This manual includes

GOAL Customizable, versatile online exam question bank. Thousands of questions!

Access your exclusive ***StudyPlus***$^{+}$ bonus content:

GOAL | Flashcards | Career Guides

* Key Code Inside *

a/s/m

Exam MAS-I Study Manual

1st Edition, Fourth Printing

Abraham Weishaus, Ph.D., F.S.A., C.F.A., M.A.A.A.

NO RETURN IF OPENED

a/s/m
Actuarial Study Materials
Learning Made Easier

TO OUR READERS:

Please check A.S.M.'s web site at **www.studymanuals.com** for errata and updates. If you have any comments or reports of errata, please e-mail us at **mail@studymanuals.com**.

Contents

Lesson 4

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 $X(t)$.[1] 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 non-homogeneous. We will always assume our chains are homogeneous, unless specified otherwise.[2]

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:

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

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.

Example 4A 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:

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

[1] 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.

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

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

Solution: 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 $0.5625(0.25) = \boxed{\mathbf{0.140625}}$. □

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

Example 4B 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.

Solution: 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:

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

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. □

4.1 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

$$P_{ij}^{t} = \sum_{k=1}^{n} P_{ik}^{u} P_{kj}^{t-u} \tag{4.1}$$

Let $\mathbf{P}$ be the transition probability matrix, and let $\mathbf{P}^{(k)}$ be the k-step transition probability matrix. In other words, $\mathbf{P}_{ij}^{(k)}$ 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,[3] 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 6.

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 4C 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.

[3]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.

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

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

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:

$$0.6(0.2) + 0.22(0.1) + 0.18(0.2) = \boxed{\mathbf{0.178}}$$ □

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 4D 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.

Solution: In the previous example, we calculated the state probability vector at time 2: $\begin{pmatrix}0.6 & 0.22 & 0.18\end{pmatrix}$. Also, $P_{21} = 0.7$. So the answer is $(0.22)(0.7) = \boxed{\mathbf{0.154}}$. □

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 4E 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.

Solution: 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 $\begin{pmatrix}0.6 & 0.4\end{pmatrix}$ by the transition probability matrix, we get

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

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

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

4.2 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 $0, 1, \ldots, N$. 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 $j = i, i-1, \ldots 1$, 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 $j = N$,

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

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

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

and it follows that

$$P_j = \begin{cases} \dfrac{j}{N} & r = 1 \\ \dfrac{r^j - 1}{r^N - 1} & r \neq 1 \end{cases} \tag{4.2}$$

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

Example 4F 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.

Solution: Applying formula (4.2) with $r = q/p = 0.4/0.6 = 2/3$ and $N = 25$,

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

□

Quiz 4-1 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?

4.3 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 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^{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^{th} best solution to the best solution, then

$$\mathbf{E}[N_j] = \sum_{i=1}^{j-1} \frac{1}{i} \tag{4.3}$$

$$\operatorname{Var}(N_j) = \sum_{i=1}^{j-1} \left(\frac{1}{i}\right)\left(1 - \frac{1}{i}\right) \tag{4.4}$$

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

Table 4.1: Summary of Formulas in this Lesson

Chapman-Kolmogorov equations

$$P_{ij}^{t} = \sum_{k=1}^{n} P_{ik}^{u} P_{kj}^{t-u} \tag{4.1}$$

Gambler's ruin probabilities

$$P_j = \begin{cases} \dfrac{j}{N} & r = 1 \\ \dfrac{r^j - 1}{r^N - 1} & r \neq 1 \end{cases} \tag{4.2}$$

where p is the probability of success at each round and $r = q/p$.

Algorithmic efficiency, with N_j = number of steps from j^{th} solution to best solution.

$$\mathbf{E}[N_j] = \sum_{i=1}^{j-1} \frac{1}{i} \tag{4.3}$$

$$\operatorname{Var}(N_j) = \sum_{i=1}^{j-1} \left(\frac{1}{i}\right)\left(1 - \frac{1}{i}\right) \tag{4.4}$$

As $j \to \infty$, $\mathbf{E}[N_j] \to \ln j$ and $\operatorname{Var}(N_j) \to \ln j$.

Exercises

4.1. 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.

4.2. 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.

Exercises continue on the next page . . .

4.3. 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.

4.4. 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 $t+1$.

4.5. **[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

4.6. 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

4.7. In your organization, officers transition between the states (1) Director, (2) Vice President, (3) President, and (4) Fired. Transitions occur at the end of each year in a homogeneous Markov chain with the following transition probability matrix:

$$\begin{pmatrix} 0.8 & 0.2 & 0 & 0 \\ 0.05 & 0.75 & 0.1 & 0.1 \\ 0 & 0.05 & 0.75 & 0.2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Calculate the probability that a person who is Vice President at the beginning of 2009 gets fired no later than the end of 2011.

A. Less than 0.10
B. At least 0.10, but less than 0.15
C. At least 0.15, but less than 0.20
D. At least 0.20, but less than 0.25
E. At least 0.25

Exercises continue on the next page . . .

