

---

---

# Contents

---

|           |  |           |
|-----------|--|-----------|
| <b>I</b>  | <b>Probability Review</b>  | <b>1</b>  |
| <b>1</b>  | <b>Probability Review</b>  | <b>3</b>  |
| 1.1       | Functions and moments . . . . .  | 3         |
| 1.2       | Probability distributions . . . . .  | 5         |
| 1.2.1     | Bernoulli distribution . . . . .   | 5         |
| 1.2.2     | Uniform distribution . . . . .   | 5         |
| 1.2.3     | Exponential distribution . . . . .   | 6         |
| 1.3       | Variance . . . . .   | 6         |
| 1.4       | Normal approximation . . . . .   | 7         |
| 1.5       | Conditional probability and expectation . . . . .                              | 8         |
| 1.6       | Conditional variance . . . . .   | 10        |
|           | Exercises . . . . .  | 13        |
|           | Solutions . . . . .  | 16        |
| <b>II</b> | <b>Stochastic Processes</b>  | <b>21</b> |
| <b>2</b>  | <b>Markov Chains: Chapman-Kolmogorov Equations</b>                             | <b>23</b> |
| 2.1       | Chapman-Kolmogorov equations . . . . .   | 25        |
| 2.2       | Gambler's ruin . . . . .   | 27        |
| 2.3       | Algorithmic efficiency . . . . .   | 28        |
|           | Exercises . . . . .  | 30        |
|           | Solutions . . . . .  | 38        |
| <b>3</b>  | <b>Markov Chains: Classification of States</b>                                 | <b>45</b> |
| 3.1       | Random walks . . . . .   | 46        |
|           | Exercises . . . . .  | 47        |
|           | Solutions . . . . .  | 49        |
| <b>4</b>  | <b>Discrete Markov Chains: Long-Run Proportions and Limiting Probabilities</b> | <b>51</b> |
|           | Exercises . . . . .  | 53        |
|           | Solutions . . . . .  | 55        |
| <b>5</b>  | <b>Markov Chains: Time in Transient States</b>                                 | <b>61</b> |
|           | Appendix: Inverting a Matrix . . . . .   | 65        |
|           | Exercises . . . . .  | 67        |
|           | Solutions . . . . .  | 70        |
| <b>6</b>  | <b>Markov Chains: Branching Processes</b>                                      | <b>73</b> |
|           | Exercises . . . . .  | 75        |
|           | Solutions . . . . .  | 76        |
| <b>7</b>  | <b>Markov Chains: Time Reversible</b>  | <b>79</b> |
|           | Exercises . . . . .  | 81        |

|  |            |
|--|------------|
| Solutions . . . . .  | 82         |
| <b>8 Markov Chains: Monte Carlo Methods</b>  | <b>85</b>  |
| Exercises . . . . .  | 88         |
| Solutions . . . . .  | 90         |
| <b>9 Exponential Distribution</b>  | <b>93</b>  |
| 9.1 Lack of memory . . . . .   | 93         |
| 9.2 Minimum . . . . .  | 95         |
| 9.3 Sum . . . . .  | 96         |
| Exercises . . . . .  | 96         |
| Solutions . . . . .  | 97         |
| <b>10 The Poisson Process: Probabilities of Events</b>                                 | <b>101</b> |
| 10.1 Introduction . . . . .  | 101        |
| 10.2 Probabilities—homogeneous process . . . . .                                       | 102        |
| 10.3 Probabilities—non-homogeneous process . . . . .                                   | 104        |
| Exercises . . . . .  | 106        |
| Solutions . . . . .  | 111        |
| <b>11 The Poisson Process: Time To Next Event</b>                                      | <b>117</b> |
| Exercises . . . . .  | 120        |
| Solutions . . . . .  | 122        |
| <b>12 The Poisson Process: Thinning</b>  | <b>127</b> |
| 12.1 Constant probabilities . . . . .  | 127        |
| 12.2 Non-constant probabilities . . . . .  | 129        |
| Exercises . . . . .  | 131        |
| Solutions . . . . .  | 137        |
| <b>13 The Poisson Process: Other Characteristics</b>                                   | <b>143</b> |
| 13.1 Probability of events from one system before events from another system . . . . . | 143        |
| 13.2 Conditional Distribution of Event Time . . . . .                                  | 143        |
| Exercises . . . . .  | 145        |
| Solutions . . . . .  | 146        |
| <b>14 The Poisson Process: Sums and Mixtures</b>                                       | <b>149</b> |
| 14.1 Sums of Poisson processes . . . . .   | 149        |
| 14.2 Mixtures of Poisson processes . . . . .   | 150        |
| Exercises . . . . .  | 155        |
| Solutions . . . . .  | 158        |
| <b>15 Compound Poisson Processes</b>   | <b>163</b> |
| 15.1 Definition and moments . . . . .  | 163        |
| 15.2 Sums of compound distributions . . . . .  | 165        |
| Exercises . . . . .  | 166        |
| Solutions . . . . .  | 174        |
| <b>16 Continuous Markov Chains: Introduction</b>                                       | <b>181</b> |
| Exercises . . . . .  | 182        |
| Solutions . . . . .  | 184        |

|   |            |
|---|------------|
| <b>17 Markov Chains: Birth and Death Processes</b>                        | <b>189</b> |
| Exercises . . . . .   | 191        |
| Solutions . . . . .   | 192        |
| <b>18 Continuous Markov Chains: Limiting Probabilities</b>                | <b>195</b> |
| 18.1 Probability of state at time $t$ . . . . .                           | 195        |
| 18.2 Limiting probabilities . . . . .                                     | 197        |
| Exercises . . . . .   | 200        |
| Solutions . . . . .   | 205        |
| <b>19 Reliability: Structure Functions</b>                                | <b>209</b> |
| 19.1 Expressing a system as a parallel set of series structures . . . . . | 211        |
| 19.2 Expressing a system as a series set of parallel structures . . . . . | 211        |
| Exercises . . . . .   | 213        |
| Solutions . . . . .   | 214        |
| <b>20 Reliability: Probabilities</b>                                      | <b>217</b> |
| 20.1 Exact calculation of the reliability function . . . . .              | 217        |
| 20.2 Upper and lower bounds for the reliability function . . . . .        | 219        |
| 20.2.1 Inclusion/Exclusion Bounds . . . . .                               | 219        |
| 20.2.2 Estimating Intersection Probabilities . . . . .                    | 222        |
| Exercises . . . . .   | 225        |
| Solutions . . . . .   | 226        |
| <b>21 Reliability: Time to Failure</b>                                    | <b>229</b> |
| 21.1 Expected amount of time to failure . . . . .                         | 229        |
| 21.2 Increasing failure rates . . . . .                                   | 230        |
| Exercises . . . . .   | 231        |
| Solutions . . . . .   | 233        |
| <b>III Life Contingencies</b>   | <b>235</b> |
| <b>22 Survival Models</b>   | <b>237</b> |
| 22.1 Survival probabilities . . . . .                                     | 237        |
| 22.2 Special situations . . . . .   | 239        |
| 22.3 Life expectancy . . . . .  | 240        |
| Exercises . . . . .   | 242        |
| Solutions . . . . .   | 245        |
| <b>23 Contingent Payments</b>   | <b>249</b> |
| 23.1 Life insurances . . . . .  | 249        |
| 23.2 Life annuities . . . . .   | 251        |
| 23.3 Premiums . . . . .   | 253        |
| Exercises . . . . .   | 254        |
| Solutions . . . . .   | 265        |
| <b>IV Statistics</b>  | <b>273</b> |
| <b>24 Estimator Quality</b>   | <b>275</b> |

|   |            |
|---|------------|
| 24.1 Bias . . . . .   | 275        |
| 24.2 Consistency . . . . .  | 277        |
| 24.3 Efficiency and mean square error . . . . .                     | 277        |
| Exercises . . . . .   | 280        |
| Solutions . . . . .   | 286        |
| <b>25 Maximum Likelihood</b>  | <b>293</b> |
| 25.1 Likelihood . . . . .   | 293        |
| 25.2 Maximum likelihood estimation . . . . .                        | 294        |
| Exercises . . . . .   | 300        |
| Solutions . . . . .   | 312        |
| <b>26 Variance of Maximum Likelihood Estimator</b>                  | <b>325</b> |
| Exercises . . . . .   | 326        |
| Solutions . . . . .   | 327        |
| <b>27 Sufficient Statistics</b>                                     | <b>331</b> |
| Exercises . . . . .   | 334        |
| Solutions . . . . .   | 336        |
| <b>28 Hypothesis Testing</b>  | <b>339</b> |
| 28.1 Background . . . . .   | 339        |
| 28.2 Typical exam questions . . . . .                               | 342        |
| 28.2.1 Calculate significance or power . . . . .                    | 342        |
| 28.2.2 Determine critical values . . . . .                          | 344        |
| Exercises . . . . .   | 346        |
| Solutions . . . . .   | 355        |
| <b>29 Confidence Intervals and Sample Size</b>                      | <b>363</b> |
| 29.1 Confidence intervals . . . . .                                 | 363        |
| 29.2 Sample size . . . . .  | 364        |
| Exercises . . . . .   | 365        |
| Solutions . . . . .   | 367        |
| <b>30 Confidence Intervals for Means</b>                            | <b>371</b> |
| 30.1 $\chi^2$ distribution . . . . .                                | 371        |
| 30.2 Student's $t$ distribution . . . . .                           | 372        |
| 30.3 Testing the mean of a Bernoulli population . . . . .           | 373        |
| 30.4 Testing the difference of means from two populations . . . . . | 374        |
| 30.4.1 Two unpaired normal populations . . . . .                    | 374        |
| 30.4.2 Two paired normal populations . . . . .                      | 376        |
| 30.4.3 Two Bernoulli populations . . . . .                          | 376        |
| Exercises . . . . .   | 377        |
| Solutions . . . . .   | 385        |
| <b>31 Chi Square Tests</b>  | <b>391</b> |
| 31.1 One-dimensional chi-square . . . . .                           | 391        |
| 31.2 Two-dimensional chi-square . . . . .                           | 394        |
| Exercises . . . . .   | 396        |
| Solutions . . . . .   | 402        |

---

|  |            |
|--|------------|
| <b>32 Confidence Intervals for Variances</b>           | <b>407</b> |
| 32.1 Testing variances                                 | 407        |
| 32.2 Testing ratios of variances; the $F$ distribution | 408        |
| Exercises  | 409        |
| Solutions  | 415        |
| <b>33 Uniformly Most Powerful Critical Regions</b>     | <b>419</b> |
| Exercises  | 421        |
| Solutions  | 424        |
| <b>34 Likelihood Ratio Tests</b>                       | <b>427</b> |
| Exercises  | 429        |
| Solutions  | 433        |
| <b>35 Bayesian Estimation</b>                          | <b>437</b> |
| 35.1 Background  | 437        |
| 35.2 Loss functions                                    | 439        |
| 35.3 Interval estimates                                | 440        |
| Exercises  | 440        |
| Solutions  | 443        |
| <b>36 Beta-Bernoulli Conjugate Prior Pair</b>          | <b>445</b> |
| Exercises  | 446        |
| Solutions  | 450        |
| <b>37 Normal-Normal Conjugate Prior Pair</b>           | <b>453</b> |
| Exercises  | 454        |
| Solutions  | 456        |
| <b>38 Gamma-Poisson Conjugate Prior Pair</b>           | <b>459</b> |
| Exercises  | 460        |
| Solutions  | 467        |
| <b>39 Order Statistics</b>                             | <b>471</b> |
| Exercises  | 475        |
| Solutions  | 477        |
| <b>40 Sign Test</b>                                    | <b>485</b> |
| Exercises  | 488        |
| Solutions  | 490        |
| <b>41 Wilcoxon Tests</b>                               | <b>493</b> |
| 41.1 Signed rank test                                  | 493        |
| 41.2 Rank sum test                                     | 496        |
| Exercises  | 500        |
| Solutions  | 505        |
| <b>42 Rank Correlation Coefficients</b>                | <b>509</b> |
| Exercises  | 512        |
| Solutions  | 515        |

|           |  |            |
|-----------|--|------------|
| <b>43</b> | <b><i>q</i>-<i>q</i> Plots</b>   | <b>519</b> |
|           | Exercises . . . . .  | 520        |
|           | Solutions . . . . .  | 521        |
| <b>V</b>  | <b>Generalized Linear Model</b>  | <b>523</b> |
| <b>44</b> | <b>Types of Variables and Graphs</b>   | <b>525</b> |
|           | 44.1 Types of variables . . . . .  | 525        |
|           | 44.2 Graphs . . . . .  | 525        |
|           | Exercises . . . . .  | 527        |
|           | Solutions . . . . .  | 528        |
| <b>45</b> | <b>How a Generalized Linear Model Works</b>  | <b>529</b> |
|           | 45.1 Basic Models . . . . .  | 529        |
|           | 45.2 Exponential Family . . . . .  | 532        |
|           | 45.3 Poisson Response . . . . .  | 535        |
|           | Exercises . . . . .  | 535        |
|           | Solutions . . . . .  | 539        |
| <b>46</b> | <b>How a Generalized Linear Model Works: Categorical Response</b>                          | <b>541</b> |
|           | 46.1 Binomial Response . . . . .   | 541        |
|           | 46.2 Nominal Response . . . . .  | 545        |
|           | 46.3 Ordinal Response . . . . .  | 545        |
|           | Exercises . . . . .  | 549        |
|           | Solutions . . . . .  | 551        |
| <b>47</b> | <b>Generalized Linear Model: Estimating Parameters</b>                                     | <b>555</b> |
|           | Exercises . . . . .  | 561        |
|           | Solutions . . . . .  | 565        |
| <b>48</b> | <b>Generalized Linear Model: Measures of Fit</b>   | <b>569</b> |
|           | 48.1 Deviance . . . . .  | 569        |
|           | 48.2 Three tests for evaluating significance of explanatory variables . . . . .            | 572        |
|           | 48.3 Penalized loglikelihood tests . . . . .   | 575        |
|           | Exercises . . . . .  | 576        |
|           | Solutions . . . . .  | 579        |
| <b>49</b> | <b>Normal Linear Model: Standard Error, <math>R^2</math>, and <math>t</math> statistic</b> | <b>583</b> |
|           | 49.1 Standard error of the regression . . . . .  | 583        |
|           | 49.2 $R^2$ : the coefficient of determination . . . . .                                    | 584        |
|           | 49.3 $t$ statistic . . . . .   | 585        |
|           | Exercises . . . . .  | 588        |
|           | Solutions . . . . .  | 595        |
| <b>50</b> | <b>Normal Linear Model: <math>F</math> and VIF</b>   | <b>601</b> |
|           | 50.1 $F$ statistic . . . . .   | 601        |
|           | 50.2 Collinearity of explanatory variables; VIF . . . . .                                  | 603        |
|           | Exercises . . . . .  | 604        |
|           | Solutions . . . . .  | 614        |

|   |            |
|---|------------|
| <b>51 Normal Linear Model: Validation</b>               | <b>621</b> |
| 51.1 Validating model assumptions . . . . .             | 621        |
| 51.2 Influential points . . . . .                       | 623        |
| Exercises . . . . .                                     | 624        |
| Solutions . . . . .                                     | 627        |
| <b>52 Normal Linear Model: Predictions</b>              | <b>629</b> |
| 52.1 Matrix multiplication of $x^T Ax$ . . . . .        | 630        |
| Exercises . . . . .                                     | 630        |
| Solutions . . . . .                                     | 632        |
| <b>53 ANOVA</b>   | <b>635</b> |
| 53.1 One-factor ANOVA . . . . .                         | 635        |
| 53.2 Randomized block design . . . . .                  | 640        |
| 53.3 Two-factor ANOVA with replication . . . . .        | 641        |
| 53.4 ANCOVA . . . . .                                   | 644        |
| 53.5 Log-linear models . . . . .                        | 644        |
| Exercises . . . . .                                     | 644        |
| Solutions . . . . .                                     | 655        |
| <b>54 Generalized Linear Models: Measures of Fit II</b> | <b>663</b> |
| 54.1 Goodness of fit tests . . . . .                    | 663        |
| 54.2 Residuals . . . . .                                | 666        |
| 54.3 Overdispersion . . . . .                           | 667        |
| Exercises . . . . .                                     | 667        |
| Solutions . . . . .                                     | 671        |
| <b>VI Time Series</b>                                   | <b>675</b> |
| <b>55 Time Series: Trend and Seasonality</b>            | <b>677</b> |
| Exercises . . . . .                                     | 679        |
| Solutions . . . . .                                     | 681        |
| <b>56 Time Series: Correlation</b>                      | <b>683</b> |
| 56.1 Second order properties of a time series . . . . . | 683        |
| 56.2 Relationships of different time series . . . . .   | 684        |
| Exercises . . . . .                                     | 686        |
| Solutions . . . . .                                     | 689        |
| <b>57 Time Series: White Noise and Random Walks</b>     | <b>691</b> |
| 57.1 White noise . . . . .                              | 691        |
| 57.2 Random walks . . . . .                             | 691        |
| Exercises . . . . .                                     | 692        |
| Solutions . . . . .                                     | 693        |
| <b>58 Time Series: Autoregressive Models</b>            | <b>695</b> |
| Exercises . . . . .                                     | 699        |
| Solutions . . . . .                                     | 701        |
| <b>59 Time Series: Regression</b>                       | <b>703</b> |

