2\} = \mathrm{E}[E(S|\Theta)^2] = \mathrm{E}[(\Lambda\alpha\theta)^2] = (0.2)(200)^2 + (0.4)(360)^2 + (0.4)(240)^2 = 82,880.$$

The expected value of the process variance is

$$\mu_{\mathrm{PV}} = \mathrm{E}[\sigma_S^2(\Theta)] = \mathrm{E}[\mathrm{Var}(S|\Theta)] = \mathrm{E}[\Lambda\alpha\theta^2(1+\alpha)] = (0.2)(2,400)+(0.4)(5,400)+(0.4)(1,920) = 3,408,$$

the variance of the hypothetical means is

$$\sigma_{\mathrm{HM}}^2 = \mathrm{Var}[\mu_S(\Theta)] = \mathrm{E}\{[\mu_S(\Theta)]^2\} - \{\mathrm{E}[\mu_S(\Theta)]\}^2 = \mathrm{E}[\mathrm{E}(S|\Theta)^2] - \mathrm{E}[\mathrm{E}(S|\Theta)]^2 = 82,880 - (280)^2 = 4,480,$$

and thus the Bühlmann's k is

$$k = \frac{\mu_{\mathrm{PV}}}{\sigma_{\mathrm{HM}}^2} = \frac{3,408}{4,480} = 0.7607.$$

We have one year of experience for aggregate loss, i.e. $n = 1$, and thus the credibility factor is

$$Z = \frac{n}{n+k} = \frac{1}{1+0.7607} = 0.568.$$

The aggregate loss last year is $(26)(12) = 312$, i.e. $D = 312$, and the overall mean of S is $\mathrm{E}(S) = 280$, i.e. $M = 280$, and thus the updated prediction of the aggregate loss is

$$Z\,D + (1-Z)\,M = (0.568)(312) + (1-0.568)(280) = 298.2.$$

Recap: The total variance can be written as $\text{Var}(X) = \mu_{\text{PV}} + \sigma^2_{\text{HM}}$. The Bühlmann's k is $k = \mu_{\text{PV}}/\sigma^2_{\text{HM}}$. The credibility factor is $Z = n/(n+k)$ and the updated prediction is $Z\,D + (1-Z)\,M$.

Example 3.1: Determine the Bühlmann credibility factor Z.

Example 3.2: Determine the Bühlmann credibility factor Z for aggregate losses.

Example 3.3: Determine the updated prediction of the **claim frequency** for a discrete prior.

Example 3.4: Determine the updated prediction of the **average claim size** for a discrete prior.

Example 3.5: Determine the updated prediction of the **aggregate loss** for a discrete prior.

3.1 Problem Set

Question 3.1. * (C 2007 Spring #21) You are given:

(i) Losses in a given year follow a gamma distribution with parameters α and θ, where θ does not vary by policyholder.

(ii) The prior distribution of α has mean 50.

(iii) The Bühlmann credibility factor based on two years of experience is 0.25.

Calculate Var(α).

Question 3.2. *** (C 2006 Spring #6) For a group of policies, you are given:

(i) The annual loss on an individual policy follows a gamma distribution with parameters $\alpha = 4$ and θ.

(ii) The prior distribution of θ has mean 600.

(iii) randomly selected policy had losses of 1400 in Year 1 and 1900 in Year 2.

(iv) Loss data for Year 3 was misfiled and unavailable.

(v) Based on the data in (iii), the Bühlmann credibility estimate of the loss on the selected policy in Year 4 is 1800.

(vi) After the estimate in (v) was calculated, the data for Year 3 was located. The loss on the selected policy in Year 3 was 2763.

Calculate the Bühlmann credibility estimate of the loss on the selected policy in Year 4 based on the data for Years 1, 2 and 3.

Question 3.3. (C 2005 Spring #20) For a particular policy, the conditional probability of the annual number of claims given $\Theta = \theta$, and the probability distribution of Θ are as follows:

Number of Claims	0	1	2
Probability	2θ	θ	$1-3\theta$

θ	0.05	0.30
Probability	0.80	0.20

Two claims are observed in Year 1.

Calculate the Bühlmann credibility estimate of the number of claims in Year 2.

Question 3.4. ** (C 2001 Spring #38) You are given the following information about workers compensation coverage:

(i) The number of claims for an employee during the year follows a Poisson distribution with mean $(100 - \lambda)/100$, where λ is the salary (in thousands) for the employee.

(ii) The distribution of λ is uniform on the interval (0, 100].

An employee is selected at random. During the last 4 years, the employee has had a total of 5 claims.

Determine the Bühlmann credibility estimate for the expected number of claims the employee will have next year.

Question 3.5. *** (C 2000 Spring #37) You are given:

(i) X_i is the claim count observed for driver i for one year.

(ii) X_i has a negative binomial distribution with parameters $\beta = 0.5$ and r_i.

(iii) μ_i is the expected claim count for driver i for one year.

(iv) The μ_i's have an exponential distribution with mean 0.2.

Determine the Bühlmann credibility factor for an individual driver for one year.

Question 3.6. * (4B 1990 Spring #35) The underlying expected loss for each individual insured is assumed to be constant over time. The Bühlmann credibility factor assigned to the pure premium for an insured observed for one year is 1/2.

Determine the Bühlmann credibility factor to be assigned to the pure premium for an insured observed for 3 years.

Question 3.7. * (4B 1991 Spring #25) Assume that the expected pure premium for an individual insured is constant over time. The Bühlmann credibility factor for two years of experience is equal to 0.4.

Determine the Bühlmann credibility factor for three years of experience.

Question 3.8. * (4B 1996 Spring #3) Given a first observation with a value of 2, the Bühlmann credibility estimate for the expected value of the second observation would be 1. Given a first observation with a value of 5, the Bühlmann credibility estimate for the expected value of the second observation would be 2.

Determine the Bühlmann credibility factor.

Question 3.9. * (4B 1996 Fall #10) The Bühlmann credibility factor of n observations of the loss experience of a single risk is 1/3. The Bühlmann credibility factor of $n + 1$ observations of the loss experience of a single risk is 2/5.

Determine the Bühlmann credibility factor of $n + 2$ observations of the loss experience of this risk.

Question 3.10. ** (4B 1996 Fall #4) You are given the following:

(i) A portfolio of independent risks is divided into three classes.

(ii) Each class contains the same number of risks.

(iii) For each risk in Classes 1 and 2, the probability of exactly one claim during one exposure period is 1/3, while the probability of no claim is 2/3.

(iv) For each risk in Class 3, the probability of exactly one claim during one exposure period is 2/3, while the probability of no claim is 1/3.

A risk is selected at random from the portfolio. During the first two exposure periods, two claims are observed for this risk (one in each exposure period).

Determine the Bühlmann credibility estimate of the probability that a claim will be observed for this same risk during the third exposure period.

Question 3.11. *** (4B 1996 Fall #20) You are given the following:

(i) The number of claims for a single risk follows a Poisson distribution with mean $\theta\mu$.

(ii) θ and μ have a prior probability distribution with joint density function:
$f(\theta, \mu) = 1, \quad 0 < \theta < 1, \quad 0 < \mu < 1.$

Determine the value of Bühlmann's k.

Question 3.12. * (4B 1998 Spring #2) You are given the following:

(i) The number of claims for a single insured follows a Poisson distribution with mean λ.

(ii) λ varies by insured and follows a Poisson distribution with mean μ.

Determine the value of Bühlmann's k.

Question 3.13. * (4B 1990 Spring #52) The number of claims each year for an individual insured has a Poisson distribution. The expected annual claim frequency of the entire population of insureds is uniformly distributed over the interval (0,1). An individual expected claim frequency is constant through time. A particular insured had 3 claims during the prior three years.

Determine the Bühlmann credibility estimate of this insured's future annual claim frequency.

Question 3.14. * (4B 1998 Fall #21) You are given the following:

(i) The number of claims for a single risk follows a Poisson distribution with mean λ.

(ii) The amount of an individual claim is always 1000λ.

(iii) λ is a random variable with the density function: $f(\lambda) = \dfrac{4}{\lambda^5}$, $1 < \lambda < \infty$.

Determine the expected value of the process variance of the aggregate losses for a single risk.

Question 3.15. *** (4B 1998 Spring #7) You are given the following:

(i) The number of claims during one exposure period follows a Bernoulli distribution with mean p.

(ii) The prior density function of p is assumed to be

$$f(p) = \frac{\pi}{2}\sin\frac{\pi p}{2}, \qquad 0 < p < 1.$$

(iii) Hint: $\int_0^1 \frac{\pi p}{2}\sin\frac{\pi p}{2}\mathrm{d}p = \frac{2}{\pi}$ and $\int_0^1 \frac{\pi p^2}{2}\sin\frac{\pi p}{2}\mathrm{d}p = \frac{4(\pi-2)}{\pi^2}$

Determine the expected value of the process variance.

Question 3.16. *** (4B 1998 Spring #26) You are given the following:

(i) The number of claims follows a Poisson distribution with mean m.

(ii) Claim sizes follow a distribution with mean $20m$ and variance $400m^2$.

(iii) m is a gamma random variable with density function $f(m) = (0.5)m^2e^{-m}$, $0 < m < \infty$.

(iv) For any value of m, the number of claims and the claim sizes are independent.

Determine the expected value of the process variance of the aggregate losses.

Question 3.17. *** (C 2007 Spring #6) An insurance company sells two types of policies with the following characteristics:

Type of Policy	Proportion of Total Policies	Poission Annual Claim Frequency
I	θ	$\lambda = 0.5$
II	$1 - \theta$	$\lambda = 1.5$

A randomly selected policyholder is observed to have one claim in Year 1.

For the same policyholder, determine the Bühlmann credibility factor Z for Year 2.

A. $\dfrac{\theta - \theta^2}{1.5 - \theta^2}$

B. $\dfrac{1.5 - \theta}{1.5 - \theta^2}$

C. $\dfrac{2.25 - 2\theta}{1.5 - \theta^2}$

D. $\dfrac{2\theta - \theta^2}{1.5 - \theta^2}$

E. $\dfrac{2.25 - 2\theta^2}{1.5 - \theta^2}$

Question 3.18. *** (C 2005 Spring #32) You are given:

(i) The number of claims in a year for a selected risk follows Poisson distribution with mean λ.

(ii) The severity of claims for the selected risk follows exponential distribution with mean θ.

(iii) The number of claims is independent of the severity of claims.

(iv) The prior distribution of λ is exponential with mean 1.

(v) The prior distribution of θ is Poisson with mean 1.

(vi) A priori, λ and θ are independent.

Using Bühlmann's credibility for aggregate losses, determine k.

Question 3.19. ** (4 2000 Fall #19) For a portfolio of independent risks, you are given:

(i) The risks are divided into two classes, Class A and Class B.