4.8. The distribution of the number of accidents a driver has in a year depends only on the number of accidents in the previous year. The distribution is as follows:

Number of Accidents in Previous Year	Probability of 0 Accidents	Probability of 1 Accident	Probability of 2 Accidents
0	0.80	0.15	0.05
1	0.60	0.30	0.10
2	0.40	0.40	0.20

A driver has no accidents in the current year.

Calculate the expected number of accidents over the next 3 years.

A. Less than 0.7
B. At least 0.7, but less than 0.8
C. At least 0.8, but less than 0.9
D. At least 0.9, but less than 1.0
E. At least 1.0

Use the following information for questions 4.9 and 4.10:

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.

4.9. Calculate the probability of a currently active life being active at the end of three years.

4.10. Calculate the probability of a currently active life getting disabled and recovering within four years.

4.11. 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:

	To			
From	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.

Exercises continue on the next page . . .

4.12. 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.

4.13. **[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

4.14. **[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 year $t-2$	From year $t-2$ to year $t-1$	Probability that year t will 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

4.15. **[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 $\mathbf{E}[Y]$.

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

4.16. **[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

4.17. **[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

Exercises continue on the next page . . .

4.18. 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.

4.19. 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.

4.20. 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.

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

- The annual transition matrix is given by

$$\begin{array}{c c} & \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \end{array} & \begin{bmatrix} 0.70 & 0.20 & 0.10 \\ 0.10 & 0.65 & 0.25 \\ 0 & 0 & 1 \end{bmatrix} \end{array}$$

- 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

Exercises continue on the next page . . .

4.22. [**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

4.23. [**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

4.24. [**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

Exercises continue on the next page . . .

4.25. 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.

4.26. **[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

4.27. **[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

4.28. 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.

4.29. 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.

4.30. 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.

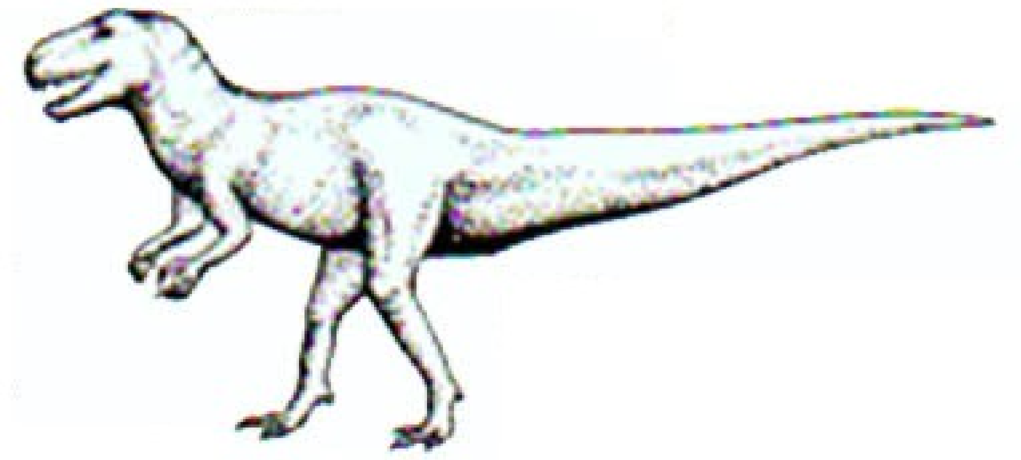


Figure 4.1: Allosaur for exercise 4.27

Additional old CAS Exam MAS-I questions: S19:8,10

Additional old CAS Exam S questions: F15:8, S16:9, F16:14, S17:11,12, F17:11,13

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

Solutions

4.1. 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, $\begin{pmatrix}0.2 & 0.6 & 0.2\end{pmatrix}$. The state vector at time 2 is:

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

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

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

4.2. 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 $(0.42)(0.5^2) = \boxed{\mathbf{0.105}}$.

4.3. 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

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

4.4. 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:

$$\begin{aligned}(x)(0.2)+(1-x)(1-x)&=0.65\\ x^2-1.8x+0.35&=0\\ x=\frac{1.8-\sqrt{1.84}}{2}&=\boxed{\mathbf{0.2218}}\end{aligned}$$

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

4.5. 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

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

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

4.6. Let p be the probability that the copier was working at the beginning of the previous day. Then

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

4.7. At the end of one year, the state vector is $\begin{pmatrix}0.05 & 0.75 & 0.1 & 0.1\end{pmatrix}$. At the end of the second year, it is

$$\begin{pmatrix}0.05 & 0.75 & 0.1 & 0.1\end{pmatrix}\begin{pmatrix}0.8 & 0.2 & 0 & 0\\ 0.05 & 0.75 & 0.1 & 0.1\\ 0 & 0.05 & 0.75 & 0.2\\ 0 & 0 & 0 & 1\end{pmatrix}=\begin{pmatrix}0.0775 & 0.5775 & 0.15 & 0.195\end{pmatrix}$$