|   |            |
|---|------------|
| Exercises . . . . .   | 706        |
| Solutions . . . . .   | 708        |
| <b>60 Time Series: Moving Average Models</b>  | <b>711</b> |
| Exercises . . . . .   | 714        |
| Solutions . . . . .   | 718        |
| <b>61 Time Series: ARMA Models</b>  | <b>723</b> |
| Exercises . . . . .   | 726        |
| Solutions . . . . .   | 728        |
| <b>62 Time Series: ARIMA and SARIMA models</b>  | <b>731</b> |
| Exercises . . . . .   | 732        |
| Solutions . . . . .   | 734        |
| <br>  |            |
| <b>VII Practice Exams</b>   | <b>737</b> |
| <br>  |            |
| <b>1 Practice Exam 1</b>  | <b>739</b> |
| <br>  |            |
| <b>2 Practice Exam 2</b>  | <b>753</b> |
| <br>  |            |
| <b>3 Practice Exam 3</b>  | <b>769</b> |
| <br>  |            |
| <b>4 Practice Exam 4</b>  | <b>783</b> |
| <br>  |            |
| <b>5 Practice Exam 5</b>  | <b>799</b> |
| <br>  |            |
| <b>6 Practice Exam 6</b>  | <b>815</b> |
| <br>  |            |
| <b>Appendices</b>   | <b>833</b> |
| <br>  |            |
| <b>A Solutions to the Practice Exams</b>  | <b>835</b> |
| Solutions for Practice Exam 1 . . . . .   | 835        |
| Solutions for Practice Exam 2 . . . . .   | 843        |
| Solutions for Practice Exam 3 . . . . .   | 853        |
| Solutions for Practice Exam 4 . . . . .   | 864        |
| Solutions for Practice Exam 5 . . . . .   | 875        |
| Solutions for Practice Exam 6 . . . . .   | 887        |
| <br>  |            |
| <b>B Solutions to Statistics and Stochastic Process Questions on Old CAS 3 and 3L Exams</b> | <b>897</b> |
| B.1 Solutions to CAS Exam 3, Spring 2005 . . . . .  | 897        |
| B.2 Solutions to CAS Exam 3, Fall 2005 . . . . .  | 901        |
| B.3 Solutions to CAS Exam 3, Spring 2006 . . . . .  | 903        |
| B.4 Solutions to CAS Exam 3, Fall 2006 . . . . .  | 908        |
| B.5 Solutions to CAS Exam 3, Spring 2007 . . . . .  | 912        |
| B.6 Solutions to CAS Exam 3, Fall 2007 . . . . .  | 915        |
| B.7 Solutions to CAS Exam 3L, Spring 2008 . . . . .   | 918        |
| B.8 Solutions to CAS Exam 3L, Fall 2008 . . . . .   | 921        |
| B.9 Solutions to CAS Exam 3L, Spring 2009 . . . . .   | 925        |
| B.10 Solutions to CAS Exam 3L, Fall 2009 . . . . .  | 928        |



---

|  |             |
|--|-------------|
| B.11 Solutions to CAS Exam 3L, Spring 2010 . . . . .                       | 932         |
| B.12 Solutions to CAS Exam 3L, Fall 2010 . . . . .                         | 936         |
| B.13 Solutions to CAS Exam 3L, Spring 2011 . . . . .                       | 939         |
| B.14 Solutions to CAS Exam 3L, Fall 2011 . . . . .                         | 943         |
| B.15 Solutions to CAS Exam 3L, Spring 2012 . . . . .                       | 946         |
| B.16 Solutions to CAS Exam 3L, Fall 2012 . . . . .                         | 949         |
| B.17 Solutions to CAS Exam 3L, Spring 2013 . . . . .                       | 952         |
| B.18 Solutions to CAS Exam 3L, Fall 2013 . . . . .                         | 956         |
| B.19 Solutions to CAS Exam LC, Spring 2014 . . . . .                       | 959         |
| B.20 Solutions to CAS Exam ST, Spring 2014 . . . . .                       | 961         |
| B.21 Solutions to CAS Exam LC, Fall 2014 . . . . .                         | 967         |
| B.22 Solutions to CAS Exam ST, Fall 2014 . . . . .                         | 968         |
| B.23 Solutions to CAS Exam LC, Spring 2015 . . . . .                       | 975         |
| B.24 Solutions to CAS Exam ST, Spring 2015 . . . . .                       | 976         |
| B.25 Solutions to CAS Exam S, Fall 2015 . . . . .                          | 981         |
| B.26 Solutions to CAS Exam LC, Fall 2015 . . . . .                         | 989         |
| B.27 Solutions to CAS Exam ST, Fall 2015 . . . . .                         | 991         |
| B.28 Solutions to CAS Exam S, Spring 2016 . . . . .                        | 996         |
| B.29 Solutions to CAS Exam LC, Spring 2016 . . . . .                       | 1004        |
| B.30 Solutions to CAS Exam ST, Spring 2016 . . . . .                       | 1006        |
| B.31 Solutions to CAS Exam S, Fall 2016 . . . . .                          | 1010        |
| B.32 Solutions to CAS Exam S, Spring 2017 . . . . .                        | 1018        |
| <br>   |             |
| <b>C Lessons Corresponding to Questions on Released and Practice Exams</b> | <b>1027</b> |



---

---

## Lesson 2

# Markov Chains: Chapman-Kolmogorov Equations

---

**Reading:** Ross 4.1–4.2, 4.5.1–4.5.2

A *stochastic process* is an infinite sequence of random variables indexed by the nonnegative number  $t$ , which represents time. Examples of stochastic processes  $X(t)$  are (I'm assuming that the following processes may be considered random.)

- (1)  $X(t)$  is the population of the world at time  $t$ .
- (2)  $X(t)$  is the number pointed to by your electricity meter at time  $t$ .
- (3)  $X(t)$  is the number of people who have arrived at a party by time  $t$ .
- (4)  $X(t)$  is your bank account at time  $t$ .

In the next few lessons, we will write  $X_t$  instead of  $X(t)$ .<sup>1</sup> When  $t$  only assumes integral values, the process is discrete. When  $t$  may assume any real value, the process is continuous. The next few lessons discuss discrete stochastic processes.

When the value of  $X_u$ ,  $u > t$ , only depends on the value of  $X_t$  and not on any  $X_i$  with  $i < t$ , the process is called a *Markov chain*. In other words, a Markov chain is a stochastic process with no memory. For a Markov chain, the value of  $X_t$  is called the *state* of the system. For a discrete Markov chain, the value of  $X_{t+1}$  depends only on the value of  $X_t$  and not on the value of  $X_u$  for any  $u < t$ .

Most of our Markov chains will be finite. That means that the number of possible states is finite. We will usually number the possible states starting with 1, as Ross does. But occasionally we will start the numbering at 0.

For a finite Markov chain, we can define the chain with a transition probability matrix. For every  $i$  and  $j$ , this matrix specifies  $P_{ij} = \Pr(X_t = j \mid X_{t-1} = i)$ . If there are  $n$  states, it is an  $n \times n$  matrix. When this matrix is not a function of  $t$ , we say that the Markov chain is *homogeneous*; otherwise the chain is nonhomogeneous. We will always assume our chains are homogeneous, unless specified otherwise.<sup>2</sup>

As an example of a transition probability matrix, suppose a Markov chain has two states, numbered 1 and 2. Suppose the following is the transition probability matrix:

$$\begin{pmatrix} 0.6 & 0.4 \\ 0.3 & 0.7 \end{pmatrix}$$

This means that for an individual in state 1 at time 0, the probability that the individual is in state 1 at time 1 is 0.6 and the probability that the individual is in state 2 at time 1 is 0.4. Also,  $P_{21} = 0.3$  and  $P_{22} = 0.7$ . The matrix entries in each row must add up to 1, but the entries in columns do not have this requirement.

---

<sup>1</sup>We will use whatever notation Ross uses. He uses  $X_t$  when discussing discrete Markov chains, but switches to  $X(t)$  when discussing Poisson processes.

<sup>2</sup>The syllabus in its learning objectives says "For discrete and continuous Markov Chains under both homogeneous and nonhomogeneous states. . .". First, there is no such thing as a nonhomogeneous state; only the chain itself can be non-homogeneous. Second, the syllabus implies that you are responsible for nonhomogeneous Markov chains, but nonhomogeneous Markov chains are not discussed in the Ross textbook, which is the only reading required by the syllabus. So go figure it out.

**EXAMPLE 2A** An auto insurance policy has two ratings, Standard (state 1) and Preferred (state 2). Moves between ratings are modeled as a Markov chain. Policyholders transition from one state to another in accordance with the following transition probability matrix:

$$\begin{pmatrix} 0.75 & 0.25 \\ 0.40 & 0.60 \end{pmatrix}$$

Calculate the probability that a Standard insured remains Standard for 2 years and then becomes Preferred.

**ANSWER:** The probability of remaining Standard for one year is 0.75. Since this occurs twice and the two years are independent (Markov chains are memoryless), the probability of remaining Standard for two years is  $(0.75^2) = 0.5625$ . The probability of becoming Preferred in the third year given Standard in the second year is 0.25. So the probability of remaining Standard for two years and then transitioning to Preferred is  $0.5625(0.25) = \boxed{0.140625}$ .  $\square$

It is possible for a Markov chain to take into account states in previous periods by incorporating them into the current period.

**EXAMPLE 2B** An auto insurance policy has two ratings, Standard (state 1) and Preferred (state 2). The probability that a policyholder is in a state depends on the states of the previous two years as follows:

| State two years ago | State one year ago | Probability of Standard |
|---------------------|--------------------|-------------------------|
| Standard            | Standard           | 0.9                     |
| Standard            | Preferred          | 0.4                     |
| Preferred           | Standard           | 0.8                     |
| Preferred           | Preferred          | 0.2                     |

Develop a Markov transition probability matrix for this situation.

**ANSWER:** In this example, the probability of Preferred is the complement of the probability of Standard. Let the four states be:

1. Standard last year, Standard this year
2. Standard last year, Preferred this year
3. Preferred last year, Standard this year
4. Preferred last year, Preferred this year

Then the following matrix gives the transition probabilities:

$$\begin{pmatrix} 0.9 & 0.1 & 0 & 0 \\ 0 & 0 & 0.4 & 0.6 \\ 0.8 & 0.2 & 0 & 0 \\ 0 & 0 & 0.2 & 0.8 \end{pmatrix}$$

Do you see how this works? For example, if the policyholder is Standard for two years, then next year the probability he will be Standard is 0.9, in which case he will be Standard for two years, while the probability he will be Preferred is the complement or 0.1, in which case in the following year he will be Standard two years ago and Preferred one year ago. It is impossible to move to the Preferred/Standard or Preferred/Preferred since he is Standard in the previous year.  $\square$

## 2.1 Chapman-Kolmogorov equations

One of the first things we want to calculate for a Markov chain is the probability of state  $j$  at time  $t$  given state  $i$  at time 0. One could calculate this probability by exhaustively listing all paths from  $i$  to  $j$  and summing up the probabilities of each path. A more systematic way to do the calculation involves the Chapman-Kolmogorov equations. These equations state that the probability of state  $j$  at time  $t$ , given state  $i$  at time 0, is the sum of the products of probabilities of being in state  $k$  at time  $u$  given state  $i$  at time 0, where  $u$  is a fixed integer satisfying  $0 < u < t$ , times probabilities of being in state  $j$  at time  $t$  given state  $k$  at time  $u$ . Let  $P_{ij}^t$  be the probability of state  $j$  at time  $t$  given state  $i$  at time 0. Then the Chapman-Kolmogorov equations are

Chapman-Kolmogorov Equations

$$P_{ij}^t = \sum_{k=1}^n P_{ik}^u P_{kj}^{t-u} \quad (2.1)$$

Let  $\mathbf{P}$  be the transition probability matrix, and let  $\mathbf{P}^{(k)}$  be the  $k$ -step transition probability matrix. In other words,  $P_{ij}^{(k)}$  is the probability of state  $j$  at time  $k$  given state  $i$  at time 0. The sum on the right of the Chapman-Kolmogorov equations are matrix multiplication,<sup>3</sup> so we can write the Chapman-Kolmogorov equations this way:

$$\mathbf{P}^{(t)} = \mathbf{P}^{(u)} \mathbf{P}^{(t-u)}$$

This implies

$$\mathbf{P}^{(t)} = \mathbf{P}^t$$

For example, suppose you are given the following transition probability matrix:

$$\mathbf{P} = \begin{pmatrix} 0.3 & 0.7 \\ 0.5 & 0.5 \end{pmatrix}$$

Then the two-step transition probability matrix is

$$\mathbf{P}^2 = \begin{pmatrix} 0.3 & 0.7 \\ 0.5 & 0.5 \end{pmatrix} \begin{pmatrix} 0.3 & 0.7 \\ 0.5 & 0.5 \end{pmatrix} = \begin{pmatrix} 0.44 & 0.56 \\ 0.40 & 0.60 \end{pmatrix}$$

This means that the probability that someone in state 1 at time 1 is in state 2 at time 3 is 0.56. The three-step transition probability matrix is

$$\mathbf{P}^3 = \begin{pmatrix} 0.44 & 0.56 \\ 0.40 & 0.60 \end{pmatrix} \begin{pmatrix} 0.3 & 0.7 \\ 0.5 & 0.5 \end{pmatrix} = \begin{pmatrix} 0.412 & 0.588 \\ 0.42 & 0.58 \end{pmatrix}$$

and the six-step transition probability matrix is

$$\mathbf{P}^6 = \begin{pmatrix} 0.412 & 0.588 \\ 0.42 & 0.58 \end{pmatrix} \begin{pmatrix} 0.412 & 0.588 \\ 0.42 & 0.58 \end{pmatrix} = \begin{pmatrix} 0.416704 & 0.583296 \\ 0.41664 & 0.58336 \end{pmatrix}$$

So the probability that someone in state 2 at time 3 is in state 2 at time 9 is 0.58336. You may notice that  $\mathbf{P}^{(6)}$  isn't much different from  $\mathbf{P}^{(3)}$ . We will discuss why they're not much different in Lesson 4.

<sup>3</sup>Recall that to multiply matrices  $\mathbf{AB} = \mathbf{C}$ , you set  $c_{ij} = \sum_{k=1}^n a_{ik} b_{kj}$ . You can only multiply matrices if the number of columns of the first equals the number of rows of the second.

Most of the time, it will be overkill to multiply matrices. Instead, you should keep track of the state probability vector at each time. The state probability vector is the  $n$ -component vector having the probabilities of each state at a given time. By keeping track of the state probability vector, you only need to multiply a vector by a matrix, rather than an entire matrix by a matrix.

**EXAMPLE 2C** The interest rate at a given time depends only on the interest rate one year ago. You are given the following probabilities:

| Previous year interest rate | Current year interest rate |     |     |
|-----------------------------|----------------------------|-----|-----|
|                             | 5%                         | 6%  | 7%  |
| 5%                          | 0.6                        | 0.2 | 0.2 |
| 6%                          | 0.7                        | 0.2 | 0.1 |
| 7%                          | 0.5                        | 0.3 | 0.2 |

Calculate the probability that the interest rate is 7% after 3 years if it is 5% currently.