(ii) Equal numbers of risks are in Class A and Class B.

(iii) For each risk, the probability of having exactly 1 claim during the year is 20% and the probability of having 0 claims is 80%.

(iv) All claims for Class A are of size 2.

(v) All claims for Class B are of size c, an unknown but fixed quantity.

One risk is chosen at random, and the total loss for one year for that risk is observed. You wish to estimate the expected loss for that same risk in the following year. Determine the limit of the Bühlmann credibility factor as c goes to infinity.

Question 3.20. ** (C 2005 Fall #19) For a portfolio of independent risks, the number of claims for each risk in a year follows a Poisson distribution with means given in the following table:

Class	Mean Number of Claims per Risk	Number of Risks
1	1	900
2	10	90
3	20	10

You observe x claims in Year 1 for a randomly selected risk. The Bühlmann credibility estimate of the number of claims for the same risk in Year 2 is 11.983.

Determine x.

Question 3.21. ** (C 2006 Fall #6) For a group of policies, you are given:

(i) The annual loss on an individual policy follows a gamma distribution with parameters $\alpha = 4$ and θ.

(ii) The prior distribution of θ has mean 600.

(iii) A randomly selected policy had losses of 1400 in Year 1 and 1900 in Year 2.

(iv) Loss data for Year 3 was misfiled and unavailable.

(v) Based on the data in (iii), the Bühlmann credibility estimate of the loss on the selected policy in Year 4 is 1800.

(vi) After the estimate in (v) was calculated, the data for Year 3 was located. The loss on the selected policy in Year 3 was 2763.

Calculate the Bühlmann credibility estimate of the loss on the selected policy in Year 4 based on the data for Years 1, 2 and 3.

Question 3.22. ** (C 2005 Spring #6) You are given:

(i) Claims are conditionally independent and identically Poisson distributed with mean Θ.

(ii) The prior distribution function of Θ is:

$$F(\theta) = 1 - \left(\frac{1}{1+\theta}\right)^{2.6}, \ \theta > 0$$

(iii) Five claims are observed.

Determine the Bühlmann credibility factor for predicting the claim severity.

Question 3.23. ** (C 2005 Fall #7) For a portfolio of policies, you are given:

(i) The annual claim amount on a policy has probability density function:

$$f(x|\theta) = \frac{2x}{\theta^2}, \quad 0 < x < \theta$$

(ii) The prior distribution of θ has density function:

$$\pi(\theta) = 4\theta^3, \quad 0 < \theta < 1$$

(iii) A randomly selected policy had claim amount 0.1 in Year 1.

Determine the Bühlmann credibility estimate of the claim amount for the selected policy in Year 2.

Question 3.24. ** (4 2003 Fall #11) You are given:

(i) Claim counts follow a Poisson distribution with mean λ.

(ii) Claim sizes follow an exponential distribution with mean 10λ.

(iii) Claim counts and claim sizes are independent, given λ.

(iv) The prior distribution has probability density function:

$$\pi(\lambda) = \frac{5}{\lambda^6}, \ \lambda > 1$$

Calculate Bühlmann's k for aggregate losses.

Question 3.25. ** (C 2005 Spring #17) You are given

(i) The annual number of claims on a given policy has a geometric distribution with parameter β.

(ii) The prior distribution of β has the Pareto density function

$$\pi(\beta) = \frac{\alpha}{(\beta+1)^{\alpha+1}}, \qquad 0 < \beta < \infty.$$

where α is a known constant greater than 2.

A randomly selected policy had x claims in Year 1.

Determine the Bühlmann credibility estimate of the number of claims for the selected policy in Year 2.

The following questions are from past MAS-II exams.

Question 3.26. (MAS-II 2018 Fall #3) An insurance company sells homeowners' policies, each of which belongs to one of two possible risk groups, S and T. You are given the following information:

(i) Risk group S occurs 20% of the time.

(ii) Risk group S has claim frequencies that are Poisson distributed with parameter $\lambda = 2$.

(iii) Risk group S has claim severity that is uniformly distributed between 100 and 1000.

(iv) Risk group T occurs 80% of the time.

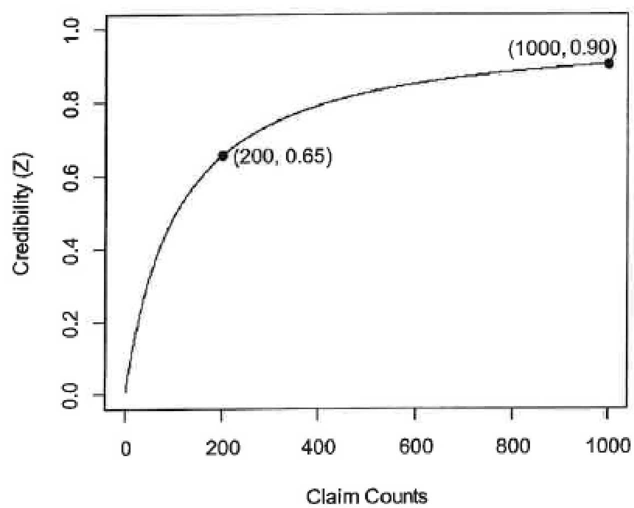
(v) Risk group T has claim frequencies that are Poisson distributed with parameter $\lambda = 1$.

(vi) Risk group T has claim severity that is uniformly distributed between 2000 and 8000.

(vii) Claim frequency and claim severity are independently distributed given a risk group.

The Bühlmann credibility method is used to calculate the next year's predicted aggregate loss given three prior years of loss experience for a given risk.
Calculate the Bühlmann credibility factor for this risk.

Question 3.27. (MAS-II 2019 Fall #3) An actuary is constructing a credibility formula to apply to claim severity. To visualize the partial credibility, a plot showing credibility as a function of claim counts is created. Determine which credibility method the actuary is using.



A. Limited-fluctuation credibility with a full-credibility standard of 1300 claims.

B. Limited-fluctuation credibility with a full-credibility standard of 800 claims.

C. Limited-fluctuation credibility with $\lambda_F = 541$ and $C_x = 1.053$.

D. Bühlmann credibility with EVPV = 0.266 and VHM = 0.016.

E. Bühlmann credibility with EVPV = 0.844 and VHM = 0.008.

Question 3.28. (MAS-II 2019 Fall #6) Policies belong to one of two possible risk groups, Risk Group R and Risk Group S. You are given the following information:

(i) 40% of policies belong to Risk Group R.

(ii) Risk Group R has claim frequencies that are Poisson distributed with $\lambda = 3$.

(iii) Risk Group R has claim severity that is uniformly distributed between 3000 and 5000.

(iv) 60% of policies belong to Risk Group S.

(v) Risk Group S has claim frequencies that are Poisson distributed with $\lambda = 1$.

(vi) Risk Group S has claim severity that is uniformly distributed between 50 and 500.

(vii) Claim frequency and claim severity are independently distributed given a risk group.

3.2 Problem Set Solutions

Question 3.1. The expected value of process variance is

$$\mu_{\text{PV}} = \text{E}[\text{Var}(X|\alpha)] = \text{E}(\alpha\theta^2) = 50\theta^2$$

and the variance of hypothetical means is

$$\sigma^2_{\text{HM}} = \text{Var}[\text{E}(X|\alpha)] = \text{Var}(\alpha\theta) = \text{Var}(\alpha)\theta^2.$$

Since the Bühlmann credibility factor based on two years of experience is 0.25. We have

$$Z = \frac{n}{n + \frac{\mu_{\text{PV}}}{\sigma^2_{\text{HM}}}} \quad \Rightarrow \quad 0.25 = \frac{2}{2 + \frac{50\theta^2}{\text{Var}(\alpha)\theta^2}} \quad \Rightarrow \quad \text{Var}(\alpha) \approx 8.3.$$

Question 3.2. The prior mean is

$$\mu = \text{E}(X) = \text{E}[\text{E}(X|\Theta)] = \text{E}[4\Theta] = 4(600) = 2400.$$

For the first calculation (based on the first two observations):

$$1800 = Z\bar{X}_2 + (1-Z)\mu = Z(1650) + (1-Z)2400 \implies Z = 0.8 = \frac{n}{n+k} \implies k = 0.5$$

k is a parameter of the model, not the data, so it is the same for the second calculation based on all three observations.

Hence

$$\begin{aligned} Z &= \frac{3}{3+0.5} = \frac{6}{7}, \\ \bar{X}_3 &= \frac{1400+1900+2763}{3} = 2021, \\ Z\bar{X}_3 + (1-Z)\mu &= \frac{6}{7}(2021) + \frac{1}{7}2400 = 2075.14. \end{aligned}$$

Question 3.3. Make the following tables:

Number of claims k	0	1	2
$\Pr(N=k\|\theta=0.05)$	0.10	0.05	0.85
$\Pr(N=k\|\theta=0.30)$	0.60	0.30	0.10

$\Pr(\Theta=\theta)$	0.80	0.20
$\text{E}(N\|\theta)$	1.75	0.50
$\text{E}(N^2\|\theta)$	3.45	0.70
$\text{Var}(N\|\theta)$	0.3875	0.45

The expected value of the process variance is

$$\mu_{\text{PV}} = \text{E}[\text{Var}(N|\Theta)] = 0.8(0.3875)) + 0.2(0.45) = 0.4.$$

The variance of hypothetical mean is (Bernoulli variance formula)

$$\sigma^2_{\mathrm{HM}} = 0.8(0.2)(1.75-0.5)^2 = 0.25,$$

and hence $k = \mu_{\mathrm{PV}}/\sigma^2_{\mathrm{HM}} = 1.6$ and the credibility factor is $Z = n/(n+k) = 1/2.6 = 0.3846$. The sample mean is $\bar{X} = 2$, the unconditional mean is $\mu = 0.8(1.75) + 0.2(0.50) = 1.5$, and hence the Bühlmann credibility estimate is

$$Z\bar{X} + (1-Z)\mu = (0.3846)(2) + (1-0.3846)(1.5) = 1.692.$$

Question 3.4. The Bühlmann estimate is a weighted average of the sample mean

$$\bar{X} = \frac{5}{4} = 1.25$$

and the population mean

$$\mathrm{E}(X) = \mathrm{E}\left[\mathrm{E}(X|\Lambda)\right] = \mathrm{E}\left[\frac{100-\Lambda}{100}\right] = \frac{100-100/2}{100} = 0.5$$

where the weight Z is calculate from

$$\mu_{\mathrm{PV}} = \mathrm{E}\left[\mathrm{Var}(X|\Lambda)\right] = \mathrm{E}\left[\frac{100-\Lambda}{100}\right] = 0.5,$$
$$\sigma^2_{\mathrm{HM}} = \mathrm{Var}\left[\mathrm{E}(X|\Lambda)\right] = \mathrm{Var}\left[\frac{100-\Lambda}{100}\right] = \frac{\mathrm{Var}(\Lambda)}{10,000} = \frac{1}{12},$$
$$Z = \frac{n}{n+\mu_{\mathrm{PV}}/\sigma^2_{\mathrm{HM}}} = \frac{4}{4+6} = 0.4.$$

Hence, the Bühlmann credibility estimate is

$$Z\bar{X} + (1-Z)\mathrm{E}[X] = (0.4)(1.25) + (0.6)(0.5) = 0.8.$$

Question 3.5. The risk classification parameter is just r. The distribution of $X|r$ is negative binomial so

$$\mathrm{E}(X|r) = r\beta = 0.5r,$$

$$\mathrm{Var}(X|r) = r\beta(1+\beta) = 0.75r.$$

Since μ_i is the expected claim count for driver i for one year, i.e. $\mu_i = 0.5r_i$, then r is exponential distributed with mean $(2)(0.2) = 0.4$. The expected value of the process variance is

$$\mu_{\mathrm{PV}} = \mathrm{E}(0.75r) = 0.75\mathrm{E}(r) = 0.75(0.4) = 0.3.$$

The variance of the hypothetical means is

$$\sigma^2_{\mathrm{HM}} = \mathrm{Var}(0.5r) = (0.25)(0.4)^2 = 0.04.$$

This gives us $k = 0.3/0.04 = 7.5$, and a Bühlmann credibility factor of $Z = 1/(1+k) = 0.1176$.