The probability of the "fired" state at the end of the third year is

$$0.5775(0.1)+0.15(0.2)+0.195=\boxed{\mathbf{0.28275}}\qquad\textbf{(E)}$$

4.8. The state probability vector at the end of 1 year is $\begin{pmatrix}0.80 & 0.15 & 0.05\end{pmatrix}$. The state probability vector at the end of 2 years is

$$\begin{pmatrix}0.80 & 0.15 & 0.05\end{pmatrix}\begin{pmatrix}0.80 & 0.15 & 0.05\\ 0.60 & 0.30 & 0.10\\ 0.40 & 0.40 & 0.20\end{pmatrix} = \begin{pmatrix}0.75 & 0.185 & 0.065\end{pmatrix}$$

The state probability vector at the end of 3 years is

$$\begin{pmatrix}0.75 & 0.185 & 0.065\end{pmatrix}\begin{pmatrix}0.80 & 0.15 & 0.05\\ 0.60 & 0.30 & 0.10\\ 0.40 & 0.40 & 0.20\end{pmatrix} = \begin{pmatrix}0.737 & 0.194 & 0.069\end{pmatrix}$$

The expected number of accidents in the first year is $0.15(1)+0.05(2) = 0.25$. The expected number of accidents in the second year is $0.185(1)+0.065(2) = 0.315$. The expected number of accidents in the third year is $0.194(1)+0.069(2) = 0.332$. Total expected number of accidents is $0.25 + 0.315 + 0.332 = \boxed{\mathbf{0.897}}$. **(C)**

4.9. Let the states by (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.

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

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

4.10. 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 $0.1176 + 0.0672 + 0.0294 = \boxed{\mathbf{0.2142}}$.

An alternative method[4] 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] provided by David Sidney

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

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

The state vector after three years is

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

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

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

4.11. 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 $\begin{pmatrix} 0.5 & 0.2 & 0.2 & 0.1 \end{pmatrix}$. The transition probability vector for two years is

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

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 $0.317 - 0.21 = \boxed{\mathbf{0.107}}$.

4.12. 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 $(0.066)(0.5) = \boxed{\mathbf{0.033}}$.

4.13. 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 $1 - (1 - 0.240625)^3 = \boxed{\mathbf{0.562106}}$. **(E)**

4.14. 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 $0.15 + 0.64 = \boxed{\mathbf{0.79}}$. **(B)**

4.15. 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

$$Y = v_1 + v_1 v_2 + v_1 v_2 Y$$

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

$$\begin{aligned}
\mathbf{E}[Y] &= \mathbf{E}[v_1] + \mathbf{E}[v_1 v_2] + \mathbf{E}[v_1 v_2 Y]\\
\mathbf{E}[Y] &= 0.94 + 0.94\big((0.9)(0.95) + (0.1)(0.93)\big) + 0.94\big((0.9)(0.95) + (0.1)(0.93)\big)\,\mathbf{E}[Y]\\
\mathbf{E}[Y] &= 0.94 + 0.89112 + 0.89112\,\mathbf{E}[Y]\\
\mathbf{E}[Y] &= \frac{0.94 + 0.89112}{1 - 0.89112} = \boxed{\mathbf{16.817781}} \qquad \textbf{(D)}
\end{aligned}$$

4.16. 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 $100{,}000(0.17145) = \boxed{\mathbf{17{,}145}}$. **(D)**

4.17. 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 $1 - 0.22 = \boxed{\mathbf{0.78}}$. **(D)**

4.18. 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 $0.05 + 0.055 = \boxed{\mathbf{0.105}}$.

4.19. The state vector at time 1 is $\begin{pmatrix}0.3 & 0.5 & 0.2\end{pmatrix}$. The state vector at time 2 is

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

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 $100(0.297)(1 - 0.297) = \boxed{\mathbf{20.8791}}$.

4.20. 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 $\begin{pmatrix}0.8 & 0.1 & 0.1\end{pmatrix}$. 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 $\begin{pmatrix}0.7 & 0.2 & 0.1\end{pmatrix}$. 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 $0.5(0.7) + 0.5(0.07) = \boxed{\mathbf{0.385}}$.

4.21. 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 $\begin{pmatrix}0.70 & 0.20 & 0.10\end{pmatrix}$ 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 $mq(1 - q) = 100(0.22)(0.78) = \boxed{\mathbf{17.16}}$. **(C)**

4.22. The state vector at the beginning of year 2 is $\begin{pmatrix}0.80 & 0.15 & 0.05\end{pmatrix}$. 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 $mq(1 - q) = 50(0.0975)(1 - 0.0975) = \boxed{\mathbf{4.400}}$. **(C)**

4.23. 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 $0.18 + 0.27 = \boxed{\mathbf{0.45}}$. **(B)**

4.24. The state probability vectors are $\begin{pmatrix}1 & 0 & 0\end{pmatrix}$ initially, $\begin{pmatrix}0.7 & 0.1 & 0.2\end{pmatrix}$ at the end of one year, and

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

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

4.25. The state probability vectors are $\begin{pmatrix}1 & 0 & 0\end{pmatrix}$ initially, $\begin{pmatrix}0.7 & 0.1 & 0.2\end{pmatrix}$ at the end of one year, and

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

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 $500(0.1)(1 + 0.7 + 0.52) = \boxed{111}$.