**ANSWER:** We number the states 1 for 5%, 2 for 6%, and 3 for 7%. The initial state probability vector is  $(1 \ 0 \ 0)$  since the system is definitely in state 1 initially. After 1 year, the state probability vector is  $(0.6 \ 0.2 \ 0.2)$ , since that is what the first row of transition probabilities is. After 2 years, multiply this vector by the transition matrix:

$$(0.6 \ 0.2 \ 0.2) \begin{pmatrix} 0.6 & 0.2 & 0.2 \\ 0.7 & 0.2 & 0.1 \\ 0.5 & 0.3 & 0.2 \end{pmatrix} = (0.6 \ 0.22 \ 0.18)$$

We don't have to multiply this by the transition probability matrix again to get the third year state probability vector, since we only want the last component of that vector, the probability of state 3 at time 3. We just have to calculate the last component of the state probability vector at time 3, which is the state probability vector at time 2 times the last column of the transition probability matrix:

$$0.6(0.2) + 0.22(0.1) + 0.18(0.2) = \boxed{0.178} \quad \square$$

To calculate the probability of a transition from state  $i$  to state  $j$  in year  $n$ , calculate the probability of state  $i$  at time  $k - 1$  and then calculate the probability of transition to state  $j$ .

**EXAMPLE 2D** The interest rate at a given time depends only on the interest rate one year ago. You are given the following probabilities:

| Previous year interest rate | Current year interest rate |     |     |
|-----------------------------|----------------------------|-----|-----|
|                             | 5%                         | 6%  | 7%  |
| 5%                          | 0.6                        | 0.2 | 0.2 |
| 6%                          | 0.7                        | 0.2 | 0.1 |
| 7%                          | 0.5                        | 0.3 | 0.2 |

Calculate the probability that the interest rate goes from 6% to 5% in year 3 if it is 5% initially.

**ANSWER:** In the previous example, we calculated the state probability vector at time 2:  $(0.6 \ 0.22 \ 0.18)$ . Also,  $P_{21} = 0.7$ . So the answer is  $(0.22)(0.7) = \boxed{0.154}$ .  $\square$

An exam question may ask an insurance-related question by putting a value on either being in a state or on transitioning from one state to another. To calculate the total value, sum up probabilities of being in a state or transitioning from state to state times the value assigned.

**EXAMPLE 2E** You are given:

- The Driveco Insurance Company classifies all of its auto customers into two classes: preferred with annual expected losses of 400 and standard with annual expected losses of 900.
- There will be no change in the expected losses for either class over the next three years.
- The one year transition matrix between driver classes is given by:

| Driver's class | Driver's class in year $k$ |          |
|----------------|----------------------------|----------|
|                | Preferred                  | Standard |
| Preferred      | 0.85                       | 0.15     |
| Standard       | 0.60                       | 0.40     |

- All drivers insured with Driveco at the start of the period will remain insured for the following  
Calculate the expected total losses paid over a period of 3 years for a driver who is initially standard.

**ANSWER:** In the first year the driver's class is Standard and expected losses are 900.

In the second year, the driver's class is Standard with probability 0.4 and Preferred with probability 0.6. Expected losses are  $0.4(900) + 0.6(400) = 600$ .

Multiplying the state probability vector at the end of year 1  $(0.6 \ 0.4)$  by the transition probability matrix, we get

$$(0.6 \ 0.4) \begin{pmatrix} 0.85 & 0.15 \\ 0.60 & 0.40 \end{pmatrix} = (0.75 \ 0.25)$$

Expected losses in the third year are  $0.25(900) + 0.75(400) = 525$ .

Total expected losses for 3 years are  $900 + 600 + 525 = \boxed{2025}$ . □

## 2.2 Gambler's ruin

An example of a Markov chain is gambler's ruin. In this chain, a gambler starts out with  $k$  chips. At each round, the gambler may win one chip with probability  $p$  or lose one chip with probability  $q = 1 - p$ . The game ends when the gambler loses all his chips or has  $N > k$  chips. Thus the game is a Markov chain with a state defined as a number of chips. The state may be  $0, 1, \dots, N$ . We will calculate the probability of reaching state  $N$ .

Let  $P_i$  be the probability of reaching  $N$  given that the gambler currently has  $i$  chips. Then for  $0 < i < N$ , by conditioning on winning one round,

$$P_i = pP_{i+1} + qP_{i-1}$$

Then we ingeniously express  $P_i$  as  $pP_i + qP_i$  and get

$$\begin{aligned} pP_i + qP_i &= pP_{i+1} + qP_{i-1} \\ p(P_{i+1} - P_i) &= q(P_i - P_{i-1}) \\ P_{i+1} - P_i &= \frac{q}{p}(P_i - P_{i-1}) \end{aligned}$$

and repeatedly substituting  $P_j - P_{j-1} = (q/p)(P_{j-1} - P_{j-2})$  into the parentheses on the right side, for  $j = i, i-1, \dots, 1$ , we get

$$P_{i+1} - P_i = \left(\frac{q}{p}\right)^i (P_1 - P_0) = \left(\frac{q}{p}\right)^i P_1$$

since  $P_0 = 0$ . Adding up these equations from  $i = 1$  to  $i = j - 1$ , we get

$$P_j = \sum_{i=0}^{j-1} \left(\frac{q}{p}\right)^i P_1$$

Let  $r = q/p$ . If  $r = 1$ , then the sum on the right is  $jP_1$ . Otherwise it is

$$P_j = \frac{r^j - 1}{r - 1} P_1$$

In particular, with  $j = N$ ,

$$P_N = \begin{cases} NP_1 & r = 1 \\ \frac{r^N - 1}{r - 1} P_1 & r \neq 1 \end{cases}$$

But  $P_N = 1$ , so we can solve for  $P_1$ :

$$P_1 = \begin{cases} \frac{1}{N} & r = 1 \\ \frac{r - 1}{r^N - 1} & r \neq 1 \end{cases}$$

and it follows that

$$P_j = \begin{cases} \frac{j}{N} & r = 1 \\ \frac{r^j - 1}{r^N - 1} & r \neq 1 \end{cases} \quad (2.2)$$

Notice that  $r = 1$  if and only if  $p = 1/2$ .

**EXAMPLE 2F** A gambler has 10 chips. The gambler bets 1 chip at each round of a game. If the gambler wins, she gets 1 chip in addition to the one she bet. Otherwise she loses the chip.

The gambler has found a strategy which gives her a 0.6 probability of winning at each round. She will keep betting until she has 25 chips or until she loses all her chips.

Determine the probability that she succeeds in having 25 chips.

**ANSWER:** Applying formula (2.2) with  $r = q/p = 0.4/0.6 = 2/3$  and  $N = 25$ ,

$$P_{10} = \frac{(2/3)^{10} - 1}{(2/3)^{25} - 1} = \boxed{0.9827} \quad \square$$



**Quiz 2-1** <sup>4</sup> You have 6 chips and your friend has 4 chips. You bet one chip against your friend. You have a 1/2 probability of winning. The gambling continues until one of you runs out of chips.

What is the probability that you will end up with 10 chips?

## 2.3 Algorithmic efficiency

Consider the simplex method of linear programming, which involves maximizing a linear expression. At each step, the algorithm moves around the corners of a convex polyhedron from one corner to an adjacent corner that is better, in the sense of increasing the value of the expression. This method is surprisingly

<sup>4</sup>Solutions to quizzes are at the end of the lesson after exercise solutions.



**Table 2.1:** Summary of Formulas in this Lesson

|   |   |
|---|---|
| Chapman-Kolmogorov equations  | $P_{ij}^t = \sum_{k=1}^n P_{ik}^u P_{kj}^{t-u} \quad (2.1)$   |
| Gambler's ruin probabilities  | $P_j = \begin{cases} \frac{j}{N} & r = 1 \\ \frac{r^j - 1}{r^N - 1} & r \neq 1 \end{cases} \quad (2.2)$ |
| where $p$ is the probability of success at each round and $r = q/p$ .                                     |   |
| Algorithmic efficiency, with $N_j =$ number of steps from $j^{\text{th}}$ solution to best solution.      |   |
|   | $\mathbf{E}[N_j] = \sum_{i=1}^{j-1} \frac{1}{i} \quad (2.3)$  |
|   | $\text{Var}(N_j) = \sum_{i=1}^{j-1} \left(\frac{1}{i}\right) \left(1 - \frac{1}{i}\right) \quad (2.4)$  |
| As $j \rightarrow \infty$ , $\mathbf{E}[N_j] \rightarrow \ln j$ and $\text{Var}(N_j) \rightarrow \ln j$ . |   |

efficient. To better understand the efficiency of this, or any similar, algorithm, we create a simple Markov chain model for any algorithm that is trying to maximize an expression and has a finite number of possible solutions. In this Markov chain model, we assume that if we are at the  $j^{\text{th}}$  best solution, the algorithm randomly selects a better solution, with a  $1/(j-1)$  probability for each possible better solution. Assuming that this holds for every  $j$ , what is the expected number of steps and the variance of the number of steps to the best solution?

It turns out that if  $N_j$  is the number of steps from the  $j^{\text{th}}$  best solution to the best solution, then

$$\mathbf{E}[N_j] = \sum_{i=1}^{j-1} \frac{1}{i} \quad (2.3)$$

$$\text{Var}(N_j) = \sum_{i=1}^{j-1} \left(\frac{1}{i}\right) \left(1 - \frac{1}{i}\right) \quad (2.4)$$

Both of these expressions approach  $\ln j$  as  $j \rightarrow \infty$ .

## Exercises

2.1. Interest rates are either 5%, 6%, or 7%. Changes in interest rate are modeled as a Markov chain with the following transition probabilities:

|      | To  |     |     |
|------|-----|-----|-----|
| From | 5%  | 6%  | 7%  |
| 5%   | 0.6 | 0.4 | 0   |
| 6%   | 0.2 | 0.6 | 0.2 |
| 7%   | 0.1 | 0.4 | 0.5 |

Interest rates were 6% both last year and this year.

Calculate the probability that they will be 5% three years from now.

2.2. A taxi services a city. Rides are between uptown, downtown, and the airport. The probabilities of destinations are as follows:

|          | To       |        |         |
|----------|----------|--------|---------|
| From     | Downtown | Uptown | Airport |
| Downtown | 0.5      | 0.3    | 0.2     |
| Uptown   | 0.3      | 0.5    | 0.2     |
| Airport  | 0.6      | 0.4    | 0       |

The taxi starts at the airport, and always picks up a passenger at the destination of the previous passenger.

Calculate the probability that the third and fourth trips will both be within downtown.

2.3. If you set an alarm clock, you wake up at 6. Otherwise you wake up at 6 with probability 0.3 and at 8 with probability 0.7. When you wake up at 6, you set the alarm clock the next day 50% of the time. When you wake up at 8, you set the alarm clock the next day 80% of the time.

You wake up at 6 today.

Calculate the probability of waking up at 6 two days from now.

2.4. For an auto insurance coverage, drivers are classified into class A and class B. You are given:

- The probability that a driver in class A at time  $t$  will be reclassified into class B at time  $t + 1$  is 0.2.
- The probability that a driver in class B at time  $t$  will be in class B at time  $t + 2$  is 0.65.
- Reclassification only occurs at integral times.

Calculate the probability that a driver in class B at time  $t$  will be reclassified into Class A at time  $t + 1$ .

2.5. [CAS3-F04:15] An auto insured who was claim-free during a policy period will be claim-free during the next policy period with probability 0.9.

An auto insured who was not claim-free during a policy period will be claim-free during the next policy period with probability 0.7.

What is the probability that an insured who was claim-free during the policy period 0 will incur a claim during policy period 3?

- A. 0.120                      B. 0.124                      C. 0.128                      D. 0.132                      E. 0.136

Use the following information for questions 2.6 and 2.7:

A life insurance policy has a disability waiver provision under which premium is waived if the insured is disabled. You are given:

- The probability of death each year is 0.1.
- For a life active (not disabled) at the beginning of a year, the probability of being disabled at the end of the year is 0.2.
- For a life disabled at the beginning of a year, the probability of recovery from disability in that year is 0.3.

All changes in state (deaths, disabilities, recoveries) occur at the end of the year.

- 2.6. Calculate the probability of a currently active life being active at the end of three years.
- 2.7. Calculate the probability of a currently active life getting disabled and recovering within four years.

2.8. In a Continuing Care Retirement Community (CCRC), a resident may be in the Independent Living Unit, Temporary Health Care Unit, or Permanent Health Care unit. The probabilities of transferring between these units, or leaving the CCRC, are as follows:

| From                  | To                 |                       |                       |              |
|-----------------------|--------------------|-----------------------|-----------------------|--------------|
|                       | Independent Living | Temporary Health Care | Permanent Health Care | Leaving CCRC |
| Independent Living    | 0.5                | 0.2                   | 0.2                   | 0.1          |
| Temporary Health Care | 0.5                | 0.3                   | 0.1                   | 0.1          |
| Permanent Health Care | 0                  | 0                     | 0.8                   | 0.2          |

Calculate the probability of a resident currently in the Independent Living unit leaving the CCRC in the third year.

2.9. In a Continuing Care Retirement Community (CCRC), a resident may be in the Independent Living unit (ILU), Temporary Health Care unit (THCU), or Permanent Health Care unit (PHCU). The conditional probabilities of transferring between these units or leaving the CCRC during each year are as follows:

| Transition                | Year 1 | Year 2 | Year 3 | Year 4 |
|---------------------------|--------|--------|--------|--------|
| From ILU to THCU          | 0.1    | 0.2    | 0.2    | 0.2    |
| From ILU to PHCU          | 0.1    | 0.2    | 0.2    | 0.2    |
| From ILU to leaving CCRC  | 0.1    | 0.2    | 0.3    | 0.4    |
| From THCU to ILU          | 0.5    | 0.5    | 0.5    | 0.5    |
| From THCU to PHCU         | 0.3    | 0.3    | 0.3    | 0.3    |
| From THCU to leaving CCRC | 0.2    | 0.2    | 0.2    | 0.2    |
| From PHCU to leaving CCRC | 0.4    | 0.4    | 0.4    | 0.4    |

Calculate the probability of a resident currently in the ILU transferring from the THCU to the ILU during Year 4.

2.10. [3-S00:38] For Shoestring Swim Club, with three possible financial states at the end of each year:

- State 0 means cash of 1500. If in state 0, aggregate member charges for the next year are set equal to operating expenses.
- State 1 means cash of 500. If in state 1, aggregate member charges for the next year are set equal to operating expenses plus 1000, hoping to return the club to state 0.
- State 2 means cash less than 0. If in state 2, the club is bankrupt and remains in state 2.
- The club is subject to four risks each year. These risks are independent. Each of the four risks occurs at most once per year, but may recur in a subsequent year.
- Three of the four risks each have a cost of 1000 and a probability of occurrence 0.25 per year.
- The fourth risk has a cost of 2000 and a probability of occurrence 0.10 per year.
- Aggregate member charges are received at the beginning of the year.
- $i = 0$

Calculate the probability that the club is in state 2 at the end of three years, given that it is in state 0 at time 0.

- A. 0.24                      B. 0.27                      C. 0.30                      D. 0.37                      E. 0.56

2.11. [3-F02:30] Nancy reviews the interest rates each year for a 30-year fixed mortgage issued on July 1. She models interest rate behavior by a Markov model assuming:

- Interest rates always change between years.
- The change in any given year is dependent on the change in prior years as follows:

| From year $t - 3$ to<br>year $t - 2$ | From year $t - 2$ to<br>year $t - 1$ | Probability that year $t$ will<br>increase from year $t - 1$ |
|--------------------------------------|--------------------------------------|--|
| Increase                             | Increase                             | 0.10   |
| Decrease                             | Decrease                             | 0.20   |
| Increase                             | Decrease                             | 0.40   |
| Decrease                             | Increase                             | 0.25   |

She notes that interest rates decreased from year 2000 to 2001 and from year 2001 to 2002.

Calculate the probability that interest rates will decrease from year 2003 to 2004.

- A. 0.76                      B. 0.79                      C. 0.82                      D. 0.84                      E. 0.87

2.12. [SOA3-F03:24] For a perpetuity-immediate with annual payments of 1:

- The sequence of annual discount factors follows a Markov chain with the following three states:

| State number                | 0    | 1    | 2    |
|-----------------------------|------|------|------|
| Annual discount factor, $v$ | 0.95 | 0.94 | 0.93 |

- The transition matrix for the annual discount factors is:

$$\begin{pmatrix} 0.0 & 1.0 & 0.0 \\ 0.9 & 0.0 & 0.1 \\ 0.0 & 1.0 & 0.0 \end{pmatrix}$$

$Y$  is the present value of the perpetuity payments when the initial state is 1.

Calculate  $E[Y]$ .