Question 3.6. For one year of experience, $n = 1$, we have $Z = n/(n+k) = 1/(1+k) = 1/2$. Hence $k = 1$. For three years of experience, $n = 3$, we have $Z = n/(n+k) = 3/(3+k) = 3/4$.

Question 3.7. For two years of experience, we are given $n = 2$ and $Z = n/(n+k) = 2/(2+k) = 0.4$. Hence $k = 3$. For three years of experience, $n = 3$, the Bühlmann credibility factor is $Z = n/(n+k) = 3/(3+3) = 0.5$.

Question 3.8. The Bühlmann credibility estimate is

$$P_c = Z\bar{X} + (1-Z)\mu.$$

We are given

$$\begin{aligned} 1 &= Z(2) + (1-Z)\mu, \\ 2 &= Z(5) + (1-Z)\mu. \end{aligned}$$

Solving the system equations we have $Z = 1/3$.

Question 3.9. For n observations, we are given

$$1/3 = n/(n+k).$$

For $n+1$ observations, we are given

$$2/5 = (n+1)/(n+1+k).$$

Solving the system equations we have $n = 3$ and $k = 6$. Hence, the credibility factor of $n + 2 = 5$ observations is $Z = 5/(5+6) = 5/11$.

Question 3.10. In a Bernoulli distribution, the probability of $X = 1$ is the expected value. Hence the question is simply asking the credibility estimate. The Bühlmann estimate is a weighted average of the sample mean

$$\bar{X} = \frac{2}{2} = 1$$

and the population mean

$$\mathrm{E}(X) = \mathrm{E}[\mathrm{E}(X|C)] = (1/3 + 1/3 + 2/3)/3 = 4/9$$

where the weight Z is calculate from

$$\begin{aligned} \mu_{\mathrm{PV}} &= \mathrm{E}[\mathrm{Var}(X|C)] = [(1/3)(2/3) + (1/3)(2/3) + (2/3)(1/3)]/3 = 2/9, \\ \sigma^2_{\mathrm{HM}} &= \mathrm{Var}[\mathrm{E}(X|C)] = (1/3)^2(1/3) + (1/3)^2(1/3) + (2/3)^2(1/3) - (4/9)^2 = 2/81, \\ Z &= \frac{n}{n + \mu_{\mathrm{PV}}/\sigma^2_{\mathrm{HM}}} = \frac{2}{2+9} = 2/11. \end{aligned}$$

Hence, the Bühlmann credibility estimate is

$$Z\bar{X} + (1-Z)\mathrm{E}[X] = (2/11)(1) + (1-2/11)(4/9) = 6/11.$$

Question 3.11. First recognize that the prior parameters λ and μ are independent and uniformly distributed. Hence

$$\begin{aligned} \mathrm{E}(\theta\mu) &= \mathrm{E}(\theta)\mathrm{E}(\mu) = (1/2)(1/2) = 1/4, \\ \mathrm{E}[(\theta\mu)^2] &= \mathrm{E}(\theta^2)\mathrm{E}(\mu^2) = (1/3)(1/3) = 1/9, \\ \mathrm{Var}(\lambda\mu)] &= 1/9 - (1/4)^2 = 7/144. \end{aligned}$$

Therefore,

$$\begin{aligned} \mu_{\mathrm{PV}} &= \mathrm{E}[\mathrm{Var}(X|\theta\mu)] = \mathrm{E}(\theta\mu) = 1/4, \\ \sigma^2_{\mathrm{HM}} &= \mathrm{Var}[\mathrm{E}(X|\theta\mu)] = \mathrm{Var}(\theta\mu) = 7/144, \\ k &= \frac{\mu_{\mathrm{PV}}}{\sigma^2_{\mathrm{HM}}} = \frac{1/4}{7/144} = 5.14. \end{aligned}$$

Question 3.12. Calculate

$$\begin{aligned} \mu_{\mathrm{PV}} &= \mathrm{E}[\mathrm{Var}(X|\lambda)] = \mathrm{E}(\lambda) = \mu \\ \sigma^2_{\mathrm{HM}} &= \mathrm{Var}[\mathrm{E}(X|\lambda)] = \mathrm{Var}(\lambda) = \mu \\ k &= \frac{\mu_{\mathrm{PV}}}{\sigma^2_{\mathrm{HM}}} = \frac{\mu}{\mu} = 1. \end{aligned}$$

Question 3.13. λ is uniformly distributed. The expected value of the process variance is

$$\mu_{\mathrm{PV}} = \mathrm{E}[\mathrm{Var}(N|\lambda)] = \mathrm{E}(\lambda) = 1/2.$$

The variance of hypothetical mean is

$$\sigma^2_{\mathrm{HM}} = \mathrm{Var}[\mathrm{E}(N|\lambda)] = \mathrm{Var}(\lambda)) = 1/12.$$

Hence $k = \mu_{\mathrm{PV}}/\sigma^2_{\mathrm{HM}} = 6$ and the credibility factor is

$$Z = n/(n+k) = 3/(3+6) = 1/3.$$

The sample mean is $\bar{X} = 3/3 = 1$, the unconditional mean is $\mu = \mathrm{E}[\mathrm{E}(N|\lambda)] = \mathrm{E}(\lambda) = 1/2$, and hence the Bühlmann credibility estimate is

$$Z\bar{X} + (1-Z)\mu = (1/3)(1) + (1-1/3)(1/2) = 2/3.$$

Question 3.14. Since the amount of an individual claim is always 1000λ, the aggregate losses can be written as $S = X_1 + \cdots + X_N = N(1000\lambda)$ where N is Poisson distributed with mean λ. The prior distribution is a single-parameter Pareto with $\alpha = 4$ and $\theta = 1$. From the Exam Tables, calculate the expected value of the process variance

$$\mu_{\mathrm{PV}} = \mathrm{E}[\mathrm{Var}(S|\lambda)] = \mathrm{E}[\mathrm{Var}(1000\lambda N|\lambda)] = \mathrm{E}(10^6\lambda^2\mathrm{Var}(N)) = 10^6\mathrm{E}(\lambda^3) = 10^6\frac{4}{4-3} = 4{,}000{,}000.$$

Question 3.15. The process variance of the given Bernoulli is $p(1-p)$. The expected value of the process variance is

$$\mu_{\text{PV}} = \text{E}[\text{Var}(N|p)] = \text{E}[p(1-p)] = \text{E}(p) - \text{E}(p^2) = \frac{2}{\pi} - \frac{4(\pi-2)}{\pi^2} = \frac{2(4-\pi)}{\pi^2}.$$

Question 3.16. The aggregate losses follow a compound Poisson model. First calculate

$$\text{E}(X^2|m) = [\text{E}(X|m)]^2 + \text{Var}(X|m) = (20m)^2 + 400m^2 = 800m^2.$$

Using Eq. (3.0.13), the process variance is

$$\text{Var}(S|m) = \text{E}(N|m)\text{E}(X^2|m) = 800m^3.$$

Observe that m is gamma distributed with parameters $\alpha = 3$ and $\theta = 1$ and hence $\text{E}(m^3) = (\alpha+2)(\alpha+1)(\alpha) = (3+2)(3+1)(3) = 60$. Therefore, the expected value of the the process variance is

$$\text{E}[\text{Var}(S|m)] = \text{E}(800m^3) = (800)(60) = 48,000.$$

Question 3.17. The expected value of the process variance is

$$\text{EPV} = \text{E}[\text{Var}(X|\Lambda)] = \theta(0.5) + (1-\theta)(1.5) = 1.5 - \theta.$$

The variance of the hypothetical means is

$$\begin{aligned}\text{VHM} &= \text{Var}[\text{E}(X|\Lambda)] = \text{E}\{[\text{E}(X|\Lambda)]^2\} - \{\text{E}[\text{E}(X|\Lambda)]\}^2 \\ &= \theta(0.5)^2 + (1-\theta)(1.5)^2 - (1.5-\theta)^2 = \theta - \theta^2.\end{aligned}$$

Therefore,

$$\begin{aligned}k &= \frac{v}{a} = \frac{1.5-\theta}{\theta-\theta^2}, \\ z &= \frac{n}{n+k} = \frac{n}{n + \dfrac{1.5-\theta}{\theta-\theta^2}} = \frac{\theta-\theta^2}{1.5-\theta^2}. \qquad \text{(A)}\end{aligned}$$

Question 3.18. Denote the frequency and severity by $N|\lambda$ and $X|\theta$, respectively, and the aggregate by $S|\lambda,\theta$.

$$\begin{aligned}\text{EPV} &= \text{E}[\text{Var}(S|\lambda,\theta)] = \text{E}[\text{E}(N|\lambda)\text{E}(X^2|\theta)] = \text{E}[\lambda(2\theta^2)] = 2\text{E}(\lambda)\text{E}(\theta^2) = 2(\text{Var}(\theta) + \text{E}(\theta)^2) = 4 \\ \text{VHM} &= \text{Var}[\text{E}(S|\lambda,\theta)] = \text{Var}[\text{E}(N|\lambda)\text{E}(X|\theta)] = \text{Var}(\lambda\theta) = \text{E}(\lambda^2\theta^2) - \text{E}(\lambda\theta)^2 \\ &= \text{E}(\lambda^2)\text{E}(\theta^2) - \text{E}(\lambda)^2\text{E}(\theta)^2 = (2)(1+1) - 1 = 3 \\ k &= \text{EPV}/\text{VHM} = 4/3\end{aligned}$$

Question 3.19. For each class the aggregate $S|\theta$ is just a multiple of the number of claims $N|\theta$ since the severity is constant within the class.