4.26. 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)**

4.27. 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

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

4.28. Using formula (4.2),

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

4.29. Using formula (4.3):

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

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

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

Quiz Solutions

4-1. Using formula (4.2) with $r = 1$, we get $\boxed{\mathbf{0.6}}$.

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^{th} 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 at least 2 windstorms will occur before the third wildfire.

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^{th} birthday to the 65^{th} 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 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:

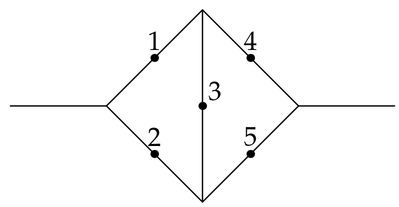
- The limiting probability of state 1 is 0.6.
- The limiting probability of state 3 is 0.3.
- The probability of transition from state 2 to state 1 is 0.8.
- The probability of transition from state 3 to state 1 is 0.

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

Exam questions continue on the next page . . .

9. 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

10. 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

11. 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

Exam questions continue on the next page . . .

12. 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 $\mathbf{E}[Y]$.

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

13. You are given:

- Loss sizes follow a paralogistic distribution with $\alpha = 3$, $\theta = 10$.
- The time of a loss follows a distribution with density function

$$f(t) = 2t \qquad 0 < t < 1$$

- Time of loss is independent of loss size.
- The interest rate is 0.06.
- Z, the present value of one loss, is simulated.
- Loss size is simulated using the random number 0.3 drawn from a uniform distribution on $[0, 1)$.
- Time of loss is simulated using the random number 0.6 drawn from a uniform distribution on $[0, 1)$.

Calculate the simulated value of Z.

A. Less than 4.75
B. At least 4.75, but less than 4.85
C. At least 4.85, but less than 4.95
D. At least 4.95, but less than 5.05
E. At least 5.05

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

•

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

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

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

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

Exam questions continue on the next page . . .

15. The observations 4, 8, 18, 21, 49 are fitted to a distribution with density

$$f(x;\theta,d)=\frac{1}{\theta}e^{-(x-d)/\theta} \qquad x\geq d$$

by matching the first and second moments.

Determine the median of the fitted distribution.

A. Less than 13
B. At least 13, but less than 14
C. At least 14, but less than 15
D. At least 15, but less than 16
E. At least 16

16. 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

17. For an insurance coverage with policy limit 100, there are five observed losses of sizes 30, 50, 60, 70, and 100. In addition, there are three losses for amounts above 100.

Loss sizes are fitted to a Pareto distribution with parameters $\theta = 50$ and α.

Calculate the maximum likelihood estimate of α.

A. 0.55 B. 0.58 C. 0.61 D. 0.66 E. 0.69

18. From a mortality study, you have five observations of time to death: 2, 3, 5, 8, 10.

Survival time is estimated using these observations with kernel-density smoothing. A rectangular kernel with bandwidth 2 is used.

Determine the 30$^{\text{th}}$ percentile of the kernel-density smoothed distribution.

A. $3\frac{1}{6}$ B. $3\frac{1}{5}$ C. $3\frac{1}{4}$ D. $3\frac{1}{3}$ E. $3\frac{1}{2}$

19. 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

Exam questions continue on the next page . . .

20. 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

21. 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.

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

$$\sum X_i = 120$$
$$\sum X_i^2 = 1100$$

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

23. X is a random variable having probability density function

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

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. $\frac{1}{x_1} + \frac{1}{x_2} < k$

Exam questions continue on the next page . . .

24. A random variable follows a two-parameter Pareto distribution with $\alpha = 1$ and $\theta = 5$.

Let Y be the minimum of a sample of 10 drawn from this random variable.

Calculate $\mathbf{E}[Y]$.

A. 5/12 B. 5/11 C. 5/10 D. 5/9 E. 5/8

25. 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

26. 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

Using the likelihood ratio test, 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.

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

$$\mathbf{y} = \beta_0 + \beta_1\mathbf{x}_1 + \beta_2\mathbf{x}_2 + \beta_3\mathbf{x}_3 + \beta_4\mathbf{x}_4 + \beta_5\mathbf{x}_5 + \varepsilon$$

The error sum of squares for this model is 156.

If the variables $\mathbf{x}_4$ and $\mathbf{x}_5$ are removed, the error sum of squares is 310.

Calculate the F ratio to determine the significance of the variables $\mathbf{x}_4$ and $\mathbf{x}_5$.

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

Exam questions continue on the next page . . .

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

I. The chi-square distribution is a poor approximation for deviance if cell frequencies are too low.

II. The Hosmer-Lemeshow method is a method of calculating the deviance statistic when cell frequencies are low.

III. PseudoR^2 often is alarmingly low even when other measures indicate the model fits well.

A. I only B. I and II only C. III only D. I and III only E. I, II, and III

29. 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 $|\hat{\varepsilon}_1|$, 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

30. 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 **W**, as a function of μ_i, the mean of the response variable.