- A. 15.67                      B. 15.71                      C. 15.75                      D. 16.82                      E. 16.86

**2.13. [CAS3-S04:18]** Loans transition through five states (Current, 30, 60, 90, and Foreclosed) based on the following matrix:

|            | Current | 30   | 60   | 90   | Foreclosed |
|------------|---------|------|------|------|------------|
| Current    | 0.80    | 0.20 | 0.00 | 0.00 | 0.00       |
| 30         | 0.50    | 0.00 | 0.50 | 0.00 | 0.00       |
| 60         | 0.25    | 0.00 | 0.00 | 0.75 | 0.00       |
| 90         | 0.10    | 0.00 | 0.00 | 0.00 | 0.90       |
| Foreclosed | 0.00    | 0.00 | 0.00 | 0.00 | 1.00       |

The transitions happen monthly.

Out of 100,000 Current loans, how many are expected to be Foreclosed in six months?

- A. Less than 16,500
- B. At least 16,500, but less than 16,750
- C. At least 16,750, but less than 17,000
- D. At least 17,000, but less than 17,250
- E. At least 17,250

**2.14. [CAS3-S04:23]** A customer service department receives 0 or 1 complaint each day, depending on the number of complaints on the previous 2 days, as follows:

- If there were no complaints the past 2 days, then there will be no complaints today with probability 0.75.
- If there were no complaints 2 days ago but 1 complaint yesterday, then there will be no complaints today with probability 0.40.
- If there was 1 complaint 2 days ago but no complaints yesterday, then there will be no complaints today with probability 0.55.
- If there was 1 complaint on each of the past 2 days, then there will be no complaints today with probability 0.10.

Suppose there were no complaints 2 days ago and 1 complaint yesterday.

Calculate the probability that there will be at least 1 complaint over the next 2 days.

- A. 0.4375      B. 0.5700      C. 0.6975      D. 0.7800      E. 0.8400

**2.15.** On an auto insurance coverage, drivers are classified into two classes, 0 and 1. The probabilities that a driver in one class will be reclassified into the other class at the beginning of the next year are as follows:

| Transfer                | Year 1 | Year 2 | Year 3 | Year 4 |
|-------------------------|--------|--------|--------|--------|
| From class 0 to class 1 | 0.4    | 0.3    | 0.2    | 0.1    |
| From class 1 to class 0 | 0.2    | 0.2    | 0.1    | 0.1    |

Calculate the probability that a driver currently in class 0 will move from class 1 to class 0 in Year 3 or Year 4.

2.16. The transition matrix for a Markov chain with three states is

$$\begin{pmatrix} 0.3 & 0.5 & 0.2 \\ 0.6 & 0.3 & 0.1 \\ 0 & 0 & 1 \end{pmatrix}$$

For a group of 100 individuals in state 1 at time 0, calculate the variance of the number of individuals in state 1 at time 3.

2.17. You are given:

- The transition matrix for a Markov chain with three states, numbered 1, 2, and 3, is:

$$\begin{pmatrix} 0.8 & 0.1 & 0.1 \\ 0.7 & 0.2 & 0.1 \\ 0 & 0.1 & 0.9 \end{pmatrix}$$

- Time is measured in years.
- All individuals are in state 1 at time 0.
- Transitions occur in the middle of each year.

Calculate the probability that an individual in state 2 or 3 at time 1 is in state 1 at time 3.

2.18. [M-S05:11] For a Markov model with three states, Healthy (0), Disabled (1), and Dead (2):

- The annual transition matrix is given by

$$\begin{array}{c} 0 \quad 1 \quad 2 \\ \begin{array}{l} 0 \left[ \begin{array}{ccc} 0.70 & 0.20 & 0.10 \\ 1 \left[ \begin{array}{ccc} 0.10 & 0.65 & 0.25 \\ 2 \left[ \begin{array}{ccc} 0 & 0 & 1 \end{array} \right] \end{array} \right] \end{array} \right] \end{array}$$

- There are 100 lives at the start, all Healthy. Their future states are independent.

Calculate the variance of the number of the original 100 lives who die within the first two years.

- A. 11                      B. 14                      C. 17                      D. 20                      E. 23

**2.19. [M-F06:14]** A homogeneous Markov model has three states representing the status of the members of a population.

State 1 = healthy, no benefits

State 2 = disabled, receiving Home Health Care benefits

State 3 = disabled, receiving Nursing Home benefits

The annual transition matrix is given by:

$$\begin{pmatrix} 0.80 & 0.15 & 0.05 \\ 0.05 & 0.90 & 0.05 \\ 0.00 & 0.00 & 1.00 \end{pmatrix}$$

Transitions occur at the end of each year.

At the start of year 1, there are 50 members, all in state 1, healthy.

Calculate the variance of the number of those 50 members who will be receiving Nursing Home benefits during year 3.

- A. 2.3                      B. 2.7                      C. 4.4.                      D. 4.5                      E. 5.6

**2.20. [MLC-S07:16]** The number of coins Lucky Tom finds in successive blocks as he walks to work follows a homogeneous Markov model:

- States 0, 1, 2 correspond to 0, 1, or 2 coins found in a block.
- The transition matrix is:

$$\mathbf{P} = \begin{pmatrix} 0.2 & 0.5 & 0.3 \\ 0.1 & 0.6 & 0.3 \\ 0.1 & 0.5 & 0.4 \end{pmatrix}$$

- Tom found one coin in the first block today.

Calculate the probability that Tom will find at least 3 more coins in the next two blocks today.

- A. 0.43                      B. 0.45                      C. 0.47                      D. 0.49                      E. 0.51

2.21. [SOA3-F04:14] For a Markov model for an insured population:

- Annual transition probabilities between health states of individuals are as follows:

|            | Healthy | Sick | Terminated |
|------------|---------|------|------------|
| Healthy    | 0.7     | 0.1  | 0.2        |
| Sick       | 0.3     | 0.6  | 0.1        |
| Terminated | 0.0     | 0.0  | 1.0        |

- The mean annual healthcare cost each year for each health state is:

|            | Mean |
|------------|------|
| Healthy    | 500  |
| Sick       | 3000 |
| Terminated | 0    |

- Transitions occur at the end of the year.
- A premium of 800 is paid each year by an insured not in the terminated state.

Calculate the expected value of premiums less healthcare costs over the first 3 years for a new healthy insured.

- A. -390                      B. -200                      C. -20                      D. 160                      E. 340

2.22. For a Markov model for an insured population:

- Annual transition probabilities between health states of individuals are as follows:

|            | Healthy | Sick | Terminated |
|------------|---------|------|------------|
| Healthy    | 0.7     | 0.1  | 0.2        |
| Sick       | 0.3     | 0.6  | 0.1        |
| Terminated | 0.0     | 0.0  | 1.0        |

- The cost of a transition from Healthy to Sick is 500.

Calculate the expected cost of transitions over the first 3 years for a new healthy insured.

2.23. [CAS3-F03:26] A fair coin is flipped by a gambler with 10 chips. If the outcome is “heads”, the gambler wins 1 chip; if the outcome is “tails”, the gambler loses 1 chip.

The gambler will stop playing when he either has lost all of his chips or he reaches 30 chips.

Of the first ten flips, 4 are “heads” and 6 are “tails”.

Calculate the probability that the gambler will lose all of his chips, given the results of the first ten flips.

- A. Less than 0.75  
 B. At least 0.75, but less than 0.80  
 C. At least 0.80, but less than 0.85  
 D. At least 0.85, but less than 0.90  
 E. At least 0.90





**Figure 2.1:** Allosaur for exercise 2.24

2.24. [3-F02:7] For an allosaur with 10,000 calories stored at the start of a day:

- The allosaur uses calories uniformly at a rate of 5,000 per day. If his stored calories reach 0, he dies.
- Each day, the allosaur eats 1 scientist (10,000 calories) with probability 0.45 and no scientist with probability 0.55.
- The allosaur eats only scientists.
- The allosaur can store calories without limit until needed.

Calculate the probability that the allosaur ever has 15,000 or more calories stored.

- A. 0.54                      B. 0.57                      C. 0.60                      D. 0.63                      E. 0.66

2.25. Susan Gambler is trying to improve her gambling technique. She starts a game with two chips, and at each round has probability  $p$  of gaining one chip and probability  $1 - p$  of losing one chip. She would like to get  $p$  high enough so that her probability of gaining four chips for a total of six chips before losing all of her chips is 0.5.

Determine the smallest  $p$  she needs to accomplish this.

2.26. You are maximizing an expression using an iterative procedure. At each iteration, if there are  $n$  better solutions, the procedure selects a better solution randomly with probability  $1/n$  for each candidate solution. At the current point, there are 5 better solutions.

Calculate the expected number of iterations needed to obtain the best solution.

**2.27.** You are maximizing an expression. There are 200 potential solutions, none of them tied with any other solution. You have an algorithm. At the first iteration, the algorithm selects one of the 200 solutions at random. If that solution is not the best, then at each successive iteration, it selects a better solution, with each better solution equally likely.

Using the normal approximation with continuity correction, calculate the probability that more than 10 steps will be necessary to arrive at the best solution.

**Additional old CAS Exam S questions:** F15:8, S16:9, F16:14, S17:11,12

**Additional old CAS Exam 3/3L questions:** S05:36, F05:23, S06:29, F06:21,22, S07:28, S08:20, F08:19, S09:6, F09:9, S10:10,11, F10:8,9, S11:8, F11:8,9, S12:8, F12:8, S13:7,8, F13:8

**Additional old CAS Exam LC questions:** S14:10,11,15, F14:9,10,15, S15:9,10,15, F15:9,10,14,15, S16:9,10,15

**Sample Questions:** 2

## Solutions

**2.1.** The fact that interest rates were 6% last year is irrelevant, since Markov chains have no memory.

We need  $P_{21}^{(3)}$ . At time 1, the transition probability vector will be the second row of the transition probability matrix,  $(0.2 \ 0.6 \ 0.2)$ . The state vector at time 2 is:

$$(0.2 \ 0.6 \ 0.2) \begin{pmatrix} 0.6 & 0.4 & 0 \\ 0.2 & 0.6 & 0.2 \\ 0.1 & 0.4 & 0.5 \end{pmatrix} = (0.26 \ 0.52 \ 0.22)$$

The first component of the state vector at time 3, which is the probability of 5%, is

$$0.26(0.6) + 0.52(0.2) + 0.22(0.1) = \boxed{0.282}$$

**2.2.** We need the probability of transitioning from state #3 (the airport) to state #1 (downtown) in two periods, followed by two transitions from state #1 to state #1. The probability of transitioning from state #1 to state #1 is  $P_{11} = 0.5$ . The probability of transitioning from state #3 to state #1 in two periods is the product of the row vector for state #3, which is the state vector at the end of the first period, and the column vector for state #1, which is the vector of probabilities of transitioning from a state to state #1, or

$$(0.6)(0.5) + (0.4)(0.3) + (0)(0.6) = 0.42$$

Multiplying the three probabilities #3  $\rightarrow$  #1  $\rightarrow$  #1  $\rightarrow$  #1, we get  $(0.42)(0.5^2) = \boxed{0.105}$ .

**2.3.** Let waking up at 6 be state #1 and at 8 state #2. If you are in state #1, there's a 50% chance of setting an alarm clock and waking up at 6, and a 50% chance of not and then a 30% chance of waking up at 6, or  $0.5 + 0.5(0.3) = 0.65$  probability of remaining in state #1, and consequently 0.35 probability of moving to state #2. If you are in state #2, then the probability of moving to state #1, using the same logic, is  $0.8 + 0.2(0.3) = 0.86$ . So the transition probability matrix is

$$\mathbf{P} = \begin{pmatrix} 0.65 & 0.35 \\ 0.86 & 0.14 \end{pmatrix}$$

We want  $P_{11}^{(2)}$ , so we multiply the first row by the first column, or

$$(0.65)(0.65) + (0.35)(0.86) = \boxed{0.7235}$$

2.4. Let state #1 be classification in class A and state #2 classification in class B. Since there are only two states, the probability of changing state is 1 minus the probability of not changing state. The transition matrix is

$$\mathbf{P} = \begin{pmatrix} 0.8 & 0.2 \\ x & 1-x \end{pmatrix}$$

where  $x$  is what we want to calculate. We are given  $P_{22}^{(2)} = 0.65$ , and  $P_{22}^{(2)}$  is the product of the second row of  $\mathbf{P}$  by the second column of  $\mathbf{P}$ , so:

$$(x)(0.2) + (1-x)(1-x) = 0.65$$

$$x^2 - 1.8x + 0.35 = 0$$

$$x = \frac{1.8 - \sqrt{1.84}}{2} = \boxed{0.2218}$$

The other solution to the quadratic is more than 1 so it cannot be a probability.

2.5. The transition probability matrix is  $\mathbf{P} = \begin{pmatrix} 0.9 & 0.1 \\ 0.7 & 0.3 \end{pmatrix}$ . For two periods, the transition probabilities are

$$(0.9 \ 0.1) \begin{pmatrix} 0.9 & 0.1 \\ 0.7 & 0.3 \end{pmatrix} = (0.88 \ 0.12)$$

The probability of a claim is  $0.88(0.1) + 0.12(0.3) = \boxed{0.124}$ . (B)

2.6. Let the states be (1) Active, (2) Disabled, (3) Dead. The missing transition probabilities in each row are calculated as 1 minus the given transition probabilities. The probability of an active person staying active is  $1 - 0.1 - 0.2 = 0.7$ , and the probability of a disabled person remaining disabled is  $1 - 0.1 - 0.3 = 0.6$ . The transition probability matrix is

$$\mathbf{P} = \begin{pmatrix} 0.7 & 0.2 & 0.1 \\ 0.3 & 0.6 & 0.1 \\ 0 & 0 & 1 \end{pmatrix}$$

We want  $P_{11}^{(3)}$ . The state vector at the end of the first year is the first row of the matrix, and we multiply this by the matrix to get the state vector at the end of the second year.

$$(0.7 \ 0.2 \ 0.1) \begin{pmatrix} 0.7 & 0.2 & 0.1 \\ 0.3 & 0.6 & 0.1 \\ 0 & 0 & 1 \end{pmatrix} = (0.55 \ 0.26 \ 0.19)$$

Then  $P_{11}^{(3)} = 0.55(0.7) + 0.26(0.3) = \boxed{0.463}$ .

2.7. There are three ways of getting disabled and recovering within four years: getting disabled in the first year, the second year, or the third year. Note that the question does not indicate the state at time 4, which could be disabled or dead.

The first way has probability of disability 0.2, followed by probability of recovering within three years. The probability of recovering within three years is the probability of recovering in one year, or staying disabled one year and recovering the next year, or staying disabled for two years and recovering the next year, which is  $0.3 + 0.6(0.3) + 0.6^2(0.3) = 0.588$ . So the first way has probability  $(0.588)(0.2) = 0.1176$ .

The second way has probability of disability  $(0.7)(0.2) = 0.14$ , followed by recovery probability  $0.3 + 0.6(0.3) = 0.48$ , for a probability of  $(0.14)(0.48) = 0.0672$ .

The third way has probability of disability  $(0.7^2)(0.2) = 0.098$  and probability of recovery 0.3, for a probability of  $(0.098)(0.3) = 0.0294$ .

The answer is  $0.1176 + 0.0672 + 0.0294 = \boxed{0.2142}$ .

An alternative method<sup>5</sup> is to use a four-state Markov chain with the states:

1. Active and never disabled
2. Disabled
3. Previously disabled and recovered (even if dead)
4. Dead

The transition matrix is then

$$\begin{pmatrix} 0.7 & 0.2 & 0 & 0.1 \\ 0 & 0.6 & 0.3 & 0.1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Notice how the third state is made a sink. You cannot get out of that state even if you die or get disabled again. We want to track anyone who ever gets into that state.

Then the state vector after two years is

$$(0.7 \quad 0.2 \quad 0 \quad 0.1) \begin{pmatrix} 0.7 & 0.2 & 0 & 0.1 \\ 0 & 0.6 & 0.3 & 0.1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = (0.49 \quad 0.26 \quad 0.06 \quad 0.19)$$

The state vector after three years is

$$(0.49 \quad 0.26 \quad 0.06 \quad 0.19) \begin{pmatrix} 0.7 & 0.2 & 0 & 0.1 \\ 0 & 0.6 & 0.3 & 0.1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = (0.343 \quad 0.254 \quad 0.138 \quad 0.265)$$

and the answer is the third component of the state vector in the fourth year:

$$0.343(0) + 0.254(0.3) + 0.138(1) + 0.265(0) = \boxed{0.2142}$$

2.8. We need the probability of leaving within 3 years minus the probability of leaving within 2 years.

The state vector at the end of one year is  $(0.5 \quad 0.2 \quad 0.2 \quad 0.1)$ . The transition probability vector for two years is

$$(0.5 \quad 0.2 \quad 0.2 \quad 0.1) \begin{pmatrix} 0.5 & 0.2 & 0.2 & 0.1 \\ 0.5 & 0.3 & 0.1 & 0.1 \\ 0 & 0 & 0.8 & 0.2 \\ 0 & 0 & 0 & 1 \end{pmatrix} = (0.35 \quad 0.16 \quad 0.28 \quad 0.21)$$

The probability of leaving within two years is therefore 0.21. The probability of leaving within three years is obtained by summing up the products of the state vector at time 2 with the fourth column of the transition probability matrix:

$$0.35(0.1) + 0.16(0.1) + 0.28(0.2) + 0.21(1) = 0.317$$

The probability of leaving in the third year is the excess of the probability of state #4 at time 3 over the probability of state #4 at time 2, or  $0.317 - 0.21 = \boxed{0.107}$ .