Class θ	$\pi(\theta)$	$\mathrm{E}(N\|\theta)$	$\mathrm{E}(X\|\theta)$	$\mathrm{E}(S\|\theta)$	$\mathrm{Var}(S\|\theta)$
A	0.5	0.2	2	0.4	0.48
B	0.5	0.2	c	$0.2c$	$0.16c^2$

The expected value of the process variance is

$$\mathrm{EPV} = (0.5)(0.48) + (0.5)(0.16c^2) = 0.24 + 0.08c^2$$

and the variance of the hypothetical mean is

$$\mathrm{VPM} = (0.5)^2(0.2c - 0.4)^2 = 0.01(c^2 + 2c + 4).$$

Therefore,

$$\lim_{c\to\infty} k = \lim_{c\to\infty} \frac{0.24 + 0.08c^2}{0.01(c^2 + 2c + 4)} = 8$$

and $Z = \dfrac{1}{1+8} = \dfrac{1}{9}$.

Question 3.20. For the Poisson distribution, $\mathrm{E}(X|\Theta) = \mathrm{Var}(X|\Theta)$. We have

$$\begin{aligned}
\mu &= \mathrm{E}[\mathrm{E}(X|\Theta)] = (1)(0.9) + (10)(0.09) + (20)(0.01) = 2\\
\mathrm{EPV} &= \mathrm{E}[\mathrm{Var}(X|\Theta)] = \mathrm{E}[\mathrm{E}(X|\Theta)] = 2\\
\mathrm{VHM} &= \mathrm{Var}[\mathrm{E}(X|\Theta)] = (1-2)^2(0.9) + (10-2)^2(0.09) + (20-2)^2(0.01) = 9.9\\
k &= \frac{\mathrm{EPV}}{\mathrm{VHM}} = \frac{2}{9.9}\\
Z &= \frac{1}{1 + 2/9.9} = \frac{9.9}{11.9}\\
P_c &= Zx + (1-Z)\hat{\mu}\\
\Longrightarrow 11.983 &= \left(\frac{9.9}{11.9}\right)(x) + \left(\frac{2}{11.9}\right)(2) \implies x = 14
\end{aligned}$$

Question 3.21. The prior mean is:

$$\mu = \mathrm{E}(X) = \mathrm{E}[\mathrm{E}(X|\Theta)] = \mathrm{E}[4\Theta] = 4(600) = 2400$$

For the first calculation (based on the first two observations):

$$1800 = Z\bar{X}_2 + (1-Z)\mu = Z(1650) + (1-Z)2400 \quad \Rightarrow \quad Z = 0.8 = \frac{n}{n+k} \quad \Rightarrow k = 0.5$$

k is a parameter of the model, not the data, so it is the same for the second calculation based on all three observations.

$$\begin{aligned} Z &= \frac{3}{3+0.5} = \frac{6}{7} \\ \bar{X}_3 &= \frac{1400+1900+2763}{3} = 2021 \\ Z\bar{X}_3 + (1-Z)\mu &= \frac{6}{7}(2021) + \frac{1}{7}2400 = 2075 \end{aligned}$$

Question 3.22. $N|\Theta$ is Poisson with mean Θ following the Pareto with $\alpha = 2.6$ and scale factor 1. Using the Exam Tables, calculate

$$\begin{aligned} \mathrm{E}(\Theta) &= \frac{1}{\alpha-1} = \frac{1}{1.6} = 0.625, \\ \mathrm{E}(\Theta^2) &= \frac{1}{(\alpha-1)(\alpha-2)} = 2.08333, \\ \mathrm{Var}(\Theta) &= 2.08333 - (0.625)^2 = 1.6927. \end{aligned}$$

The credibility factor is $Z = \dfrac{n}{n+k} = \dfrac{5}{5+k}$. The k value is $k = \dfrac{\mathrm{EPV}}{\mathrm{VHM}}$ where

$$\begin{aligned} \mathrm{EPV} &= \mathrm{E}[\mathrm{Var}(N|\Theta)] = \mathrm{E}[\Theta] = 0.625, \\ \mathrm{VHM} &= \mathrm{Var}[\mathrm{E}(N|\Theta)] = \mathrm{Var}(\Theta) = 1.6927, \\ k &= 0.625/1.6927 = 0.36923, \\ Z &= n/(n+k) = 5/(5+0.36923) = 0.931. \end{aligned}$$

Question 3.23. First calculate $\mathrm{E}(\Theta) = \dfrac{4}{5}$ and $\mathrm{E}(\Theta^2) = \dfrac{2}{3}$.

$$\begin{aligned} \mathrm{E}(X|\theta) &= \int_0^\theta \frac{2x^2}{\theta^2}\mathrm{d}x = \frac{2\theta}{3} \\ \mathrm{E}(X^2|\theta) &= \int_0^\theta \frac{2x^3}{\theta^2}\mathrm{d}x = \frac{\theta^2}{2} \implies \mathrm{Var}(X|\theta) = \mathrm{E}(X^2|\theta) - [\mathrm{E}(X|\theta)]^2 = \frac{\theta^2}{18} \\ \mu &= \mathrm{E}[\mathrm{E}(X|\Theta)] = \mathrm{E}\left[\frac{2\Theta}{3}\right] = \frac{8}{15} \\ v &= \mathrm{E}[\mathrm{Var}(X|\Theta)] = \mathrm{E}\left[\frac{\Theta^2}{18}\right] = \frac{1}{27} \\ a &= \mathrm{Var}[\mathrm{E}(X|\Theta)] = \mathrm{Var}\left[\frac{2\Theta}{3}\right] = \frac{4}{9}\left[\frac{2}{3} - \left(\frac{4}{5}\right)^2\right] = \frac{8}{675} \implies k = \frac{v}{a} = \frac{25}{8} \\ Z &= \frac{n}{n+k} = \frac{1}{1+\frac{25}{8}} = \frac{8}{33} \implies P_c = Z\bar{X} + (1-Z)\mu = \frac{8}{33}(0.1) + \frac{25}{33}\frac{8}{15} = 0.4283 \end{aligned}$$

Question 3.24. The prior distribution is a single-parameter Pareto with $\alpha = 5$ and $\theta = 1$. Using the Exam Tables, calculate

$$\begin{aligned}
E(\theta) &= \frac{\alpha}{\alpha - 1} = 5/4, \\
E(\theta^2) &= \frac{\alpha}{\alpha - 2} = 5/3, \\
E(\theta^3) &= \frac{\alpha}{\alpha - 3} = 5/2, \\
E(\theta^4) &= \frac{\alpha}{\alpha - 4} = 5.
\end{aligned}$$

The expected value of the process variance is

$$v = E(\text{Var}(S)] = E(200\theta^3) = (200)(5/2) = 500.$$

The variance of the hypothetical means is

$$a = \text{Var}[E(S)] = \text{Var}[10\theta^2] = 100\{E(\theta^4) - [E(\theta^2)]^2\} = 100[5 - (5/3)^2] = 2000/9.$$

Therefore,

$$k = \frac{v}{a} = \frac{500}{2000/9} = 2.25.$$

Question 3.25. The conditional number $N|\beta$ is geometric and β is Pareto with α and scale parameter 1.

$$\begin{aligned}
E(N) &= E[E(N|\beta)] = E(\beta) = \frac{1}{\alpha - 1} \\
E[\text{Var}(N|\beta)] &= E[\beta(1+\beta)] = E(\beta) + E(\beta^2) = \frac{1}{\alpha - 1} + \frac{2}{(\alpha-1)(\alpha-2)} = \frac{\alpha}{(\alpha-1)(\alpha-2)} \\
\text{Var}[E(N|\beta)] &= \text{Var}[\beta] = E(\beta^2) - E(\beta)^2 \\
&= \frac{2}{(\alpha-1)(\alpha-2)} - \frac{1}{(\alpha-1)^2} = \frac{\alpha}{(\alpha-1)^2(\alpha-2)} \\
k &= \alpha - 1 \\
Z &= \frac{1}{\alpha} \\
Z\bar{X} + (1-Z)\mu &= \frac{1}{\alpha}x + \frac{\alpha-1}{\alpha}\frac{1}{\alpha-1} = \frac{x+1}{\alpha}
\end{aligned}$$

Question 3.26. Conditionally (if we know the risk group), the aggregate loss, $X = Y_1 + Y_2 + \cdots + Y_N$, follows a compound Poisson distribution. The claim frequency N is Poisson distributed with parameter λ. The claim severity Y_i is uniform distributed in $[a, b]$. Let Y be the common distribution of Y_i, $i = 1, \cdots, N$. The mean of the compound Poisson is $\mathrm{E}(X) = \lambda\mathrm{E}(Y)$ and the variance is $\mathrm{Var}(X) = \lambda\mathrm{E}(Y^2)$. Note that $\mathrm{E}(Y) = (a+b)/2$ and $\mathrm{E}(Y^2) = (a^2+ab+b^2)/3$.