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

Exam questions continue on the next page . . .

31. 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	β
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

32. 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. μ_i E. μ_i^2.

33. A generalized linear model of the form

$$\sqrt{\boldsymbol{\mu}} = \beta_1 + \beta_2 \mathbf{x}_2 + \beta_3 \mathbf{x}_3 + \varepsilon$$

is estimated based on 20 observations. The resulting estimate of $\boldsymbol{\beta}$ 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 β parameters are significant at the 5% level?

A. β_2 only B. β_3 only C. β_1 and β_2 only D. β_2 and β_3 only E. β_1, β_2, and β_3

Exam questions continue on the next page . . .

34. 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

35. 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

36. 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 β 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

37. Consider the vector $\{5, -3, 8, -2, 4\}$.

Calculate the absolute difference between the ℓ_2 norm and ℓ_1 norm of this vector.

A. Less than 12
B. At least 12, but less than 15
C. At least 15, but less than 18
D. At least 18, but less than 21
E. At least 21

Exam questions continue on the next page . . .

38. Which of the following statements are true?

I. Partial Least Squares is a supervised method of dimension reduction.

II. Partial Least Squares directions are linear combinations of the original variables.

III. Partial Least Squares can be used for feature selection.

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

39. A least squares model with a large number of predictors is fitted to 90 observations. To reduce the number of predictors, forward stepwise selection is performed.

For a model with k predictors, RSS = c_k.

The estimated variance of the error of the fit is $\hat{\sigma}^2 = 40$.

Determine the value of $c_d - c_{d+1}$ for which you would be indifferent between the $d+1$-predictor model and the d-predictor model based on Mallow's C_p.

A. Less than 30
B. At least 30, but less than 45
C. At least 45, but less than 60
D. At least 60, but less than 75
E. At least 75

40. 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

Exam questions continue on the next page . . .

41. For a random walk with variance parameter σ^2, which of the following are true?

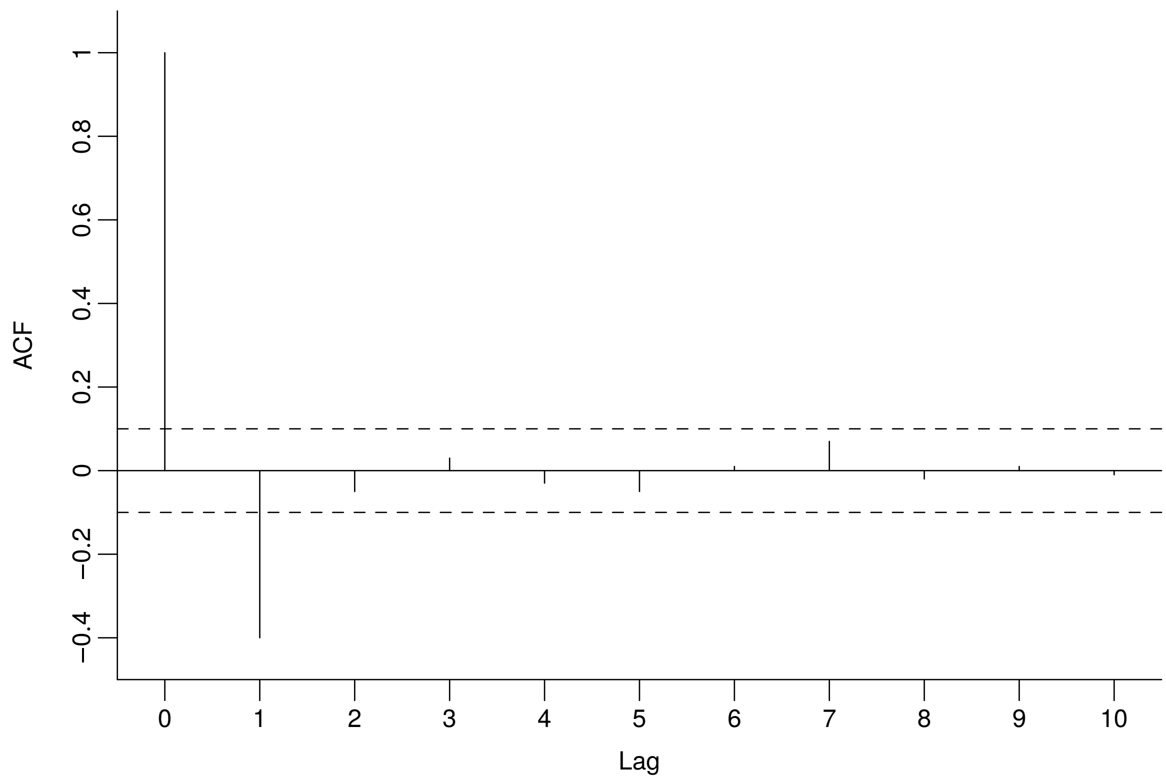
I. The random walk is stationary in the mean.