---

<sup>5</sup>provided by David Sidney

**2.9.** This exercise is made easier by the fact that the probability of staying in the THCU for two years is zero; in each year, the probabilities of leaving the THCU add up to 1. In addition you can only reach the THCU by being in the ILU in the previous year. And once in PHCU or leaving, there is no way back. So the only way that a resident in the ILU can be in the THCU at time  $t = 3$  is either to be in the ILU for 2 years and transfer, or by going to the THCU in year 1, back to the ILU in year 2, and then to the THCU. The probability of staying in the ILU is the complement of the probability of leaving, which is  $1 - 0.1 - 0.1 - 0.1 = 0.7$  in the first year and  $1 - 0.2 - 0.2 - 0.2 = 0.4$  in the second. So the probability of ILU→ILU→ILU→THCU is  $(0.7)(0.4)(0.2) = 0.056$ , and the probability of ILU→THCU→ILU→THCU is  $(0.1)(0.5)(0.2) = 0.01$ . The total probability is  $0.056 + 0.01 = 0.066$ . The probability of then transferring to the ILU in year 4 is 0.5. The answer is  $(0.066)(0.5) = \boxed{0.033}$ .

**2.10.** Notice that before considering the risks, whether the club is in state 0 or state 1, it will end up with cash of 1500. In state 0, member charges equals operating expenses and the club starts with cash of 1500, so the club ends up with cash of 1500 before considering risks. In state 1, member charges equals operating expenses plus 1000, so the club ends up with cash of  $500 + 1000 = 1500$  before considering risks. Therefore, a risk moves the club from state 0 to state 2 if and only if it moves the club from state 1 to state 2. The probability of moving from state 0 to state 2 is equal to the probability of moving from state 1 to state 2.

Since the probability of moving from state 0 to state 2 and from state 1 to state 2 are the same, it is not necessary to track the states 1 and 2 separately. Moreover, since once in state 2 the club never moves out, we just need the probability of not going to state 2. We then cube this to get the probability of not going to state 2 for three years.

The probability of no risks is  $(0.9)(0.75^3) = 0.3796875$ , and the probability of exactly one of the first three risks and not the fourth one is  $(0.9)(3)(0.75^2)(0.25) = 0.3796875$ , leaving a probability of  $1 - 2(0.3796875) = 0.240625$  of transferring to state 2. Then  $1 - (1 - 0.240625)^3 = \boxed{0.562106}$ . (E)

**2.11.** The first transition has a probability of 0.2 of increase and 0.8 of decrease (based on the second line of the table). If rates increase, the new state will be Decrease/Increase, or line 4 of the table, and the 2nd year probability of decrease will be 0.75, for a total of  $(0.2)(0.75) = 0.15$ . Otherwise, the state will still be Decrease/Decrease and the 2nd year probability of decrease will be 0.8, for a total of  $(0.8)(0.8) = 0.64$ . The answer is  $0.15 + 0.64 = \boxed{0.79}$ . (B)

**2.12.** Starting in state 1, there is a 0.9 probability of transitioning to state 0 and a 0.1 probability of transitioning to state 2. From those states, the system *must* transition back to state 1. Thus after two years, the prospective present value of perpetuity payments will be *identical* to those at time 0. In other words, if we let  $v_i$  be the discount factors for year  $i$ , then

$$Y = v_1 + v_1v_2 + v_1v_2Y$$

We take the expected values of both sides.  $v_1 = 0.94$ , since we know we're in state 1 for the first year.

$$\begin{aligned} E[Y] &= E[v_1] + E[v_1v_2] + E[v_1v_2Y] \\ E[Y] &= 0.94 + 0.94((0.9)(0.95) + (0.1)(0.93)) + 0.94((0.9)(0.95) + (0.1)(0.93)) E[Y] \\ E[Y] &= 0.94 + 0.89112 + 0.89112 E[Y] \\ E[Y] &= \frac{0.94 + 0.89112}{1 - 0.89112} = \boxed{16.817781} \quad (\text{D}) \end{aligned}$$

**2.13.** There is a 4-month pathway from Current to Foreclosed with probability  $(0.2)(0.5)(0.75)(0.9) = 0.0675$ , and that is the only way to be Foreclosed. The probability of being Current is 1 at time 0, 0.8 at time 1, and  $0.8^2 + 0.2(0.5) = 0.74$  at time 2, since at time 2 a loan can be Current since time 0 or 30 at time 1 and Current at time 2. So the answer is the sum of the three possibilities starting at Current at time 0, 1, 2, or  $0.0675(1 + 0.8 + 0.74) = 0.17145$ . The number of Foreclosed loans is  $100,000(0.17145) = \boxed{17,145}$ . (D)

**2.14.** The system is in state (ii). Let's calculate the probability of no complaints for 2 days, and then the complement will be the probability of at least one claim. The probability of no complaints the first day is 0.4. On the second day, we would then be in state (iii) and the probability of no complaints is 0.55. Thus the probability of 2 days with no complaints is  $(0.4)(0.55) = 0.22$  and the probability of at least one complaint in 2 days is  $1 - 0.22 = \boxed{0.78}$ . (D)

**2.15.** The probability of moving to class 1 within 2 years is  $(0.6)(0.3) + (0.4)(0.8) = 0.5$ , and the probability of moving back to class 0 the next year is 0.1, so this path has a probability of  $(0.5)(0.1) = 0.05$ .

The other possibility is being in class 1 after 3 years and then moving back to class 0. The probability of being in class 0 at the end of 2 years is the complement of being in class 1, which we calculated was 0.5. Then the probability of moving to class 1 in the third year is 0.2, for a product of 0.1. The other possibility is being in class 1 after 2 years, probability 0.5, and staying there another year, probability 0.9, for a probability of 0.45. The total probability of being in class 1 after 3 years is  $0.1 + 0.45 = 0.55$ . Then the probability of transferring to class 0 in the fourth year is 0.1, for a total probability of  $(0.55)(0.1) = 0.055$  for this route.

The answer is  $0.05 + 0.055 = \boxed{0.105}$ .

**2.16.** The state vector at time 1 is  $(0.3 \ 0.5 \ 0.2)$ . The state vector at time 2 is

$$(0.3 \ 0.5 \ 0.2) \begin{pmatrix} 0.3 & 0.5 & 0.2 \\ 0.6 & 0.3 & 0.1 \\ 0 & 0 & 1 \end{pmatrix} = (0.39 \ 0.30 \ 0.31)$$

The probability of state 1 at time 3 is  $(0.39)(0.3) + (0.30)(0.6) = 0.297$ . The situation of being in state 1 is Bernoulli (either you are or you aren't), so the variance for 100 individuals is  $100(0.297)(1 - 0.297) = \boxed{20.8791}$ .

**2.17.** Since Markov chains do not have memory, and this chain is homogeneous, it will suffice to calculate

1. The distribution of states at time 1.
2. The probability that someone in state 2 is in state 1 after two periods.
3. The probability that someone in state 3 is in state 1 after two periods.

After one period, the state vector is  $(0.8 \ 0.1 \ 0.1)$ . The relative probabilities of states 2 and 3 are  $0.1/0.2$  and  $0.1/0.2$ , or  $1/2$  apiece.

For someone in state 2, the state vector after one period is  $(0.7 \ 0.2 \ 0.1)$ . After two periods, the probability of state 1 is  $0.7(0.8) + 0.2(0.7) + 0.1(0) = 0.7$ .

For someone in state 3, the probability of state 1 after two periods is  $0(0.8) + 0.1(0.7) + 0.9(0) = 0.07$ .

The answer is  $0.5(0.7) + 0.5(0.07) = \boxed{0.385}$ .

**2.18.** We calculate the probability of being in state (2) at the end of two years. (Once you enter state (2), you never leave.) After the first year, the state probability vector is  $(0.70 \ 0.20 \ 0.10)$ . After the second year, the third entry of the state probability vector (multiplying the state probability vector from year one by the third column of the matrix) is  $0.7(0.1) + 0.2(0.25) + 0.1(1) = 0.22$ . The number of lives is a binomial random variable (either a life is dead or it isn't) with parameters  $m = 100$ ,  $q = 0.22$ , and therefore variance  $mq(1 - q) = 100(0.22)(0.78) = \boxed{17.16}$ . (C)

**2.19.** The state vector at the beginning of year 2 is  $(0.80 \ 0.15 \ 0.05)$ . Multiplying by the last column of the transition matrix to get the last entry of the beginning of year 3 state vector:

$$0.80(0.05) + 0.15(0.05) + 0.05(1.00) = 0.0975$$

The variance of the number of members, since that is a binomial distribution with parameters  $m = 50$ ,  $q = 0.0975$ , is  $mq(1 - q) = 50(0.0975)(1 - 0.0975) = \boxed{4.400}$ . (C)

**2.20.** Tom is in state 1. The two ways to find at least 3 coins are:

1. Find 1 coin in one block and 2 in the next block. Probabilities are 0.6, followed by (using the second row of the transition matrix) 0.3, or  $(0.6)(0.3) = 0.18$ .
2. Find 2 coins in one block and at least 1 in the next block. Probabilities are 0.3 followed by (using the last row of the transition matrix)  $1 - 0.1$  (anything other than 0 coin), or  $(0.3)(0.9) = 0.27$ .

The total probability is  $0.18 + 0.27 = \boxed{0.45}$ . **(B)**

**2.21.** The state probability vectors are  $(1 \ 0 \ 0)$  initially,  $(0.7 \ 0.1 \ 0.2)$  at the end of one year, and

$$(0.7 \ 0.1 \ 0.2) \begin{pmatrix} 0.7 & 0.1 & 0.2 \\ 0.3 & 0.6 & 0.1 \\ 0.0 & 0.0 & 1.0 \end{pmatrix} = (0.52 \ 0.13 \ 0.35)$$

at the end of 2 years. The expected value of contract premiums is  $800(1 + (0.7 + 0.1) + (0.52 + 0.13)) = 1960$ . The expected value of healthcare costs is  $500 + (500(0.7) + 3000(0.1)) + (500(0.52) + 3000(0.13)) = 1800$ . The net expected value is  $1960 - 1800 = \boxed{160}$ . **(D)**

**2.22.** The state probability vectors are  $(1 \ 0 \ 0)$  initially,  $(0.7 \ 0.1 \ 0.2)$  at the end of one year, and

$$(0.7 \ 0.1 \ 0.2) \begin{pmatrix} 0.7 & 0.1 & 0.2 \\ 0.3 & 0.6 & 0.1 \\ 0.0 & 0.0 & 1.0 \end{pmatrix} = (0.52 \ 0.13 \ 0.35)$$

at the end of two years.

The expected value of transitions is the probability of starting in Healthy, times the probability of transition to Sick (0.1), times 500, or  $500(0.1)(1 + 0.7 + 0.52) = \boxed{111}$ .

**2.23.** Since the coin is fair, the probability of heads  $p = 0.5$ . Using the gambler's ruin formula,  $r = q/p = 0.6/0.4 = 1.5$  and the gambler is left with 8 chips after the first 10 flips.

$$P_8 = \frac{i}{N} = \frac{8}{30}$$

The probability of failure, or losing all chips, is the complement of the probability of success, or  $1 - 8/30 = \boxed{0.7333333}$ . **(A)**

**2.24.** Since all we care about is whether the allosaur attains 15,000 calories, we can treat this as a Markov chain with four states: 0, 5,000 calories, 10,000 calories, and 15,000+ calories, and pretend that once the 15,000+ calorie state is reached, it is never left (even though in reality it can be). The initial state is 10,000 calories, and the question reduces to the probability of reaching the 15,000+ calorie state before reaching the 0 calorie state.

In the first day, the probability of transition to the 15,000+ calorie state is 0.45 and the probability of transition to the 5,000 calorie state is 0.55. If the allosaur moves to the 5,000 calorie state, the probability of transition to the 0 calorie state is 0.55 and the probability of transition to the 10,000 calorie state is 0.45. Once back in the 10,000 calorie state, everything starts anew, since the Markov chain has no memory,

Thus after two periods, the probability of reaching the 15,000 calorie state is 0.45 and the probability of reaching the 0 calorie state is  $(0.55)(0.55) = 0.3025$ . Thus the probability of ever reaching 15,000+ is the relative probability of 15,000+ versus the probability of 0, or  $0.45/(0.45 + 0.3025) = \boxed{0.5980}$ . **(C)**

If you remembered the gambler's ruin formulas, you could apply them directly. One step is 5000 calories. We start in state 2 (10,000 calories) and the goal is state 3 (15,000 calories). The probability of up is  $p = 0.45$ . In our notation,  $r = q/p = 11/9$ . Then

$$P_2 = \frac{r^2 - 1}{r^3 - 1} = \frac{(11/9)^2 - 1}{(11/9)^3 - 1} = \boxed{0.5980}$$

2.25. Using formula (2.2),

$$\begin{aligned}\frac{r^2 - 1}{r^6 - 1} &= 0.5 \\ \frac{1}{r^4 + r^2 + 1} &= 0.5 \\ r^4 + r^2 + 1 &= 2 \\ r^4 + r^2 - 1 &= 0 \\ r^2 &= \frac{-1 + \sqrt{5}}{2} = 0.618034 \\ r &= 0.786151 \\ \frac{1-p}{p} &= 0.786151 \\ p &= \frac{1}{1.786151} = \boxed{0.559863}\end{aligned}$$

2.26. Using formula (2.3):

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} = \boxed{2.2833}$$

2.27.  $E[N_j] = \text{Var}(N_j) = \ln 200 = 5.2983$ . We want  $\Pr(N_j > 10.5)$ , where we added 0.5 for a continuity correction.

$$1 - \Phi\left(\frac{10.5 - 5.2983}{\sqrt{5.2983}}\right) = 1 - \Phi(2.260) = \boxed{0.0119}$$

## Quiz Solutions

2-1. Using formula (2.2) with  $r = 1$ , we get  $\boxed{0.6}$ .



---

---

# Practice Exam 1

---

1. Cars arrive at a toll booth in a Poisson process at the rate of 6 per minute.

Determine the probability that the third car will arrive between 30 and 40 seconds from now.

- A. Less than 0.18
- B. At least 0.18, but less than 0.21
- C. At least 0.21, but less than 0.24
- D. At least 0.24, but less than 0.27
- E. At least 0.27

2. A business receives 50 pieces of mail every day in a Poisson process. One tenth of the mail contains checks. The logarithm of the amount of each check has a normal distribution with parameters  $\mu = 3$ ,  $\sigma^2 = 9$ .

Determine the average number of checks for amounts greater than 10,000 that the business receives in a seven day week.

- A. Less than 0.66
- B. At least 0.66, but less than 0.69
- C. At least 0.69, but less than 0.75
- D. At least 0.75, but less than 0.75
- E. At least 0.75

3. ATM withdrawals occur in a Poisson process at varying rates throughout the day, as follows:

|          |  |
|----------|--|
| 11PM–6AM | 3 per hour   |
| 6AM–8AM  | Linearly increasing from 3 per hour to 30 per hour |
| 8AM–5PM  | 30 per hour  |
| 5PM–11PM | Linearly decreasing from 30 per hour to 3 per hour |

Withdrawal amounts are uniformly distributed on (100, 500), and are independent of each other and the number of withdrawals.

Using the normal approximation, estimate the amount of money needed to be adequate for all withdrawals for a day 95% of the time.