Hence, the hypothetical mean is

$$\mathrm{E}(X|R) = \lambda\mathrm{E}(Y) = \lambda(a+b)/2 = \begin{cases} (2)(100+1,000)/2 = 1,100, & R = S, \\ (1)(2,000+8,000)/2 = 5,000, & R = T. \end{cases}$$

The process variance is

$$\mathrm{Var}(X|R) = \lambda\mathrm{E}(Y^2) = \lambda(a^2+ab+b^2)/3 = \begin{cases} (2)(100^2 + 100 * 1000 + 1000^2)/3 = 740,000, & R = S \\ (1)(2000^2 + 2000 * 8000 + 8000^2)/3 = 28,000,000, & R = T. \end{cases}$$

We have $\Pr(R = S) = 0.2$ and $\Pr(R = T) = 0.8$. The expected value of the process variance is

$$\mu_{\mathrm{PV}} = \mathrm{E}[\mathrm{Var}(X|R)] = (0.2)(740,000) + (0.8)(28,000,000) = 22,548,000,$$

the variance of the hypothetical means is

$$\sigma^2_{\mathrm{HM}} = \mathrm{Var}[\mathrm{E}(X|R)] = (0.2)(0.8)(5,000 - 1,100)^2 = 2,433,600,$$

and hence

$$k = \frac{\mu_{\mathrm{PV}}}{\sigma^2_{\mathrm{HM}}} = \frac{22,548,000}{2,433,600}.$$

The credibility factor is

$$Z = \frac{n}{n+k} = \frac{3}{3 + 22,548,000/2,433,600} = 0.2446. \qquad \text{(Answer: D)}$$

Question 3.27. The limited-fluctuation credibility is

$$Z = \sqrt{\frac{n}{\lambda_F\, C_X^2}}$$

and the Bühlmann credibility is

$$Z = \frac{n}{n + \mathrm{EVPV}/\mathrm{VHM}}.$$

Answer A is not correct. The standard for full-credibility is $\lambda_F\, C_X^2 = 1300$. The point $(n, Z) = (200, 0.65)$ doesn't fit the formula:

$$Z = \sqrt{\frac{n}{\lambda_F\, C_X^2}} \quad \text{but} \quad 0.65 \neq \sqrt{\frac{200}{1300}}$$

Answer B is not correct. The standard for full-credibility is $\lambda_F C_X^2 = 800$. The point $(n, Z) = (200, 0.65)$ doesn't fit the formula: The point $(n, Z) = (1000, 0.90)$ also doesn't fit, since 1000 is greater than 800, Z should be 1.00.

$$Z = \sqrt{\frac{n}{\lambda_F C_X^2}} \quad \text{but} \quad 0.65 \neq \sqrt{\frac{200}{800}}$$

Answer C is not correct. The standard for full-credibility is $\lambda_F C_X^2 = (541)(1.053)^2$. The point $(n, Z) = (200, 0.65)$ doesn't fit the formula:

$$Z = \sqrt{\frac{n}{\lambda_F C_X^2}} \quad \text{but} \quad 0.65 \neq \sqrt{\frac{200}{(541)(1.053)^2}}$$

Answer D is not correct. Given EVPV= 0.266 and VHM= 0.016, we have

$$Z = \frac{n}{n + \text{EVPV}/\text{VHM}} \quad \text{but} \quad 0.65 \neq \frac{200}{200 + 16.625}.$$

Answer E is correct. Given EVPV= 0.844 and VHM= 0.008, we have

$$Z = \frac{200}{200 + 0.844/0.008} = 0.65,$$
$$Z = \frac{1000}{1000 + 0.844/0.008} = 0.90. \qquad \text{(Answer: E)}$$

Question 3.28. The risk group has the probabilities $\Pr(G = R) = 0.4$ and $\Pr(G = S) = 0.6$. Denote the aggregate loss as $L = X_1 + X_2 + \cdots + X_N$. The conditional means and variances are

$$\begin{aligned}
\text{E}(L|R) &= \text{E}(N|R)\text{E}(X|R) = (3)(4,000) = 12,000,\\
\text{Var}(L|R) &= \text{E}(N|R)\text{E}(X^2|R) = (3)(98,000,000/6) = 49,000,000,\\
\text{E}(L|S) &= \text{E}(N|S)\text{E}(X|S) = (1)(275) = 275,\\
\text{Var}(L|S) &= \text{E}(N|S)\text{E}(X^2|S) = (1)(92,500) = 92,500.
\end{aligned}$$

Hence,

$$\begin{aligned}
\text{EPV} &= \text{E}[\text{Var}(L|G)] = (49,000,000)(0.4) + (92,500)(0.6) = 19,655,500,\\
\text{VHM} &= \text{Var}[\text{E}(L|G)] = (12,000 - 275)^2(0.4)(0.6) = 32,994,150, \qquad \text{(Bernoulli variance formula)}\\
Z &= \frac{2}{2 + \text{EPV}/\text{VHM}} = 0.77.
\end{aligned}$$

Part E

Practice Sets

Practice Set I Questions

The **Systolic Blood Pressure Case Study** used for Questions 1–3 can be downloaded from the following link:
https://www.casact.org/sites/default/files/2021-04/MASII_Sample_Case_Study.pdf

Question 1. In the supplemental material, you have been given a case study, "Systolic Blood Pressure Case Study", showing the results of different treatment options and the description of how that study was set up. There are different ways of setting up models to examine the benefits of the different treatment options. You have been asked which of the two model structures will give a better fit to the experience.

Model A has:

(i) All eight treatment options in the fixed effects section of the model

(ii) A random effect of doctors nested within hospitals

(iii) An assumption of constant variance across treatment effects

Model B has:

(i) All eight treatment options in the fixed effects section of the model

(ii) A random effect of doctors nested within hospitals

(iii) An assumption that the variance by treatment can be grouped under Variance Group #3

The null hypothesis is that variance is constant across all treatment effects. Determine the p-value, the level of significance at which one would reject the null hypothesis, using a likelihood ratio test.

A. Less than 0.005

B. At least 0.005, but less than 0.01

C. At least 0.01, but less than 0.025

D. At least 0.025, but less than 0.05

E. At least 0.05

Question 2. In the supplemental material, you have been given a case study, "Systolic Blood Pressure Case Study", showing the results of different treatment options and the description of how that study was set up. Determine the number of the parameters in Model 1.

A. 11 B. 12 C. 13 D. 14 E. 15

Question 3. In the supplemental material, you have been given a case study, "Systolic Blood Pressure Case Study", showing the results of different treatment options and the description of how that study was set up.

Determine the marginal predicted "change_systolic", $\hat{Y}$ for a patient in the second treatment group using Model 2.

A. Less than -1
B. At least -1, but less than 0
C. At least 0, but less than 1
D. At least 1, but less than 2
E. At least 2

Question 4. An insurance company sells automobile policies, each of which belongs to one of two possible risk groups, R_1 and R_2. You are given the following information:

(i) Risk group R_1 occurs 30% of the time.
(ii) Risk group R_1 has claim frequencies that are Poisson distributed with parameter $\lambda = 3$.
(iii) Risk group R_1 has claim severity that is exponential distributed with parameter $\theta = 3000$.
(iv) Risk group R_2 occurs 70% of the time.
(v) Risk group R_2 has claim frequencies that are Poisson distributed with parameter $\lambda = 4$.
(vi) Risk group R_2 has claim severity that is exponential distributed with parameter $\theta = 2000$.
(vii) Claim frequency and claim severity are independently distributed given a risk group.

The Bühlmann credibility method is used to calculate the next year's predicted aggregate loss given 4 prior years of loss experience for a given risk. Calculate the Bühlmann credibility factor for this risk.

A. Less than 0.05
B. At least 0.05, but less than 0.15
C. At least 0.15, but less than 0.25
D. At least 0.25, but less than 0.35
E. At least 0.35

Question 5. An insurance company writes homeowners' policies in various regions across the country. You are given the following information:

(i) The company wrote 2,468,204 policies countrywide with a pure premium of 900.

(ii) In Region A, this company wrote 72,000 policies with a pure premium of 500.

(iii) The expected variance of the pure premium within each region is 1,400,000,000.

(iv) The variance of the region pure premium means is 50,000.

(v) Within each region, the losses for each automobile policy are identically distributed.

Calculate the credibility-weighted pure premium for Region A using Bühlmann credibility.

A. Less than 500

B. At least 500, but less than 550

C. At least 550, but less than 600

D. At least 600, but less than 650

E. At least 650

Question 6. You are given:

(i) Claim frequency each month has mean λ and variance λ.

(ii) λ follows a gamma distribution with $\alpha = 10$ and $\theta = 0.1$

(iii) The following table of claim experience for a company:

Month	Number of Insureds	Number of Claims
1	5	6
2	10	12
3	15	15
4	20	?

Calculate the estimated claim count for this company in Month 4 using the Bühlmann-Straub credibility approach.

A. Less than 22

B. At least 22, but less than 24

C. At least 24, but less than 26

D. At least 26, but less than 28

E. At least 28

Question 7. You are given the following information:

(i) The claims had a mean loss of 40 and variance of loss of 6,400.

(ii) The mean frequency of these claims is 0.08 per policy.

(iii) The block has 18,000 policies.

(iv) Full credibility is based on a coverage probability of 90% for a range of within 6% deviation from the true mean.

Calculate the partial-credibility factor for pure premium, Z_p , using the limited-fluctuation credibility method.

A. Less than 0.5

B. At least 0.5, but less than 0.6

C. At least 0.6, but less than 0.7

D. At least 0.7, but less than 0.8

E. At least 0.8

Question 8. You are given:

(i) X is the claim severity random variable which can take values 100, 200, or 500.

(ii) The distribution of X differs by the risk group, θ.

(iii) The following data table:

		$\Pr(X = x\|\theta)$		
θ	$\Pr(\Theta = \theta)$	$x = 100$	$x = 200$	$x = 500$
1	0.4	0.3	0.3	0.4
2	0.6	0.2	0.2	0.6

A sample of three claims with claim severities of 200, 200, and 500 is observed. Calculate the Bayesian estimate of X given the observed severities.

A. Less than 230

B. At least 230, but less than 310

C. At least 310, but less than 390

D. At least 390, but less than 470

E. At least 470

Question 9. Suppose an insurer pursues two classes of business in the auto insurance market: class A and class B. Given the following information:

(i) There is a 20% chance of writing a policy from class A and 80% chance of writing a policy from class B.

(ii) Claim counts arising from a policy within a class follow a Geometric distribution with parameter $\beta = 0.02$ if in class A, and $\beta = 0.01$ if in class B.

(iii) The insurer writes a policy but does not know to which class the policyholder belongs.

(iv) The insurer experiences one loss from this policy in the first year.

(v) The policy renews for a second year.

Calculate the probability of the insurer experiencing one loss from this policy in the second year. (fill in the blank)

[ ] (round to the nearest three decimal points)

Question 10. An individual birth weight observation Y_{ij} on rat pup i within the j-th litter is

$$Y_{ij} = \beta_0 + \beta_1 X_j^{(1)} + \beta_2 X_j^{(2)} + \beta_3 X_{ij}^{(3)} + \beta_4 X_j^{(4)} + \beta_5 X_{ij}^{(5)} + \beta_6 X_{ij}^{(6)} + u_j + \epsilon_{ij}$$

where Y_{ij} is the dependent variable (rat pup weight), $X_j^{(1)}$ is indicator variable for the high-dose treatment, $X_j^{(2)}$ is indicator variable for the low-dose treatment, $X_{ij}^{(3)}$ is the indicator for female rat pups, $X_{ij}^{(4)}$ is the litter size, and $X_{ij}^{(5)} = X_j^{(1)} X_{ij}^{(3)}$ and $X_{ij}^{(6)} = X_j^{(2)} X_{ij}^{(3)}$ are interaction terms.