II. At time 50, the variance is $50\sigma^2$.

III. At time 50, the lag 1 autocorrelation is 0.99.

A. I only
B. II only
C. III only
D. I, II, and III
E. The correct answer is not given by A. , B. , C. , or D.

42. 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. $x_t = 2w_{t-1} + w_t$

43. 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

44. A time series $\{x_t\}$ can be expressed as

$$x_t = \alpha_0 + \alpha_1 t + w_t$$

where w_t is Gaussian white noise.

Determine the type of process followed by ∇x_t, the differences of x_t.

A. White noise
B. Random walk
C. AR(1)
D. MA(1)
E. ARMA(1,1)

45. `R` provides the following estimate for the coefficients of an MA(3) time series:

```
  ma1    ma2    ma3  intercept
0.842  0.621  0.200       -3.5
```

You are given that the residuals for periods 18, 19, and 20 are 6, −4, and 10 respectively.

Forecast the value of the time series in period 21.

A. Less than 3.7
B. At least 3.7, but less than 3.8
C. At least 3.8, but less than 3.9
D. At least 3.9, but less than 4.0
E. At least 4.0

Solutions to the above questions begin on page 985.

Appendix A. Solutions to the Practice Exams

Answer Key for Practice Exam 1

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

Practice Exam 1

1. [**Lesson 11**] 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 $0.423190-0.238103=\boxed{\mathbf{0.185087}}$. **(B)**

2. [**Lesson 13**] 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 $7(50)(0.1)(0.0192)=\boxed{\mathbf{0.672}}$. **(B)**

3. [**Lesson 16**] 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 (16.2), is $\lambda\,\mathbf{E}[X^2]=(423)(103{,}333\frac{1}{3})=43{,}710{,}000$.

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

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

4. **[Section 14.2]** The event time is uniformly distributed on [0,20], so the 90th percentile is $\boxed{\textbf{18 minutes}}$ **(D)**

5. **[Section 14.1]** 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 $\left(\frac{5}{8}\right)^4$ and the probability of 1 windstorm in the next 4 events is $\binom{4}{1}\left(\frac{5}{8}\right)^3\left(\frac{3}{8}\right)$, so the probability of 2 or more windstorms in the next 4 is

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

6. **[Lesson 10]** 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.

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

7. **[Section 4.2]** Use formula (4.2). Here, $r = q/p = (1/3)/(2/3) = 1/2$.

$$P_2 = \frac{(1/2)^2 - 1}{(1/2)^6 - 1} = \boxed{\mathbf{0.7619}} \qquad \textbf{(E)}$$

8. **[Lesson 6]** 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

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

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

9. **[Lesson 17]** {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. $\boxed{\mathbf{4}}$ **(C)**

10. **[Subsection 18.2.1]** For the graph to be connected, two of the three arcs must exist. The probability of that is

$$(0.8)(0.7) + (0.7)(0.6) + (0.8)(0.6) - 2(0.8)(0.7)(0.6) = \boxed{\mathbf{0.788}} \qquad \textbf{(D)}$$

11. **[Section 20.1]** $0.07508 = {}_2q_{52} = (d_{52} + d_{53})/l_{52} = 72/l_{52}$, so $l_{52} = 72/0.07508 = 959$. But $l_{52} = l_{50} - d_{50} - d_{51} = 1000 - 20 - d_{51}$, so $d_{51} = \boxed{\mathbf{21}}$. **(B)**

12. **[Section 21.2]** 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.

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

13. **[Lesson 23]** For loss size,

$$\begin{aligned}
F(x) &= 1 - \left(\frac{1}{1+(x/\theta)^\alpha}\right)^\alpha\\
&= 1 - \left(\frac{1}{1+(x/10)^3}\right)^3\\
0.3 &= 1 - \left(\frac{1}{1+(x/10)^3}\right)^3\\
0.7 &= \left(\frac{1}{1+(x/10)^3}\right)^3\\
\frac{1}{\sqrt[3]{0.7}} - 1 &= \left(\frac{x}{10}\right)^3\\
0.126248 &= \left(\frac{x}{10}\right)^3\\
x &= 10\sqrt[3]{0.126248} = 5.016583
\end{aligned}$$

For time of loss, $F(t) = \int_0^t 2w\,dw = t^2$, so $0.6 = t^2$, $t = \sqrt{0.6} = 0.7746$. The simulated discounted value of the loss is $\frac{5.016583}{1.06^{0.7746}} = \boxed{\mathbf{4.79519}}$. **(B)**

14. **[Lesson 25]** $\mathbf{E}\left[\frac{1}{2}(\hat\theta + \tilde\theta)\right] = \frac{1}{2}(4+5) = 4.5$, so the bias is $4.5 - 5 = -0.5$. The variance of the estimator is

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

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

15. **[Lesson 27]** This is an exponential shifted by d. The mean is $\theta + d$ and the variance is θ^2, since shifting doesn't affect the variance. The observed mean is

$$\bar{x} = \frac{4+8+18+21+49}{5} = 20$$

and the observed variance is

$$\mu_2 = \frac{4^2+8^2+18^2+21^2+49^2}{5} - 20^2 = 249.2$$

Equating the moments,

$$\theta + d = 20$$

$$\begin{aligned}\theta^2 &= 249.2\\ \theta &= 15.78607\\ d &= 4.21393\end{aligned}$$

(Notice that the method of moments estimate is implausible, since one of the observations is lower than 4.21393, yet the fitted distribution sets the probability of being less than d to 0.)