- A. Less than 137,500
- B. At least 137,500, but less than 138,000
- C. At least 138,000, but less than 138,500
- D. At least 138,500, but less than 139,000
- E. At least 139,000

4. In a Poisson process, arrivals occur at the rate of 5 per hour. Exactly one event has occurred within the last 20 minutes, but the time of the event is unknown. Estimate the 90<sup>th</sup> percentile of the time, in minutes, of the event
- A. Less than 16 minutes
  - B. At least 16 minutes, but less than 17 minutes
  - C. At least 17 minutes, but less than 18 minutes
  - D. At least 18 minutes, but less than 19 minutes
  - E. At least 19 minutes
5. The amount of time between windstorms causing losses of 100 million or more is exponentially distributed with a mean of 10 years. The amount of time between wildfires causing losses of 100 million or more is exponentially distributed with a mean of 6 years. Determine the probability that 2 windstorms will occur before 3 wildfires.
- A. Less than 0.2
  - B. At least 0.2, but less than 0.3
  - C. At least 0.3, but less than 0.4
  - D. At least 0.4, but less than 0.5
  - E. At least 0.5
6. For a certain population, lifetime is exponentially distributed with mean 70. Every member of the population earns 100,000 per year from the 20<sup>th</sup> birthday to the 65<sup>th</sup> birthday. Calculate expected lifetime earnings for a newborn.
- A. Less than 2,700,000
  - B. At least 2,700,000, but less than 3,000,000
  - C. At least 3,000,000, but less than 3,300,000
  - D. At least 3,300,000, but less than 3,600,000
  - E. At least 3,600,000
7. An insurance company currently sells 20 million of inland marine insurance. The company has devised a strategy for expanding this line of business, but the strategy is risky. In any year, if the strategy is successful, sales will increase by 10 million. If the strategy is unsuccessful, sales will decrease by 10 million. If sales go down to 0, the company will exit the business. The probability in each year that the strategy is successful is  $\frac{2}{3}$ . The company's goal is to increase sales to 60 million. Calculate the probability that the company reaches its goal.
- A. Less than 0.60
  - B. At least 0.60, but less than 0.65
  - C. At least 0.65, but less than 0.70
  - D. At least 0.70, but less than 0.75
  - E. At least 0.75

8. For a discrete irreducible Markov chain with 3 states:

- The limiting probability of state 1 is 0.6.
- The limiting probability of state 3 is 0.3.
- The probability of transition from state 2 to state 1 is 0.8.
- The probability of transition from state 3 to state 1 is 0.

Calculate the probability of staying in state 1 for one transition.

- A. Less than 0.85
- B. At least 0.85, but less than 0.88
- C. At least 0.88, but less than 0.91
- D. At least 0.91, but less than 0.94
- E. At least 0.94

9. In a continuous Markov chain with 3 states, the only transitions possible are from state 1 to state 2 and from state 1 to state 3. The process transitions out of state 1 at the rate of 0.4 per hour, with 70% of the transitions going to state 2.

The process is currently in state 1.

Calculate the probability that the process is in state 3 at the end of 2 hours.

- A. Less than 0.15
- B. At least 0.15, but less than 0.18
- C. At least 0.18, but less than 0.21
- D. At least 0.21, but less than 0.24
- E. At least 0.24

10. A copier is either working or out of order.

If the copier is working at the beginning of a day, the probability that it is working at the beginning of the next day is 0.8.

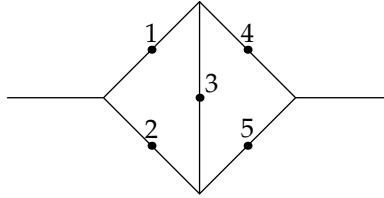
If the copier is out of order at the beginning of a day, the probability that it is fixed during the day and working at the beginning of the next day is 0.6.

You are given that the probability that the copier is working at the beginning of Tuesday is 0.77.

Determine the probability that the copier was working at the beginning of the previous day, Monday.

- A. Less than 0.75
- B. At least 0.75, but less than 0.78
- C. At least 0.78, but less than 0.81
- D. At least 0.81, but less than 0.84
- E. At least 0.84

11. You are given the following system of 5 components:



Determine the number of minimal cut sets in this system.

- A. 2                      B. 3                      C. 4                      D. 5                      E. 6

12. A graph consists of 3 nodes numbered 1, 2, 3 and arcs connecting them. The probability that an arc connects two nodes is 0.8 for nodes 1 and 2, 0.7 for nodes 1 and 3, and 0.6 for nodes 2 and 3.

Calculate the probability that the graph is connected.

- A. Less than 0.65  
 B. At least 0.65, but less than 0.70  
 C. At least 0.70, but less than 0.75  
 D. At least 0.75, but less than 0.80  
 E. At least 0.80

13. You are given:

- The following life table.

| $x$ | $l_x$ | $d_x$ |
|-----|-------|-------|
| 50  | 1000  | 20    |
| 51  |       |       |
| 52  |       | 35    |
| 53  |       | 37    |

- ${}_2q_{52} = 0.07508$ .

Determine  $d_{51}$ .

- A. Less than 20  
 B. At least 20, but less than 22  
 C. At least 22, but less than 24  
 D. At least 24, but less than 26  
 E. At least 26

14. For a 30-year deferred whole life annuity on (35):

- The annuity will pay 100 per year at the beginning of each year, starting at age 65.
- If death occurs during the deferral period, the contract will pay 1000 at the end of the year of death.
- Mortality follows the Illustrative Life Table.
- $i = 0.06$ .
- $Y$  is the present value random variable for the contract.

Calculate  $E[Y]$ .

- A. Less than 204
- B. At least 204, but less than 205
- C. At least 205, but less than 206
- D. At least 206, but less than 207
- E. At least 207

15. For 2 estimators of  $\theta$ ,  $\hat{\theta}$  and  $\tilde{\theta}$ , you are given:

•

|                |                |                  |
|----------------|----------------|------------------|
|                | $\hat{\theta}$ | $\tilde{\theta}$ |
| Expected value | 4              | 5                |
| Variance       | 2              | 3                |

- $\theta = 5$
- $\text{Cov}(\hat{\theta}, \tilde{\theta}) = -1$

Determine the mean square error of  $\frac{1}{2}(\hat{\theta} + \tilde{\theta})$  as an estimator of  $\theta$ .

- A. Less than 1.25
- B. At least 1.25, but less than 1.75
- C. At least 1.75, but less than 2.25
- D. At least 2.25, but less than 2.75
- E. At least 2.75

16. For a set of 3 biased coins, the probability of head is  $p$ . The 3 coins are tossed 10 times, with the following results:

| Number of heads | Number of times |
|-----------------|-----------------|
| 0               | 4               |
| 1               | 3               |
| 2               | 2               |
| 3               | 1               |

Determine the maximum likelihood estimate of  $p$ .

- A. 1/5
- B. 1/4
- C. 1/3
- D. 2/5
- E. 1/2

17. A sample of 6 observed claim sizes is

10    25    30    52    70    90

These observations are fitted to a Lognormal distribution with  $\mu = 2$  using maximum likelihood. Determine the variance of the fitted distribution.

- A. Less than 21,000
- B. At least 21,000, but less than 23,000
- C. At least 23,000, but less than 25,000
- D. At least 25,000, but less than 27,000
- E. At least 27,000

18. For two baseball teams  $A$  and  $B$ :

- Team  $A$  wins 7 out of 10 games.
- Team  $B$  wins  $x$  out of 14 games.
- The null hypothesis is that the two teams are equally likely to win games.
- The alternative hypothesis is that the two teams are not equally likely to win games.

Determine the highest value of  $x$  for which the null hypothesis is accepted at 5% significance.

- A. 10                      B. 11                      C. 12                      D. 13                      E. 14

19. For a Normally distributed variable  $X$  with  $\sigma^2 = 2500$ , you test  $H_0: \mu = 100$  against  $H_1: \mu < 100$  using the sample mean of 30 observations. The test is constructed to have 1% significance.

Determine the power of the test at 70.

- A. Less than 0.72
- B. At least 0.72, but less than 0.76
- C. At least 0.76, but less than 0.80
- D. At least 0.80, but less than 0.84
- E. At least 0.84

20. In Territory 1, you have 130 policies and experience aggregate losses of 100,000, with sample standard deviation 2000.

In Territory 2, you have 80 policies and experience aggregate losses of 20,000, with sample standard deviation 1500.

You test the null hypothesis that underlying average aggregate losses per policy in both territories is equal. You assume that aggregate losses are normally distributed.

Determine the results of the test.

- A. Reject the null hypothesis at 1% significance.
- B. Accept the null hypothesis at 1% significance, but not at 2.5% significance.
- C. Accept the null hypothesis at 2.5% significance, but not at 5% significance.
- D. Accept the null hypothesis at 5% significance, but not at 10% significance.
- E. Accept the null hypothesis at 10% significance.

21. A sample of 20 items from a normal distribution yields the following summary statistics:

$$\begin{aligned}\sum X_i &= 120 \\ \sum X_i^2 &= 1100\end{aligned}$$

Construct a 99% confidence interval of the form  $(0, a)$  for the variance.

Determine  $a$ .

- A. 10.0                      B. 10.1                      C. 10.5                      D. 48.5                      E. 49.8

22.  $X$  is a random variable having probability density function

$$f(x) = \alpha x^{\alpha-1} \quad 0 < x < 1$$

You test  $H_0: \alpha = 1$  against  $H_1: \alpha > 1$  using 2 observations,  $x_1$  and  $x_2$ .

Determine the form of the uniformly most powerful critical region for this test.

- A.  $x_1 + x_2 < k$   
 B.  $x_1 + x_2 > k$   
 C.  $x_1 x_2 < k$   
 D.  $x_1 x_2 > k$   
 E.  $\frac{1}{x_1} + \frac{1}{x_2} < k$

23. You are given a sample of size 4 from a distribution with probability density function

$$f(x) = 2x \quad 0 \leq x \leq 1$$

$Y_1, \dots, Y_4$  are the order statistics.

Determine  $\Pr(Y_2 > 0.5)$ .

- 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

24. You are given the following information from a group of students regarding time spent studying for an exam and the score on the exam:

|                |     |     |     |     |     |
|----------------|-----|-----|-----|-----|-----|
| Time (minutes) | 372 | 405 | 428 | 457 | 500 |
| Score          | 85  | 78  | 82  | 100 | 92  |

Calculate Spearman's  $\rho$  relating study time and exam score.

- A. Less than 0.3  
 B. At least 0.3, but less than 0.4  
 C. At least 0.4, but less than 0.5  
 D. At least 0.5, but less than 0.6  
 E. At least 0.6

25. In a certain town, the natural logarithm of annual wages is hypothesized to be symmetrically distributed with mean 10.5. The wages of six people are

20,000 30,000 50,000 80,000 110,000 200,000

Using the Wilcoxon signed rank test, calculate the  $p$ -value of the hypothesis.

- A. Less than 0.05
- B. At least 0.05, but less than 0.10
- C. At least 0.10, but less than 0.15
- D. At least 0.15, but less than 0.20
- E. At least 0.20

26. You are given:

- Claim counts follow a Poisson distribution. The probability of 0 claims is  $\theta$ .
- The distribution of  $\theta$  over the entire population has density function

$$f(\theta) = 3\theta^2 \quad 0 < \theta < 1$$

A policyholder submits no claims for 2 years.

Calculate the posterior probability that this policyholder submits no claims in the third year.

- A. 0.50                      B. 0.67                      C. 0.75                      D. 0.80                      E. 0.83

27. A small department store chain has 5 stores. At each store, daily sales are normally distributed with variance 5,000,000. Mean daily sales at each store are  $\mu$ , where  $\mu$  is normally distributed with mean 50,000 and variance 10,000,000.

At one of the stores, total sales over a 7 day week are 420,000.

Determine the posterior mean daily sales for this store.

- A. Less than 52,000
- B. At least 52,000, but less than 54,000
- C. At least 54,000, but less than 56,000
- D. At least 56,000, but less than 58,000
- E. At least 58,000

28. The monthly number of losses on an insurance coverage follows a Poisson distribution with mean  $\lambda$ . The prior distribution of  $\lambda$  is gamma with parameters  $\alpha$  and  $\theta$ .

A randomly selected insured is observed for  $n$  months and submits no claims.

Determine the smallest  $n$  such that the expected number of claims for this policyholder is half of the expected number of claims for the general population.

- A.  $\theta$                       B.  $1/\theta$                       C.  $\alpha\theta$                       D.  $\alpha/\theta$                       E.  $\theta/\alpha$

29. Auto liability claim size is modeled using a generalized linear model. Based on an analysis of the data, it is believed that the coefficient of variation of claim size is constant.

Which of the following response distributions would be most appropriate to use?

- A. Poisson                      B. Normal                      C. Gamma                      D. Inverse Gamma                      E. Inverse Gaussian



30. A generalized linear model for automobile insurance with 40 observations has the following explanatory variables:

- SEX (male or female)
- AGE (4 levels)
- TYPE OF VEHICLE (sedan, coupe, SUV, van)
- MILES DRIVEN (continuous variable)
- USE (business, pleasure, farm)

Model I includes all of these variables and an intercept. Model II is the same as Model I except that it excludes USE. You have the following statistics from these models:

|          | Deviance | AIC   |
|----------|----------|-------|
| Model I  | 23.12    | 58.81 |
| Model II |          | 62.61 |

Which of the following statements is correct?

- A. Accept USE at 0.5% significance.
- B. Accept USE at 1.0% significance but not at 0.5% significance.
- C. Accept USE at 2.5% significance but not at 1.0% significance.
- D. Accept USE at 5.0% significance but not at 2.5% significance.
- E. Reject USE at 5.0% significance.

31. You are given the following regression model, based on 22 observations.

$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_4x_4 + \beta_5x_5 + \varepsilon$$

The error sum of squares for this model is 156.

If the variables  $x_4$  and  $x_5$  are removed, the error sum of squares is 310.

Calculate the  $F$  ratio to determine the significance of the variables  $x_4$  and  $x_5$ .

- A. Less than 4.0
- B. At least 4.0, but less than 5.0
- C. At least 5.0, but less than 6.0
- D. At least 6.0, but less than 7.0
- E. At least 7.0

32. Which of the following statements are true regarding goodness-of-fit testing for a logistic regression?

1. The chi-square distribution is a poor approximation for deviance if cell frequencies are too low.
2. The Hosmer-Lemeshow method is a method of calculating the deviance statistic when cell frequencies are low.
3. Pseudo $R^2$  often is alarmingly low even when other measures indicate the model fits well.

- A. 1. only      B. 1. and 2. only      C. 3. only      D. 1. and 3. only      E. 1., 2., and 3.

33. A normal linear model with 2 variables and an intercept is based on 45 observations.  $\hat{y}_j$  is the fitted value of  $y_j$ , and  $\hat{y}_{j(i)}$  is the fitted value of  $y_j$  if observation  $i$  is removed. You are given:

- $\sum_{j=1}^{45} (y_j - y_{j(1)})^2 = 4.1$ .
- The leverage of the first observation is 0.15.

Determine  $|\hat{\epsilon}_1|$ , the absolute value of the first residual of the regression with no observation removed.

- A. Less than 4
- B. At least 4, but less than 5
- C. At least 5, but less than 6
- D. At least 6, but less than 7
- E. At least 7

34. A gamma regression model with  $\alpha = 1$  is fitted to data. The identity link is used.

This regression is equivalent to a weighted least squares regression.

Express the weights, the entries in the diagonal matrix  $\mathbf{W}$ , as a function of  $\mu_i$ , the mean of the response variable.

- A.  $1/\mu_i^2$
- B.  $1/\mu_i$
- C.  $\ln \mu_i$
- D.  $\mu_i$
- E.  $\mu_i^2$

35. You are given the following output from a GLM to estimate loss size:

- Distribution selected is Inverse Gaussian.
- The link is  $g(\mu) = 1/\mu^2$ .

| Parameter           | $\beta$  |
|---------------------|----------|
| Intercept           | 0.00279  |
| Vehicle Body        |          |
| Coupe               | 0.002    |
| Sedan               | -0.001   |
| SUV                 | 0.003    |
| Vehicle Value (000) | -0.00007 |
| Area                |          |
| B                   | -0.025   |
| C                   | 0.015    |
| D                   | 0.005    |

Calculate mean loss size for a sedan with value 25,000 from Area A.

- A. Less than 100
- B. At least 100, but less than 500
- C. At least 500, but less than 1000
- D. At least 1000, but less than 5000
- E. At least 5000

36. You have fit 20 observations to two nested normal models, Model I and Model II. The following ANOVA table summarizes the results:

| Source of variation         | Degrees of freedom | Sum of squares | Mean square |
|-----------------------------|--------------------|----------------|-------------|
| Model I                     | 3                  | 25,123         |             |
| Improvement due to Model II | 1                  | 509            | 509.0       |
| Residual                    | 16                 | 1,240          | 77.5        |
| Total                       | 20                 | 26,872         |             |

You are also given that  $\bar{Y} = 35$ .