You are given the following information:

```
Fixed effects: weight ~ treatment + sex1 + litsize + treatment:sex1

                       Value  Std.Error  DF   t-value  p-value
(Intercept)         8.317282 0.24805726 283  33.52969   0.0000
treatmentHigh      -0.879354 0.17455026  23  -5.03783   0.0000
treatmentLow       -0.443022 0.14504899  23  -3.05429   0.0056
sex1               -0.289334 0.05906798 283  -4.89833   0.0000
litsize            -0.129684 0.01700908  23  -7.62437   0.0000
treatmentHigh:sex1 -0.023173 0.10455808 283  -0.22163   0.8248
treatmentLow:sex1  -0.041230 0.08414936 283  -0.48996   0.6245
```

Construct the 90% symmetric confidence interval for $\hat{\beta}_5$ assuming normal distribution.

A. (-0.175, 0.149)

B. (-0.185, 0.149)

C. (-0.195, 0.149)

D. (-0.195, 0.139)

E. (-0.195, 0.129)

Question 11. A hierarchical model is given below:

$$Y_{tij} = b_{0i|j} + b_{1j}\,X_t^{(1)} + \epsilon_{tij} \qquad \text{(Level 1 model)}$$
$$b_{0i|j} = b_{0j} + \beta_2\,X_{ij}^{(2)} + \beta_3\,X_{ij}^{(3)} + u_{0i|j} \qquad \text{(Level 2 model)}$$
$$b_{0j} = \beta_0 + \beta_4\,X_j^{(4)} + u_{0j} \qquad \text{(Level 3 models)}$$
$$b_{1j} = \beta_1 + u_{1j}$$

You are given the following models:

$$Y_{tij} = \beta_0 + \beta_1\,X_t^{(1)} + \beta_2\,X_{ij}^{(2)} + \beta_3\,X_{ij}^{(3)} + \beta_4\,X_j^{(4)} + u_{0i|j} + \epsilon_{tij} \qquad \text{(Model W)}$$
$$Y_{tij} = \beta_0 + \beta_1\,X_t^{(1)} + \beta_2\,X_{ij}^{(2)} + \beta_3\,X_{ij}^{(3)} + \beta_4\,X_j^{(4)} + u_{1j}\,X_t^{(1)} + u_{0i|j} + \epsilon_{tij} \qquad \text{(Model X)}$$
$$Y_{tij} = \beta_0 + \beta_1\,X_t^{(1)} + \beta_2\,X_{ij}^{(2)} + \beta_3\,X_{ij}^{(3)} + \beta_4\,X_j^{(4)} + u_{0j} + u_{0i|j} + \epsilon_{tij} \qquad \text{(Model Y)}$$
$$Y_{tij} = \beta_0 + \beta_1\,X_t^{(1)} + \beta_2\,X_{ij}^{(2)} + \beta_3\,X_{ij}^{(3)} + \beta_4\,X_j^{(4)} + u_{0j} + u_{1j}\,X_t^{(1)} + u_{0i|j} + \epsilon_{tij} \qquad \text{(Model Z)}$$

Determine which model is equivalent to the hierarchical model.

A. W B. X C. Y D. Z E. None of above

Use the following information for questions 12 - 13.

The daily miles traveled for 20 working adults who live on the same street are measured for a month. The adults were classified into four groups based on their commute to work (NOCOMMUTE, SHORT, MEDIUM, or LONG).

The table below illustrates the data:

Person ID	Length of Commute	Day	Miles Traveled
1	NOCOMMUTE	1	1.2
1	NOCOMMUTE	2	0.5
1	NOCOMMUTE	3	3.7
...	...	...	...
1	NOCOMMUTE	30	2.6
2	SHORT	1	3.4
...	...	...	...
13	MEDIUM	1	10.9
...	...	...	...
20	LONG	30	25.6

Question 12. The hierarchical Linear Mixed Model given below is used to model MILES_{ti}, where MILES_{ti} is the t^{th} daily mileage of the i^{th} adult.

$$\mathsf{MILES}_{ti} = b_{0i} + b_{1i}\mathsf{DAY}_{ti} + \epsilon_{ti} \qquad \text{(Level 1 Model: Time)}$$

$$\text{where } \epsilon_{ti} \sim N(0, \sigma_\epsilon^2)$$

$$b_{0i} = \beta_0 + \beta_2\mathsf{SHORT}_i + \beta_3\mathsf{MEDIUM}_i + \beta_4\mathsf{LONG}_i + \mu_{0i}$$

$$b_{1i} = \beta_1 + \mu_{1i} \qquad \text{(Level 2 Model: Individual adult)}$$

$$\text{where } \boldsymbol{\mu}_i \sim N(0, \boldsymbol{D})$$

$\boldsymbol{\mu}_i = (\mu_{0i}, \mu_{1i})'$. Determine the Linear Mixed Model that is equivalent to the above hierarchical model.

A. $\mathsf{MILES}_{ti} = \beta_0 + \beta_1 + \beta_2\mathsf{SHORT}_i\mathsf{DAY}_{ti} + \beta_3\mathsf{MEDIUM}_i\mathsf{DAY}_{ti} + \beta_4\mathsf{LONG}_i\mathsf{DAY}_{ti} + \mu_{0i} + \mu_{1i} + \epsilon_{ti}$

B. $\mathsf{MILES}_{ti} = \beta_0 + \beta_1\mathsf{DAY}_{ti} + \beta_2\mathsf{SHORT}_i\mathsf{DAY}_{ti} + \beta_3\mathsf{MEDIUM}_i\mathsf{DAY}_{ti} + \beta_4\mathsf{LONG}_i\mathsf{DAY}_{ti} + \mu_{0i} + \mu_{1i}\mathsf{DAY}_{ti} + \epsilon_{ti}$

C. $\mathsf{MILES}_{ti} = \beta_0 + \beta_1 + \beta_2\mathsf{SHORT}_i + \beta_3\mathsf{MEDIUM}_i + \beta_4\mathsf{LONG}_i + \mu_{0i} + \mu_{1i} + \epsilon_{ti}$

D. $\mathsf{MILES}_{ti} = \beta_0 + \beta_1\mathsf{DAY}_{ti} + \beta_2\mathsf{SHORT}_i + \beta_3\mathsf{MEDIUM}_i + \beta_4\mathsf{LONG}_i + \mu_{0i} + \mu_{1i}\mathsf{DAY}_{ti} + \epsilon_{ti}$

E. $\mathsf{MILES}_{ti} = \beta_0 + \beta_1\mathsf{DAY}_{ti} + \beta_2\mathsf{SHORT}_i + \beta_3\mathsf{MEDIUM}_i + \beta_4\mathsf{LONG}_i + (\mu_{0i} + \mu_{1i})\mathsf{DAY}_{ti} + \epsilon_{ti}$

Question 13. The hierarchical Linear Mixed Model given below is used to model MILES_{ti}, where MILES_{ti} is the t^{th} daily mileage of the i^{th} adult.

$$\mathsf{MILES}_{ti} = b_{0i} + b_{1i}\mathsf{DAY}_{ti} + \epsilon_{ti} \qquad \text{(Level 1 Model: Time)}$$

$$\text{where } \epsilon_{ti} \sim N(0, \sigma_\epsilon^2)$$

$$b_{0i} = \beta_0 + \beta_2\mathsf{SHORT}_{1i} + \beta_3\mathsf{MEDIUM}_{1i} + \beta_4\mathsf{LONG}_{1i} + \mu_{0i}$$

$$b_{1i} = \beta_1 + \mu_{1i} \qquad \text{(Level 2 Model: ndividual adult)}$$

$$\text{where } \boldsymbol{\mu}_i \sim N(0, \boldsymbol{D})$$

$\boldsymbol{\mu}_i = (\mu_{0i}, \mu_{1i})'$. What's wrong with the above hierarchical model?

A. There are two random factors (μ_{0i} and μ_{1i}) in the model. The model is not estimable.

B. There are two intercepts (β_0 and β_5) in the model. The model is not estimable.

C. Only 3 categories for "Length of Commute" are included in the model. The model is estimable but not complete.

D. The subscript on the "Length of Commute" variables in the Level 2 Model should be i instead of $1i$.

E. There is nothing wrong with the model.

Question 14. You are given the following model:

$$Y_{ijk} = \beta_0 + \beta_1 X_{ijk}^{(1)} + \beta_2 X_{ijk}^{(2)} + \beta_3 X_{ijk}^{(3)} + \beta_4 X_{ijk}^{(4)} + u_k + u_{j|k} + \epsilon_{ijk}$$

where Y_{ijk} represents the value of the dependent variable mathgain for student i in classroom j nested within school k, $X^{(1)}$ to $X^{(4)}$ are student level (Level 1) covariates representing mathkind, sex, minority, and ses, respectively.

We assume that the random effects, u_k, associated with schools, the random effects, $u_{j|k}$, associated with classrooms nested within schools, and the residuals, ϵ_{ijk}, are all mutually independent. The following information is given:

```
> summary(model.ml.fit)
Random effects:
 Formula: ~1 | schoolid
        (Intercept)
StdDev:        7.56

 Formula: ~1 | classid %in% schoolid
        (Intercept) Residual
StdDev:        8.60    23.79

Fixed effects: mathgain ~ mathkind + sex + minority + ses
                 Value Std.Error  DF    t-value p-value
(Intercept)     277.01 10.280665 770  26.945216  0.0000
mathkind         -0.46  0.021091 770 -21.813731  0.0000
sex              -1.76  1.555295 770  -1.130839  0.2585
minority         -6.47  2.144891 770  -3.016381  0.0026
ses               4.95  1.153175 770   4.293239  0.0000
```

Consider a boy (sex=0) student whose mathkind is 500 and ses is 1. Determine the probability that the minority (minority=1) boy's mathgain is less than 80 using the marginal distribution.

A. 0.5 B. 0.6 C. 0.7 D. 0.8 E. 0.9

Question 15. Consider the following model:

$$Y_{ti} = \beta_0 + \beta_1 X_{ti}^{(1)} + \beta_2 X_{ti}^{(2)} + \beta_3 X_{ti}^{(3)} + \beta_4 X_{ti}^{(4)} + \beta_5 X_{ti}^{(5)} + u_{0i} + u_{3i} X_{ti}^{(3)} + \epsilon_{ti}$$

where $Y_{ti} = \text{ACTIVATE}_{ti}$ is the activate value of the tth observation on the ith animal, $X_{ti}^{(1)} = \text{REGION1}$ and $X_{ti}^{(2)} = \text{REGION2}$ are indicators which represent the BST and LS regions, respectively. When both indicators are zero, the region is VBD, $X_{ti}^{(3)} = \text{TREATMENT}$ is an indicator variable that indicates the Carbachol treatment if 1, the Basal treatment if 0. Two interaction terms $X_{ti}^{(4)} = X_{ti}^{(1)} X_{ti}^{(3)}$ and $X_{ti}^{(5)} = X_{ti}^{(2)} X_{ti}^{(3)}$ are also included in the model.

The residuals are iid and follow a normal distribution: $\epsilon_{ti} \sim N(0, \sigma^2)$. The u_{0i} term represents the random intercept associated with animal i and the term u_{3i} represents the random effect associated with treatment for animal i.