The median is x such that

$$\begin{aligned}F(x) &= 1 - e^{-(x-d)/\theta} = 0.5\\ x &= \theta \ln 2 + d\\ &= 15.78607 \ln 2 + 4.21393 = \boxed{\mathbf{15.156}} \qquad \textbf{(D)}\end{aligned}$$

16. **[Lesson 29]** The likelihood function in terms of the 6 observations x_i, dropping multiplicative constants such as $\frac{1}{x_i\sqrt{2\pi}}$, is

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

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

$$\begin{aligned}l(\sigma) &= -6\ln\sigma - \frac{18.652701}{2\sigma^2}\\ \frac{\mathrm{d}l}{\mathrm{d}\sigma} &= -\frac{6}{\sigma} + \frac{18.652701}{\sigma^3} = 0\\ -6\sigma^2 &+ 18.652701 = 0\\ \sigma^2 &= \frac{18.652701}{6} = 3.108784\end{aligned}$$

The moments of the fitted distribution are

$$\begin{aligned}\mathbf{E}[X] &= e^{2+3.108784/2} = 34.9666\\ \mathbf{E}[X^2] &= e^{4+2(3.108784)} = 27{,}380\\ \mathrm{Var}(X) &= 27{,}380 - 34.9666^2 = \boxed{\mathbf{26{,}157}} \qquad \textbf{(D)}\end{aligned}$$

17. **[Lesson 29]** The important point is that the observed claim of 100 is not censored, and therefore its likelihood is its density function rather than its distribution function.

For the five observed claims x_i, the likelihood is proportional to

$$f(x_i) \sim \frac{\alpha 50^\alpha}{(50 + x_i)^\alpha}$$

where as usual the $+1$ in the exponent of the denominator $(\alpha + 1)$ can be dropped since doing so multiplies by a constant. For the three claims above 100, the likelihood is

$$S(x_i) = \frac{50^\alpha}{150^\alpha}$$

We compute $(150^3)\prod_{i=1}^{5}(50 + x_i) = 5.346 \times 10^{16}$. Then

$$L(\alpha) = \frac{\alpha^5 50^{8\alpha}}{(5.346 \times 10^{16})^\alpha}$$

$$\begin{aligned}
l(\alpha) &= 5\ln\alpha + \alpha\left(8\ln 50 - \ln 5.346\times 10^{16}\right)\\
\frac{\mathrm{d}l}{\mathrm{d}\alpha} &= \frac{5}{\alpha} + 8(3.9120) - 38.5177 = 0\\
\alpha &= \frac{5}{38.5177 - 8(3.9120)} = \boxed{\mathbf{0.6924}} \qquad \textbf{(E)}
\end{aligned}$$

18. **[Lesson 26]** We need $F(x) = 0.3$, where $F(x)$ is the kernel-smoothed distribution. We know that in the empirical distribution, $F(x)$ increases by 0.2 for each point, so by the time we get to 5 (past the span for 3), $F(x) \geq 0.4$. Let's try $x = 3$. We have kernel weights of 0.5 from 3 and 0.75 from 2, making $F(3) = (0.5 + 0.75)/5 = 0.25$. From 3 to 4, $F(x)$ will increase at the rate of $3(0.25)/5 = 0.15$, since 2, 3, and 5 each contribute at a rate of 0.25 (since the bandwidth is 2, so the kernel density is $1/(2b) = 0.25$). Thus we need $0.25 + 0.15x = 0.3$, or $x = 1/3$. Thus $F(3\frac{1}{3}) = 0.3$ and the answer is $\boxed{\mathbf{3\frac{1}{3}}}$. **(D)**

19. **[Subsection 35.4.3]** 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$^{\text{th}}$ 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.

$$\begin{aligned}
\frac{x}{14} - 0.7 &= \frac{1.96}{24}\sqrt{0.171429(7 + x)(17 - x)}\\
2.112446x - 20.701967 &= \sqrt{(7 + x)(17 - x)}\\
4.462426x^2 - 87.463556x + 428.5714 &= -x^2 + 10x + 119\\
5.462426x^2 - 97.463556x + 309.5714 &= 0\\
x &= 13.71, 4.13
\end{aligned}$$

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

20. **[Lesson 33]** 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

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

21. **[Section 35.4]** We are testing the difference of means, $\mu_1 - \mu_2$. As discussed in Section 35.4, 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)**.