Calculate  $R^2$  for Model II.

- A. Less than 0.55
- B. At least 0.55, but less than 0.65
- C. At least 0.65, but less than 0.75
- D. At least 0.75, but less than 0.85
- E. At least 0.85

37. The response variable of a generalized linear model follows a normal distribution. The link is  $g(\mu) = \ln \mu$ .

The method of scoring is used to fit the coefficients. At each iteration, weighted least squares is performed.

Which of the following is proportional to the weights?

- A.  $1/\mu_i^2$
- B.  $1/\mu_i$
- C. 1
- D.  $\mu_i$
- E.  $\mu_i^2$ .

38. A generalized linear model of the form

$$\sqrt{\mu} = \beta_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon$$

is estimated based on 20 observations. The resulting estimate of  $\beta$  is  $b_1 = 1.80$ ,  $b_2 = 3.28$ ,  $b_3 = 3.21$ . You are given that

$$(\mathbf{X}^T \mathbf{W} \mathbf{X})^{-1} = \begin{pmatrix} 0.85 & 0.75 & 0.23 \\ 0.75 & 1.60 & 0.85 \\ 0.23 & 0.85 & 1.32 \end{pmatrix}$$

Based on the Wald statistic, which of the  $\beta$  parameters are significant at the 5% level?

- A.  $\beta_2$  only
- B.  $\beta_3$  only
- C.  $\beta_1$  and  $\beta_2$  only
- D.  $\beta_2$  and  $\beta_3$  only
- E.  $\beta_1$ ,  $\beta_2$ , and  $\beta_3$

39. For an inverse Gaussian regression, you are given

- $y_5 = 652$ .
- $\hat{y}_5 = 530$
- The inverse Gaussian has parameter  $\theta = 1/2$

Calculate the deviance residual of the fifth observation,  $d_5$ .

- A. Less than 0.01
- B. At least 0.01, but less than 0.02
- C. At least 0.02, but less than 0.03
- D. At least 0.03, but less than 0.04
- E. At least 0.04

40. For a generalized linear model,

- There are 72 observations.
- There are 25 parameters.
- The loglikelihood is  $-361.24$

You are considering adding a cubic polynomial variable.

Determine the lowest loglikelihood for which this additional variable would be accepted at 1% significance.

- A. Less than  $-356$
- B. At least  $-356$ , but less than  $-354$
- C. At least  $-354$ , but less than  $-352$
- D. At least  $-352$ , but less than  $-350$
- E. At least  $-350$

41. An insurance company is modeling the probability of a claim using logistic regression. The explanatory variable is vehicle value. Vehicle value is banded, and the value of the variable is 1, 2, 3, 4, 5, or 6, depending on the band. Band 1 is the reference level.

The fitted value of the  $\beta$  corresponding to band 4 is  $-0.695$ .

Let  $O_1$  be the odds of a claim for a policy in band 1, and  $O_4$  the odds of a claim for a policy in band 4.

Determine  $O_4/O_1$ .

- A. Less than 0.35
- B. At least 0.35, but less than 0.40
- C. At least 0.40, but less than 0.45
- D. At least 0.45, but less than 0.50
- E. At least 0.50

42. A monthly time series has seasonal patterns. Seasonality of the series is modeled with an additive model. When the 12-month centered moving average is subtracted from the series, the average result by month is

|           |      |
|-----------|------|
| January   | -5.3 |
| February  | -8.5 |
| March     | -3.2 |
| April     | 1.0  |
| May       | 1.0  |
| June      | 4.4  |
| July      | 2.1  |
| August    | 0.8  |
| September | 0.6  |
| October   | -3.5 |
| November  | -1.1 |
| December  | 6.9  |

For January 2014, the unadjusted value of the time series is 102.8.

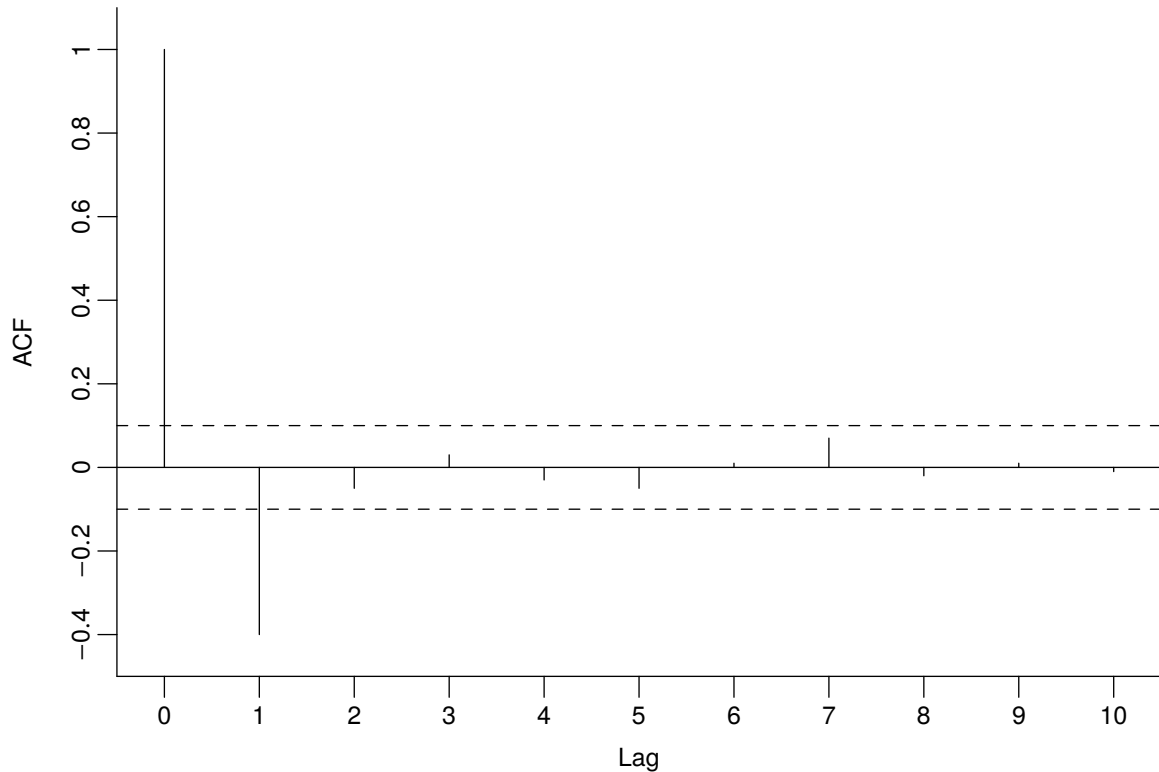
Calculate the seasonally adjusted value.

- A. Less than 99
- B. At least 99, but less than 102
- C. At least 102, but less than 105
- D. At least 105, but less than 108
- E. At least 108

43. For a random walk with variance parameter  $\sigma^2$ , which of the following are true?

- 1. The random walk is stationary in the mean.
  - 2. At time 50, the variance is  $50\sigma^2$ .
  - 3. At time 50, the lag 1 autocorrelation is 0.99.
- A. 1 and 2 only      B. 1 and 3 only      C. 2 and 3 only      D. 1, 2, and 3  
E. The correct answer is not given by A., B., C., or D.

44. The correlogram for a time series is



An invertible MA model is fitted to this time series.

Determine the model.

- A.  $x_t = 0.4w_{t-1} + w_t$
- B.  $x_t = -0.4w_{t-1} + w_t$
- C.  $x_t = 0.5w_{t-1} + w_t$
- D.  $x_t = -0.5w_{t-1} + w_t$
- E.  $x_t = 2w_{t-1} + w_t$

45. Which of the following ARMA models have redundant parameters?

Model I:  $x_t = 1.4x_{t-1} - 0.48x_{t-2} - 0.6w_{t-1} + w_t$

Model II:  $x_t = 1.4x_{t-1} - 0.48x_{t-2} + 0.6w_{t-1} + w_t$

Model III:  $x_t = 1.4x_{t-1} + 0.48x_{t-2} - 0.6w_{t-1} + w_t$

- A. Model I only
- B. Model II only
- C. Model III only
- D. Model I and II only
- E. Model II and III only

*Solutions to the above questions begin on page 835.*

---

---

## Appendix A. Solutions to the Practice Exams

---

### Answer Key for Practice Exam 1

|    |   |    |   |    |   |    |   |    |   |
|----|---|----|---|----|---|----|---|----|---|
| 1  | B | 11 | C | 21 | E | 31 | E | 41 | D |
| 2  | B | 12 | D | 22 | D | 32 | D | 42 | D |
| 3  | B | 13 | B | 23 | D | 33 | B | 43 | D |
| 4  | D | 14 | C | 24 | E | 34 | A | 44 | D |
| 5  | D | 15 | A | 25 | E | 35 | B | 45 | A |
| 6  | A | 16 | C | 26 | E | 36 | A |    |   |
| 7  | E | 17 | D | 27 | E | 37 | E |    |   |
| 8  | B | 18 | D | 28 | B | 38 | D |    |   |
| 9  | B | 19 | D | 29 | C | 39 | A |    |   |
| 10 | E | 20 | C | 30 | C | 40 | B |    |   |

### Practice Exam 1

1. [Lesson 10] The probability that the third car will arrive in the interval (30, 40) is the probability of at least 3 cars in 40 seconds minus the probability of at least 3 cars in 30 seconds. For 40 seconds, the Poisson parameter is 4 and the probability is

$$1 - e^{-4} \left( 1 + 4 + \frac{4^2}{2} \right) = 1 - 0.238103$$

For 30 seconds, the Poisson parameter is 3 and the probability is

$$1 - e^{-3} \left( 1 + 3 + \frac{3^2}{2} \right) = 1 - 0.423190$$

The difference is  $0.423190 - 0.238103 = \boxed{0.185087}$ . (B)

2. [Lesson 12] The probability of a check greater than 10,000 is

$$1 - \Phi \left( \frac{\ln 10,000 - 3}{3} \right) = 1 - \Phi(2.07) = 1 - 0.9808 = 0.0192$$

The Poisson distribution of just the checks over 10,000 in one week has parameter  $7(50)(0.1)(0.0192) = \boxed{0.672}$ . (B)

3. [Lesson 15] The Poisson parameter per day is computed by adding up the rates over the 4 periods. For 11PM–6AM, we have 7 hours times 3 per hour, or 21. For 8AM–5PM we have 9 hours times 30 per hour, or 270. For the other two periods, because of the linear increase or decrease, the average per hour is the midpoint, or  $(30 + 3)/2 = 16.5$ , and there are 8 hours with varying rates, for a total of  $8 \times 16.5 = 132$ .

The total number of withdrawals per day is  $21 + 270 + 132 = 423$ . The mean aggregate withdrawals is  $(423)(300) = 126,900$ .

The second moment of the uniform distribution on  $(100, 500)$  is the variance plus the mean squared. The variance of a uniform distribution is the range squared divided by 12, or  $400^2/12$ . Therefore, the second moment of the uniform distribution is  $400^2/12 + 300^2 = 103,333\frac{1}{3}$ . The variance of aggregate withdrawals, by the compound variance formula (15.2), is  $\lambda E[X^2] = (423)(103,333\frac{1}{3}) = 43,710,000$ .

The amount of money needed to be adequate 95% of the time is

$$126,900 + 1.645\sqrt{43,710,000} = \boxed{137,775.68} \quad (\text{B})$$

4. [Section 13.2] The event time is uniformly distributed on  $[0, 20]$ , so the 90<sup>th</sup> percentile is  $\boxed{18 \text{ minutes}}$  (D)

5. [Section 13.1] Both windstorms and wildfires are Poisson processes. The probability of a windstorm before a wildfire is

$$\frac{\frac{1}{10}}{\frac{1}{6} + \frac{1}{10}} = \frac{\frac{3}{30}}{\frac{5}{30} + \frac{3}{30}} = \frac{3}{8}$$

We are looking for the probability that 2 or more of the next 4 events will be windstorms. The number of events out of 4 that are windstorms is binomial with parameters  $m = 4$  and  $q = \frac{3}{8}$ . The probability of 0 windstorms in the next 4 events is  $(\frac{5}{8})^4$  and the probability of 1 windstorm in the next 4 events is  $\binom{4}{1}(\frac{5}{8})^3(\frac{3}{8})$ , so the probability of 2 or more windstorms in the next 4 is

$$1 - \left(\frac{5}{8}\right)^4 - \binom{4}{1}\left(\frac{5}{8}\right)^3\left(\frac{3}{8}\right) = 1 - 0.1526 - 0.3662 = \boxed{0.4812} \quad (\text{D})$$

6. [Lesson 9] At age 20 and at age 65, expected number of years to death is 70. Thus at age 0, expected number of years starting at age 20 is  $70 \Pr(T > 20) = 70e^{-20/70}$  and expected number of years starting at 65 is  $70e^{-65/70}$ . The difference is the expected number of years from 20 to 65, and we multiply that by 100,000 per year.

$$7,000,000(e^{-20/70} - e^{-65/70}) = \boxed{2,494,517} \quad (\text{A})$$

7. [Section 2.2] Use formula (2.2). Here,  $r = q/p = (1/3)/(2/3) = 1/2$ .

$$P_2 = \frac{(1/2)^2 - 1}{(1/2)^6 - 1} = \boxed{0.7619} \quad (\text{E})$$

8. [Lesson 4] The limiting probability of state 2 is the complement of the limiting probabilities of the other states,  $1 - 0.6 - 0.3 = 0.1$ . The equation for limiting probability of state 1 is

$$\begin{aligned} \pi_1 &= p_{11}\pi_1 + p_{21}\pi_2 + p_{31}\pi_3 \\ 0.6 &= p_{11}(0.6) + (0.8)(0.1) \end{aligned}$$

It follows that  $p_{11} = 0.52/0.6 = \boxed{0.8667}$ . (B)



9. [Lesson 16] The random variable for staying in state 1 is exponential distribution with mean  $1/0.4$ , so the probability of staying in state 1 for 2 hours is  $e^{-2(0.4)}$ . The probability of leaving state 1 is  $1 - e^{-0.8}$ . The probability of going to state 3 in 2 hours is 30% of that, or  $0.3(1 - e^{-0.8}) = \boxed{0.1652}$ . (B)

10. [Lesson 2] Let  $p$  be the probability that the copier was working at the beginning of the previous day. Then

$$\begin{aligned} 0.8p + 0.6(1 - p) &= 0.77 \\ 0.6 + 0.2p &= 0.77 \\ p &= \frac{0.17}{0.2} = \boxed{0.85} \quad (\text{E}) \end{aligned}$$

11. [Lesson 19]  $\{1,2\}$  and  $\{4,5\}$  are minimal cut sets. If a minimal cut set has 1 but not 2, it must have 3 and 5. Similarly, if a minimal cut set has 2 but not 1, it must have 3 and 4. That exhausts all possibilities.  $\boxed{4}$  (C)

12. [Subsection 20.2.1] For the graph to be connected, two of the three arcs must exist. The probability of that is

$$(0.8)(0.7) + (0.7)(0.6) + (0.8)(0.6) - 2(0.8)(0.7)(0.6) = \boxed{0.788} \quad (\text{D})$$

13. [Section 22.1]  $0.07508 = {}_2q_{52} = (d_{52} + d_{53})/l_{52} = 72/l_{52}$ , so  $l_{52} = 72/0.07508 = 959$ . But  $l_{52} = l_{50} - d_{50} - d_{51} = 1000 - 20 - d_{51}$ , so  $d_{51} = \boxed{21}$ . (B)

14. [Section 23.2] Let  $Y_1$  be the value of the benefits if death occurs after age 65 and  $Y_2$  the value of the benefits if death occurs before age 65.

$$\begin{aligned} {}_{30}E_{35} &= {}_{10}E_{35} {}_{20}E_{45} = (0.54318)(0.25634) = 0.139239 \\ \mathbf{E}[Y_1] &= 100 {}_{30}E_{35} \ddot{a}_{65} = 100(0.139239)(9.8969) = 137.803 \\ \mathbf{E}[Y_2] &= 1000 (A_{35} - {}_{30}E_{35} A_{65}) \\ &= 1000(0.12872 - 0.139239(0.43980)) = 67.483 \\ \mathbf{E}[Y] &= 137.803 + 67.483 = \boxed{205.286} \quad (\text{C}) \end{aligned}$$