We assume that the distribution of the random effects u_{0i} and u_{3i} is bivariate normal:

$$u_i = \begin{bmatrix} u_{0i} \\ u_{3i} \end{bmatrix} \sim N(0, D), \qquad D = \begin{bmatrix} \sigma_{in}^2 & \rho\,\sigma_{in}\,\sigma_{tr} \\ \rho\,\sigma_{in}\,\sigma_{tr} & \sigma_{tr}^2 \end{bmatrix}$$

where σ_{in}^2 is the variance of the random intercepts, σ_{tr}^2 is the variance of the random treatment effect, and ρ is the correlation of the two random effects. We assume that the ϵ_i's are independent of u_{0i}'s and u_{3i}'s. Which of the followings represents the variance of Y_{ti} if the treatment is the Carbachol?

A. σ^2

B. $\sigma^2 + \sigma_{in}^2$

C. $\sigma^2 + \sigma_{tr}^2$

D. $\sigma^2 + \sigma_{in}^2 + \sigma_{tr}^2$

E. $\sigma^2 + \sigma_{in}^2 + \sigma_{tr}^2 + 2\rho\sigma_{in}\sigma_{tr}$

Question 16. You are given the following model:

$$Y_{ti} = \beta_0 + \beta_1 X_{ti}^{(1)} + \beta_2 X_{ti}^{(2)} + \beta_3 X_{ti}^{(3)} + \beta_4 X_{ti}^{(4)} + \beta_5 X_{ti}^{(5)} + u_{0i} + u_{3i} X_{ti}^{(3)} + \epsilon_{ti}$$

where $u_{0i} \sim N(0, \sigma_{in}^2)$, $u_{3i} \sim N(0, \sigma_{tr}^2)$, $\text{cor}(u_{0i}, u_{3i}) = \rho$, $\epsilon_{ti} \sim N(0, \sigma^2)$, and ϵ_{ti} are independent if u_{0i} and u_{3i}. The term u_{3i} represents the random effect associated with treatment for animal i.

From the following partial summary output, the estimated standard errors of the random effects are shown below.

```
Fixed effects: activate ~ region.f * treat

                      Value Std.Error DF    t-value p-value
(Intercept)         212.294  19.09551 20  11.117483  0.0000
region.f1           216.212  14.68203 20  14.726304  0.0000
region.f2            25.450  14.68203 20   1.733412  0.0984
treat1              360.026  38.59808 20   9.327561  0.0000
region.f1:treat1   -261.822  20.76352 20 -12.609710  0.0000
region.f2:treat1   -162.500  20.76352 20  -7.826225  0.0000

Random effects:
 Formula: ~treat | animal
 Structure: General positive-definite, Log-Cholesky parametrization
             StdDev   Corr
(Intercept)    35.8 (Intr)
treat1         79.8    0.8
Residual       23.2
```

Match each random effect on the left with its variance on the right.

Random effect
1. u_{0i}
2. u_{3i}

Variance
A. 1282
B. 6368
C. 538

Question 17. Decision trees for regression and classification have a number of advantages over the more classical approaches:

I. Trees are very easy to explain to people. In fact, they are even easier to explain than linear regression!

II. Some people believe that decision trees more closely mirror human decision-making than do the regression and classification approaches seen in previous chapters.

III. Trees can be displayed graphically, and are easily interpreted even by a non-expert.

IV. Trees can easily handle qualitative predictors without the need to create dummy variables.

Determine which of the preceding statements are true. (select all that apply)

Question 18. Decision trees for regression and classification have some disadvantages to the corresponding classical approaches:

I. Decision trees generally do not have the same level of predictive accuracy as some of the other regression and classification approaches.

II. Decision trees can be very non-robust. In other words, a small change in the data can cause a large change in the final estimated tree.

III. Trees cannot be displayed graphically.

Determine which of the preceding statements are true. (select all that apply)

Question 19. You are provided the following data set with a single variable X.

i	X
1	9
2	15
3	4
4	2
5	18

A dendrogram is built from this data set using agglomerative hierarchical clustering with single linkage and Euclidean distance as the dissimilarity measure. Calculate the tree height at which observation $i = 5$ fuses.

A. Less than 1.5

B. At least 1.5, but less than 3.5

C. At least 3.5, but less than 5.5

D. At least 5.5, but less than 7.5

E. At least 7.5

Question 20. You are given the following data to train a K-nearest Neighbors classifier with $K = 3$:

i	X_{i1}	X_{i2}	Y
1	0	1	Low
2	0	2	Low
3	1	1	Low
4	1	2	High
5	3	3	High
6	3	4	Low
7	3	5	Low
8	4	4	Low

Assign the first 4 observations to the training set and the last 4 observations to the test set. Using the K-nearest Neighbors classifier with $K = 3$, calculate

$$\Pr(Y = \text{“Low”} | X_{81} = 4, X_{82} = 4)$$

using the **training set**.

A. Less than 0.1

B. At least 0.1, but less than 0.3

C. At least 0.3, but less than 0.5

D. At least 0.5, but less than 0.7

E. At least 0.7

Question 21. You are given the following data to train a K-nearest Neighbors classifier with $K = 3$:

i	$X_i = (X_{i1}, X_{i2})$	Y	E.d(X_i, X_5)	E.d(X_i, X_6)	E.d(X_i, X_7)
1	(0,1)	Low	1	1	$\sqrt{5}$
2	(0,2)	High	2	$\sqrt{2}$	2
3	(1,1)	Low	$\sqrt{2}$	0	$\sqrt{2}$
4	(1,2)	High	1	1	1
5	(0,0)	Low			
6	(1,1)	Low			
7	(2,2)	High			

where E.d.$(X_i, X_j) = \sqrt{(X_{i1} - X_{j1})^2 + (X_{i2} - X_{j2})^2}$. Assign the first 4 observations to the training set and the last 3 observations to the test set. Determine the test error rate of the of the K-nearest Neighbors classifier with $K = 3$.

A. Less than 33%

B. At least 33%, but less than 55%

C. At least 55%, but less than 77%

D. At least 77%, but less than 99%

E. At least 99%

Question 22. Cluster A consists of 3 points: A=$\{1, 3, 5\}$. Cluster B consists of 3 points: B=$\{2, 4, 6\}$. Calculate the Euclidean distance between the two clusters using the average linkage.

A. Less than 1

B. At least 1, but less than 2

C. At least 2, but less than 3

D. At least 3, but less than 4

E. At least 4

Question 23. Cluster A consists of 3 points: A=$\{21, 13, 35\}$. Cluster B consists of 3 points: B=$\{32, 24, 16\}$.

Calculate the Euclidean distance between the two clusters using the centroid linkage.

A. Less than 1

B. At least 1, but less than 2

C. At least 2, but less than 3

D. At least 3, but less than 4

E. At least 4

Question 24. What is the purpose of the activation function in a single layer neural network?

I. To compute the weighted sum of input features.

II. To transform and introduce nonlinearity to the hidden layer.

III. To determine the number of hidden units.

IV. To estimate the parameters using squared-error loss.

Determine which of the preceding statements are true. (select all that apply)

Question 25. In a multilayer neural network architecture designed for digit classification using the MNIST dataset, what is a key advantage of using multiple hidden layers with modest-sized units?

A. It reduces the overall complexity of the model.
B. It allows for the direct application of squared-error loss for training.
C. It significantly speeds up the learning process.
D. It helps the network build up complex transformations of the input data.
E. It leads to a decrease in the number of output variables.

Question 26. Match each object regarding convolutional neural network on the left with the corresponding description on the right.

Object	Definition
1. Convolutional Neural Networks	A. A technique that expands the training dataset by replicating images with random distortions to prevent overfitting.
2. Convolution Layer	B. Specialized layers in a convolutional neural network used to detect local features in an image.
3. Pooling Layer	C. Layers in a CNN responsible for condensing a large image into a smaller summary image, often using max pooling.
4. Data Augmentation	D. A special family of networks designed for image classification tasks, which operate in a hierarchical manner, starting with low-level features.

Question 27. What are Recurrent Neural Networks (RNNs) primarily suited for?

A. Image classification

B. Natural language processing

C. Predicting stock prices

D. Solving Sudoku puzzles

E. Digit recognition

Question 28. Determine which of the following statements are true. (select all that apply)

I. Recurrent Neural Networks (RNNs) are well-suited for sequential data, such as time series or text, because they can leverage the sequential nature of the input data.

II. In deep learning, regularization techniques like dropout are used to prevent overfitting and improve the generalization of neural network models.

III. The bias-variance tradeoff suggests that interpolating the training data (achieving zero training error) will always result in better test performance compared to slightly less complex models that don't interpolate the data.

Question 29. Match each object regarding recurrent neural network on the left with the corresponding description on the right.

Object	Definition
1. Backpropagation	A. A regularization technique that randomly disables a fraction of units within a neural network layer during training to prevent overfitting.
2. Dropout	B. An optimization algorithm used to train neural networks by iteratively updating the network's parameters in the direction of steepest decrease in the loss function.
3. Stochastic Gradient Descent	C. A method for calculating gradients in neural networks by applying the chain rule of differentiation, allowing for efficient model training and weight updates.

Question 30. You are given the test scores of 8 students:

Student	Test score
1	1.3
2	2.2
3	3.1
4	5.8
5	6.8
6	6.8
7	8.8
8	9.6

You are to use a hierarchical clustering algorithm with centroid linkage and Euclidean distance to group students whose test scores are similar.

You will stop the algorithm once you have obtained 4 clusters. Those with the lowest test scores will be placed into the first cluster. Calculate the average test score of this cluster. (fill in the blank)

[ ] (round to the nearest one decimal point)

Question 31. You are given the output of the R functions, rpart() and printcp():

```
> printcp(rtree)

Regression tree:
rpart(formula = survived ~ . , data = data)

Variables actually used in tree construction:
[1] fare sex family

Root node error: 0.2857

         CP  nsplit  rel error   xerror      xstd
1  0.279516       0    1.00000  -------  0.013462
2  0.026767       1    0.72048  -------  0.029045
3  0.014423       2    0.69372  w        0.028255
4  0.013534       3    0.67929  -------  0.027093
5  0.010938       4    0.66576  -------  0.026900
6  0.010000       5    0.65482  -------  0.024873
```

When the tree has a size of 3, the cross-validated error rate is 0.2. Find the value of w. (fill in the blank)

[ ] (round to the nearest one decimal point)

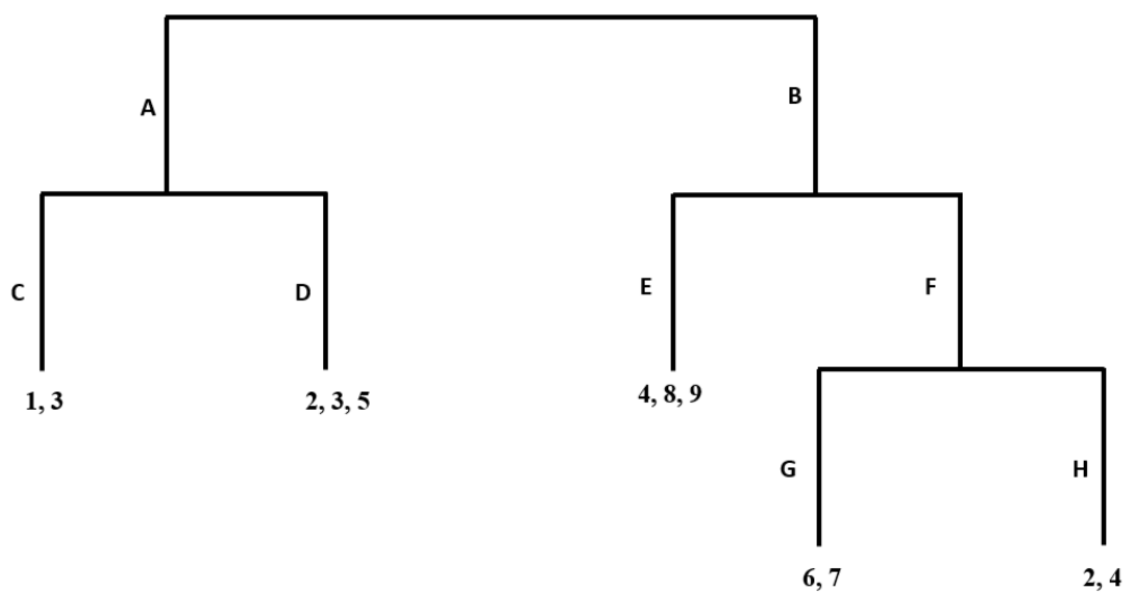
Question 32. Which of the following statements are correct about the eigenvalues in Principal Component Analysis? (select all that apply)

I. The size of the eigenvalues is used to determine the number of principal components.

II. The principal components with eigenvalues that are greater than 1 are retained.

III. The acceptable eigenvalues are between 0.5 and 1 to retain the principal components.

Question 33. You are given the following regression tree with response values at each terminal node: You consider pruning the tree using the cost complexity pruning method with tuning parameter $\alpha = 3.5$.



Which of the following pruning strategy is optimal?

A. Do not prune the tree.

B. Prune the tree at node A only.

C. Prune the tree at node B only.

D. Prune the tree at node F only.

E. Prune the tree at nodes A and F only.

Question 34. You are given the following five ordered sample values from a time series $\{x_t\}$.

t	x_t
1	5.3
2	4.8
3	5.2
4	4.5
5	5.2

Calculate the sample lag 2 autocorrelation.

A. Less than 0.4

B. At least 0.4, but less than 0.5

C. At least 0.5, but less than 0.6

D. At least 0.6, but less than 0.7

E. At least 0.7

Question 35. Differencing and the log-transformation are commonly used operations in time series analysis.

Match an operation on the left with each use on the right.

Operation
1. Differencing
2. Log-transformation

Use
A. To stabilize the variance of a time series
B. To remove stochastic trends
C. To remove seasonal effects

Question 36. Consider a sample of size 3 from a stationary time series $\{x_t\}$ with variance $\sigma^2 = 2$ and autocorrelation function $\rho_k = 0.7^k$ for $k = 0, 1, 2, \ldots$.

Calculate the variance of the sample mean $\bar{x}$.

A. Less than 1.2

B. At least 1.2, but less than 1.4

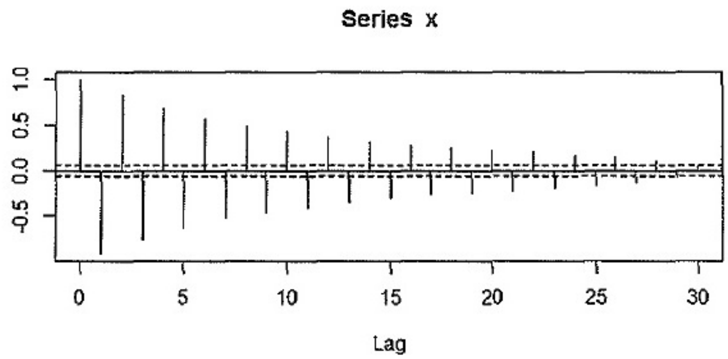
C. At least 1.4, but less than 1.6

D. At least 1.6, but less than 1.8

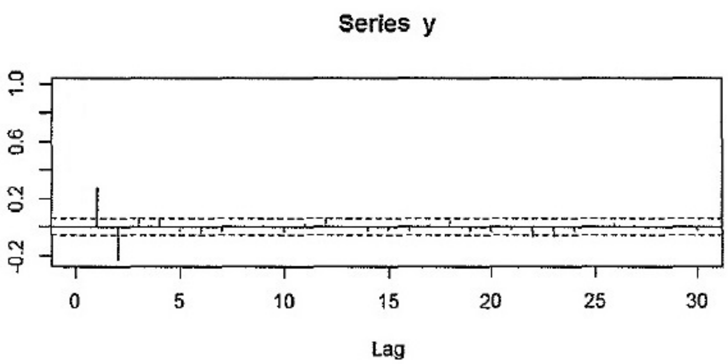
E. At least 1.8

Question 37. You are given the following information:

- x and y are two stationary time series.
- ACF of x:



- PACF of y:



- The dashed lines above and below zero indicate the range within which the ACF or PACF results are considered not significantly different than zero.
- The symbol w_t is white noise with zero mean.

Determine which of the following statements best displays the model structure that describes series x and series y.

A. $x_t = 0.9x_{t-1} + w_t$ and $y_t = 0.9y_{t-1} + w_t + 0.6w_{t-1} - 0.3w_{t-2}$

B. $x_t = -0.9x_{t-1} + w_t$ and $y_t = w_t + 0.6w_{t-1} - 0.3w_{t-2}$

C. $x_t = 0.9x_{t-1} + w_t$ and $y_t = w_t + 0.6w_{t-1}$

D. $x_t = -0.9x_{t-1} + w_t$ and $y_t = 0.6y_{t-1} - 0.3y_{t-2} + w_t$

E. $x_t = 0.9x_{t-1} + w_t$ and $y_t = 0.6y_{t-1} - 0.3y_{t-2} + w_t$

Question 38. You are given the following autoregressive model of order one:

$$\begin{aligned} x_t &= 1.2 - 0.8x_{t-1} + w_t \\ x_{10} &= 1.0 \end{aligned}$$

where $\{w_t\}$ is a white noise process.

Calculate $\hat{x}_{30}$, the forecast for x_{30}.

A. Less than 0.70
B. At least 0.70, but less than 0.75
C. At least 0.75, but less than 0.80
D. At least 0.80, but less than 0.85
E. At least 0.85

Question 39. You are given the following AR(2) process:

$$x_t = 33 + 0.5x_{t-2} + w_t,$$

where $\{w_t\}$ is a white noise process.

After observing $x_{76} = 64$ and $x_{77} = 68$, your forecast for x_{80} is a.

After observing $x_{76} = 64$, $x_{77} = 68$ and $x_{78} = 66$, your forecast for x_{80} becomes b.

Calculate $b - a$.

A. Less than 0.45
B. At least 0.45, but less than 0.55
C. At least 0.55, but less than 0.65
D. At least 0.65, but less than 0.75
E. At least 0.75

Question 40. You fit an MA(1) model $x_t = w_t + \beta w_{t-1}$ to the following data:

$$x_1 = 0, \quad x_2 = 2, \quad x_3 = 3.$$

Calculate the conditional sum of squared residuals for $\hat{\beta} = 1.5$. (round to the nearest whole number)

Question 41. Identify the following time series model as a specific ARIMA model:

$$x_t = 0.5x_{t-1} + 0.5x_{t-2} + w_t - 0.5w_{t-1} + 0.25w,$$

where $\{w_t\}$ is a white noise process.

A. ARIMA$(0, 1, 1)$
B. ARIMA$(1, 1, 1)$
C. ARIMA$(1, 1, 2)$
D. ARIMA$(2, 1, 1)$
E. ARIMA$(1, 2, 1)$

Question 42. Determine which of the following time series processes is/are stationary. (select all that apply)

A. Model I: $x_t = 1 + x_{t-1} - 0.25x_{t-2} + w_t$
B. Model II: $x_t = 0.5 + w_t - 0.8w_{t-1} + 0.5w_{t-2}$
C. Model III: $x_t = 0.4x_{t-5} + 0.6x_{t-10} + w_t - 0.56w_{t-1}$

Practice Set I Solutions

Question 1. Model A is Model 1 (see page 9 of the Case Study):

(i) We are given that Model A includes all 8 treatment options in the fixed section. Page 9 shows "Treatment group definition: Full".

(ii) We are given an assumption of constant variance across treatment effects. Page 9 shows "Variance grouping: None".

(iii) The REML is preferred to ML estimation when estimating covariance parameters because it produces unbiased estimates of covariance estimates (see Sec. 10 of the Study Manual). Page 9 shows "Computation method: restricted maximum likelihood".

Model B is Model 7 (see page 21 of the Case Study):

(i) Same as (i) above. Page 21 shows "Treatment group definition: Full".

(ii) We are given that the variance by treatment can be grouped under Variance Group #3. Page 21 shows "Variance grouping: Variance Group #3".

(iii) Same as (iii) above. Page 21 shows "Computation method: restricted maximum likelihood".

There are 4 groups under Variance Group #3 (see page 5) under the column "Var_F_G_3"). Therefore, the null and alternative hypotheses are:

$$H_0 : \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2 = \sigma^2 \qquad \text{(Model A)}$$

$$H_A : \text{At least one pair of variances is not equal to each other} \qquad \text{(Model B)}$$

where σ_i^2 is the variance for group i, $i = 1, 2, 3, 4$. The model under the null hypothesis is the nested model which has homogeneous variance. The model under the alternative is the reference model which has heterogeneous variance. Using a likelihood ratio test, the test statistic is:

$$T = 2 \times \{\mathsf{logLik}(\text{reference}) - \mathsf{logLik}(\text{nested})\}$$

The asymptotic null distribution of the test statistic is χ^2 distributed with 3 degrees of freedom corresponding to the 3 additional residual variances in Model B compared to the null model. The log-likelihood value under Model 1 (nested) is -4639.327 (page 41) and the log-likelihood value under Model 7 (reference) is -4637.635 (page 45). Hence,

$$T = (2)[(-4637.635) - (-4639.327)] = 3.384.$$