15. [Lesson 24]  $\mathbf{E}\left[\frac{1}{2}(\hat{\theta} + \tilde{\theta})\right] = \frac{1}{2}(4+5) = 4.5$ , so the bias is  $4.5 - 5 = -0.5$ . The variance of the estimator is

$$\left(\frac{1}{2}\right)^2 (\text{Var}(\hat{\theta} + \tilde{\theta})) = \left(\frac{1}{4}\right) (\text{Var}(\hat{\theta}) + \text{Var}(\tilde{\theta}) + 2 \text{Cov}(\hat{\theta}, \tilde{\theta})) = \left(\frac{1}{4}\right) (2 + 3 + 2(-1)) = 0.75$$

Therefore, the mean square error is  $0.5^2 + 0.75 = \boxed{1}$ . (A)

16. [Lesson 25] For a binomial with fixed  $m = 3$ , maximum likelihood estimates  $q$  the same way as the method of moments. For 30 tosses (10 tosses of 3 coins) we have  $(1)(3) + (2)(2) + (3)(1) = 10$  heads, so  $q = \frac{10}{30} = \boxed{\frac{1}{3}}$ . (C)

17. [Lesson 25] The likelihood function in terms of the 6 observations  $x_i$ , dropping multiplicative constants such as  $\frac{1}{x_i\sqrt{2\pi}}$ , is

$$L(\sigma) = \frac{1}{\sigma^6} e^{-\frac{\sum_{i=1}^6 (\ln x_i - 2)^2}{2\sigma^2}}$$

$$\sum_{i=1}^6 (\ln x_i - 2)^2 = 0.091558 + 1.485658 + 1.963354 + 3.807352 + 5.055731 + 6.249048 = 18.652701$$

$$l(\sigma) = -6 \ln \sigma - \frac{18.652701}{2\sigma^2}$$

$$\frac{dl}{d\sigma} = -\frac{6}{\sigma} + \frac{18.652701}{\sigma^3} = 0$$

$$-6\sigma^2 + 18.652701 = 0$$

$$\sigma^2 = \frac{18.652701}{6} = 3.108784$$

The moments of the fitted distribution are

$$E[X] = e^{2+3.108784/2} = 34.9666$$

$$E[X^2] = e^{4+2(3.108784)} = 27,380$$

$$\text{Var}(X) = 27,380 - 34.9666^2 = \boxed{26,157} \quad (\text{D})$$

18. [Subsection 30.4.3] The number of games won is binomial. The pooled mean games won is  $(7+x)/24$ . For a two-sided test with 5% significance, we need the  $Z$  statistic to be no higher than 1.96, the 97.5<sup>th</sup> percentile of a standard normal distribution. The  $Z$  statistic is

$$Z = \frac{\frac{x}{14} - \frac{7}{10}}{\sqrt{\left(\frac{7+x}{24}\right)\left(\frac{17-x}{24}\right)\left(\frac{1}{10} + \frac{1}{14}\right)}}$$

We set this equal to 1.96 and solve for  $x$ .

$$\frac{x}{14} - 0.7 = \frac{1.96}{24} \sqrt{0.171429(7+x)(17-x)}$$

$$2.112446x - 20.701967 = \sqrt{(7+x)(17-x)}$$

$$4.462426x^2 - 87.463556x + 428.5714 = -x^2 + 10x + 119$$

$$5.462426x^2 - 97.463556x + 309.5714 = 0$$

$$x = 13.71, 4.13$$

Thus we accept the null hypothesis when  $x$  is between 4 and  $\boxed{13}$ . (D)

It may be easier to solve this question by plugging in the answer choices for  $x$  in the original equation setting  $Z$  equal to 1.96.

19. [Lesson 28] To achieve 1% significance, the critical value for a normal random variable must be 2.326 times the standard deviation below the mean, or  $100 - 2.326\left(\frac{50}{\sqrt{30}}\right) = 78.76$ . The power of the test at 70 is the probability of rejecting the null hypothesis if  $\mu = 70$ , or

$$\Pr(X < 70) = \Phi\left(\frac{78.76 - 70}{50/\sqrt{30}}\right) = \Phi(0.960) = \boxed{0.831} \quad (\text{D})$$

20. [Section 30.4] We are testing the difference of means,  $\mu_1 - \mu_2$ . As discussed in Section 30.4, we calculate the pooled variance

$$s^2 = \frac{129(2000^2) + 79(1500^2)}{130 + 80 - 2} = 3,335,337$$

and the standard deviation for the combined sample is then  $\sqrt{3,335,337 \left(\frac{1}{130} + \frac{1}{80}\right)} = 259.52$ . The means are  $100,000/130 = 769.23$  and  $20,000/80 = 250$ . Thus we need

$$1 - \Phi\left(\frac{769.23 - 250}{259.52}\right) = 1 - \Phi(2.00) = 0.0228.$$

Since it is a two-sided test, we double 0.0228 and get 0.0456, so the answer is (C).

21. [Lesson 32] The sample variance is

$$S^2 = \frac{20}{19} \left( \frac{1100}{20} - \left( \frac{120}{20} \right)^2 \right) = 20$$

$\sigma^2 = \frac{19S^2}{W}$ , where  $W$  is chi-square with 19 degrees of freedom. To make  $\sigma^2$  large, make  $W$  small: pick its 1<sup>st</sup> percentile, 7.633. Then  $\sigma^2 = \frac{19(20)}{7.633} = \boxed{49.8}$  is the upper bound of the interval. (E)

22. [Lesson 33] The likelihood ratio is ( $\alpha_0 = 1$ )

$$\frac{(x_1 x_2)^{\alpha_0 - 1}}{\alpha^2 (x_1 x_2)^{\alpha - 1}} = \left( \frac{1}{\alpha^2} \right) (x_1 x_2)^{1 - \alpha}$$

This should be less than a constant  $k$ . The first factor is a positive constant and can be incorporated in  $k$ . Since  $1 - \alpha < 0$ , we will have this expression less than a constant if  $x_1 x_2 > k$ . (D)

23. [Lesson 39] The probability that one item  $X$  is greater than 0.5 is

$$\Pr(X > 0.5) = \int_{0.5}^1 2x \, dx = 1 - 0.5^2 = 0.75$$

The probability that  $Y_2$  is greater than 0.5 is the probability that three or four items are above 0.5, or

$$\begin{aligned} \Pr(Y_2 \geq 0.5) &= \binom{4}{3} (0.75^3)(0.25) + \binom{4}{4} (0.75^4) \\ &= 0.421875 + 0.31640625 = \boxed{0.7383} \quad \text{(D)} \end{aligned}$$

24. [Lesson 42] The ranks of time are in order: 1, 2, 3, 4, 5. The ranks of scores are 3, 1, 2, 5, 4. Using formula (42.5),

$$\rho = 1 - \frac{6((1-3)^2 + (2-1)^2 + (3-2)^2 + (4-5)^2 + (5-4)^2)}{5(24)} = \boxed{0.6} \quad \text{(E)}$$

25. [Section 41.1] The logarithms of the six numbers are 9.9, 10.3, 10.8, 11.3, 11.6, 12.2. After subtracting 10.5, we have  $-0.6, -0.2, 0.3, 0.8, 1.1, 1.7$ . The ranks are 3, 1, 2, 4, 5, 6. The sum of the ranks of the positive numbers is  $2 + 4 + 5 + 6 = 17$ . Since the maximum statistic is  $n(n+1)/2 = (6)(7)/2 = 21$ , the probability  $\Pr(T \geq 17)$  is the same as  $\Pr(T \leq 4)$ . To get a statistic of 4 or less as a sum of unequal numbers, you'd need no numbers, or 1, 2, 3, 4, or 1 + 2, 1 + 3. That's a total of 7 ways out of  $2^6$ , or  $7/64 = 0.109375$ . Since we are performing a two-sided test, the  $p$ -value is  $2(0.109375) = \boxed{0.219}$ . (E)

26. [Lesson 36] Since there are only two possibilities (either 0 claims are submitted or not), the model is Bernoulli. The prior is beta with  $a = 3, b = 1$ , which in the posterior go to  $a' = 3 + 2 = 5, b' = 1 + 0 = 1$ , and the posterior expected value of  $\theta$  is then  $a/(a+b) = 5/6 = \boxed{0.8333}$ , which is the posterior probability of no claims. (E)

27. [Lesson 37] We are given 7 days of experience, so  $n = 7$ . We are given that  $v = 5,000,000$ ,  $a = 10,000,000$ , and  $n\bar{x} = 420,000$ . Using formula (37.1),

$$\mu_* = \frac{5,000,000(50,000) + 420,000(10,000,000)}{5,000,000 + 7(10,000,000)} = \boxed{59,333} \quad (\text{E})$$

28. [Lesson 38] Let  $\gamma = 1/\theta$ . Then after  $n$  months,  $\alpha \rightarrow \alpha$  and  $\gamma \rightarrow \gamma + n$ . We want

$$\frac{\alpha}{\gamma + n} = \frac{1}{2} \frac{\alpha}{\gamma}$$

This means  $n = \gamma$ , or  $n = \boxed{\frac{1}{\theta}}$ . (B)

29. [Section 45.2] The square of the coefficient of variation is the variance divided by the square of the mean. If it is constant, then variance is proportional to mean squared. For the Tweedie distribution family, the gamma distribution has  $p = 2$ , which means that the variance is a constant times the mean squared. (C)

30. [Lesson 48] USE has 3 levels, so Model II has 2 parameters fewer than Model I. Thus the AIC penalty on Model II is 4 less than for Model I. The AIC for Model I is 3.80 greater than for Model II, but before the penalty, twice the loglikelihood of Model I is 7.80 greater than for Model II. The critical values for chi-square with 2 degrees of freedom are 7.38 at 2.5% and 9.21 at 1%, making (C) the correct answer choice.

31. [Section 50.1] There are  $n = 22$  observations,  $p + 1 = 6$  coefficients in the unrestricted model, and  $q = 2$  restrictions.

$$F_{2,16} = \frac{(SSE_R - SSE_{UR})/q}{SSE_{UR}/(N - (p + 1))} = \frac{(310 - 156)/2}{156/16} = \boxed{7.897} \quad (\text{E})$$

32. [Section 54.1] 1. and 3. are lifted from the Dobson & Barnett textbook. The Homer-Lemeshow method is for Pearson chi-square, not deviance. (D)

33. [Section 51.2] Use the second equality of formula (51.2). The standard error of the first residual is  $s\sqrt{1 - h_{11}}$ .

$$\frac{4.1}{3s^2} = \left( \frac{\hat{\varepsilon}_1}{s\sqrt{1 - 0.15}} \right)^2 \left( \frac{0.15}{3(0.85)} \right)$$

$$4.1 = \left(\frac{\hat{\varepsilon}_1^2}{0.85}\right)\left(\frac{0.15}{0.85}\right)$$

$$\hat{\varepsilon}_1^2 = \frac{4.1(0.85^2)}{0.15} = 19.7483$$

$$|\hat{\varepsilon}_1| = \boxed{4.4439} \quad (\text{B})$$

34. [Lesson 47] With the identity link, the weights are  $1/\text{Var}(Y_i)$ . For a gamma distribution with  $\alpha = 1$ ,  $\text{Var}(Y_i) = \mu^2$ , so (A) is the correct answer.

35. [Section 45.2] Area A is the base level, so nothing is added to  $g(\mu)$  for it.

$$g(\mu) = 0.00279 - 0.001 + 25(-0.00007) = 0.00004$$

$$\frac{1}{\mu^2} = 0.00004$$

$$\mu = \sqrt{\frac{1}{0.00004}} = \boxed{158.11} \quad (\text{B})$$

36. [Section 49.1 and Lesson 50] This ANOVA table follows the Dobson & Barnett textbook's convention of including the total sum of squares in the table, not just the sum of square deviation from the mean. You can see this by noticing that the table has 20 degrees of freedom, not 19. The SST in  $R^2$ , however, is the sum of square deviations from the mean. So  $\text{SST} = 26,872 - 20(35^2) = 2,372$ . So  $R^2$  is

$$R^2 = 1 - \frac{1,240}{2,372} = \boxed{0.4772} \quad (\text{A})$$

37. [Lesson 47] The weights are  $((dg/d\mu_i)^2 \text{Var}(\mu_i))^{-1}$ . For the normal distribution,  $\text{Var}(\mu_i) = \sigma^2$ , which is constant. For the link,  $dg/d\mu = 1/\mu$ . So the weights are proportional to  $1/(1/\mu_i)^2 = \mu_i^2$ . (E)

38. [Section 48.2] The variances of the bs are the diagonal of  $(\mathbf{X}^T \mathbf{W} \mathbf{X})^{-1}$ . There is 1 degree of freedom for each  $\beta$ , so the critical value at 5% significance is 3.84. The Wald statistics are:

For  $\beta_0$ :  $1.80^2/0.85 = 3.81$ . Not significant.

For  $\beta_1$ :  $3.28^2/1.60 = 6.72$  Significant.

For  $\beta_2$ :  $3.21^2/1.32 = 7.81$  Significant.

(D)

39. [Section 54.2] We first calculate the likelihood of an observation. For an inverse Gaussian, the parameter for GLM using the parametrization in the distribution tables is  $\mu$ , not  $\theta$ . The table has

$$f(x) = \left(\frac{\theta}{2\pi x^2}\right)^{1/2} \exp\left(-\frac{\theta z^2}{2x}\right), \quad z = \frac{x - \mu}{\mu}$$

We will use the letter  $y$  instead of  $x$ . We can ignore any terms in the inverse Gaussian density not involving  $\mu$ , since we just want the difference between the likelihood in our model and the likelihood in the saturated model. So we ignore  $(\frac{\theta}{2\pi y^3})^{1/2}$ . The log of the rest of the density, replacing  $\mu$  with  $\hat{y}$ , is

$$-\frac{\theta z^2}{2y} = -\frac{\theta}{2} \left(\frac{(y - \hat{y})^2}{y \hat{y}^2}\right)$$

For a saturated model  $\hat{y} = y$  and this is 0, so the excess of the saturated loglikelihood over the loglikelihood of our fit is

$$\frac{\theta}{2} \left( \frac{(y - \hat{y})^2}{y\hat{y}^2} \right)$$

We set  $\theta = 1/2$  and double this expression to get the deviance component for this observation and take the square root, with positive sign since  $y_5 - \hat{y}_5 > 0$ . We get

$$d_5 = \sqrt{\frac{1}{2} \frac{(652 - 530)^2}{652(530^2)}} = \boxed{0.00637} \quad (\text{A})$$

40. [Section 48.2] A cubic polynomial adds 3 parameters. The 99<sup>th</sup> percentile of chi-square at 3 degrees of freedom is 11.34. Twice the difference in loglikelihoods must exceed 11.34, so the loglikelihood must increase by 5.67. Then  $-361.24 + 5.67 = \boxed{-355.57}$ . (B)

41. [Lesson 46] In logistic regression,  $g(\mu)$  is the logarithm of the odds, so we must exponentiate  $\beta$  to obtain odds.

$$e^{-0.695} = \boxed{0.4991} \quad (\text{D})$$

42. [Lesson 55] The sum of the 12 averages is  $-5.3 - 8.5 + \dots + 6.9 = -4.8$ . Divide by 12:  $-4.8/12 = -0.4$ . We add 0.4 to each adjustment so that the adjustments average 0. The adjustment for January is then  $-4.9$ . The seasonally adjusted value is  $102.8 - (-4.9) = \boxed{107.7}$ . (D)

43. [Lesson 57] All three statements are true. The random walk is not stationary, but its mean is always 0, so it is stationary in the means. The variance at time  $t$  is  $t\sigma^2$ , here  $50\sigma^2$ . The lag 1 autocorrelation is  $1/\sqrt{1 + 1/50} = 0.99$ . (D)

44. [Lesson 60] The first autocorrelation is  $-0.4$ , and by equation (60.5) equals  $\beta/(1 + \beta^2)$ .

$$\begin{aligned} \frac{\beta}{1 + \beta^2} &= -0.4 \\ 0.4\beta^2 + \beta + 0.4 &= 0 \\ \beta &= \frac{-1 \pm \sqrt{0.36}}{0.8} = -2, -0.5 \end{aligned}$$

$\beta = -2$  would result in a non-invertible model, so we use  $\beta = -0.5$ . (D)

45. [Lesson 61] In Model I,

$$(1 - 1.4\mathbf{B} + 0.48\mathbf{B}^2)x_t = (1 - 0.6\mathbf{B})w_t$$

The left side factors as  $(1 - 0.6\mathbf{B})(1 - 0.8\mathbf{B})$  so there is a redundant factor  $(1 - 0.6\mathbf{B})$  on both sides.

In Model II, the left side factors the same way but the right side is  $(1 + 0.6\mathbf{B})w_t$  so there is no redundant factor. The left side of Model III doesn't factor. (A)