22. [**Lesson 38**] 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 σ^2 large, make W small: pick its 1st percentile, 7.633. Then $\sigma^2 = \frac{19(20)}{7.633} = \boxed{49.8}$ is the upper bound of the interval. **(E)**

23. [**Lesson 39**] The likelihood ratio is ($\alpha_0 = 1$)

$$\frac{(x_1x_2)^{\alpha_0-1}}{\alpha^2(x_1x_2)^{\alpha-1}} = \left(\frac{1}{\alpha^2}\right)(x_1x_2)^{1-\alpha}$$

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 $x_1x_2 > k$. **(D)**

24. [**Lesson 41**] The survival function for Y is

$$\Pr(Y > x) = \Pr(X_1,\ldots,X_{10} > x) = \left(\frac{5}{5+x}\right)^{10}$$

so Y follows a Pareto with $\theta = 5$, $\alpha = 10$, and has mean $\boxed{5/9}$. **(D)**

25. [**Section 44.2**] 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)**

26. [**Lesson 47**] 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 less than for Model II, but before the penalty, twice the negative loglikelihood of Model I is 7.80 less 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.

27. [**Section 49.1**] There are $n = 22$ observations, $p = 6$ coefficients in the unrestricted model, and $q = 2$ restrictions. By formula (49.3),

$$F_{2,16} = \frac{(\mathrm{SSE}_R - \mathrm{SSE}_{UR})/q}{\mathrm{SSE}_{UR}/(n-p))} = \frac{(310-156)/2}{156/16} = \boxed{7.897} \qquad \textbf{(E)}$$

28. [**Section 53.1**] I and III are lifted from the *An Introduction to Generalized Linear Models* textbook. The Hosmer-Lemeshow method is for Pearson chi-square, not deviance. **(D)**

29. [**Section 50.2**] Use the second equality of formula (50.3). The standard error of the first residual is $s\sqrt{1-h_{11}}$.

$$\begin{aligned}
\frac{4.1}{3s^2} &= \left(\frac{\hat\varepsilon_1}{s\sqrt{1-0.15}}\right)^2\left(\frac{0.15}{3(0.85)}\right)\\
4.1 &= \left(\frac{\hat\varepsilon_1^2}{0.85}\right)\left(\frac{0.15}{0.85}\right)\\
\hat\varepsilon_1^2 &= \frac{4.1(0.85^2)}{0.15} = 19.7483\\
|\hat\varepsilon_1| &= \boxed{4.4439} \qquad \textbf{(B)}
\end{aligned}$$

30. **[Lesson 46]** With the identity link, the weights are $1/\mathrm{Var}(Y_i)$. For a gamma distribution with $\alpha = 1$, $\mathrm{Var}(Y_i) = \mu^2$, so **(A)** is the correct answer.

31. **[Section 44.2]** Area A is the base level, so nothing is added to $g(\mu)$ for it.

$$\begin{aligned} g(\mu) &= 0.00279 - 0.001 + 25(-0.00007) = 0.00004 \\ \frac{1}{\mu^2} &= 0.00004 \\ \mu &= \sqrt{\frac{1}{0.00004}} = \boxed{\mathbf{158.11}} \qquad \textbf{(B)} \end{aligned}$$

32. **[Lesson 46]** The weights are $\left((\mathrm{d}g/\mathrm{d}\mu_i)^2\,\mathrm{Var}(\mu_i)\right)^{-1}$. For the normal distribution, $\mathrm{Var}(\mu_i) = \sigma^2$, which is constant. For the link, $\mathrm{d}g/\mathrm{d}\mu = 1/\mu$. So the weights are proportional to $1/(1/\mu_i)^2 = \boxed{\boldsymbol{\mu_i^2}}$. **(E)**

33. **[Section 47.2]** The variances of the **b**s are the diagonal of $(\mathbf{X}^T\mathbf{W}\mathbf{X})^{-1}$. There is 1 degree of freedom for each β, so the critical value at 5% significance is 3.84. The Wald statistics are:
For β_0: $1.80^2/0.85 = 3.81$. Not significant.
For β_1: $3.28^2/1.60 = 6.72$ Significant.
For β_2: $3.21^2/1.32 = 7.81$ Significant.
(D)

34. **[Section 53.2]** We first calculate the likelihood of an observation. For an inverse Gaussian, the parameter for GLM using the parametrization in the distribution tables is μ, not θ. The table has

$$f(x) = \left(\frac{\theta}{2\pi x^3}\right)^{1/2} \exp\left(-\frac{\theta z^2}{2x}\right), \quad z = \frac{x-\mu}{\mu}$$

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

$$d_5 = \sqrt{\frac{1}{2}\,\frac{(652-530)^2}{652(530^2)}} = \boxed{\mathbf{0.00637}} \qquad \textbf{(A)}$$

35. **[Section 47.2]** A cubic polynomial adds 3 parameters. The 99th 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 $-361.24 + 5.67 = \boxed{\mathbf{-355.57}}$. **(B)**

36. **[Lesson 45]** In logistic regression, $g(\mu)$ is the logarithm of the odds, so we must exponentiate β to obtain odds.

$$e^{-0.695} = \boxed{\mathbf{0.4991}} \qquad \textbf{(D)}$$

37. **[Section 56.1]** Let v be the vector.

$$\|v\|_1 = 5+3+8+2+4 = 22$$
$$\|v\|_2 = \sqrt{5^2+3^2+8^2+2^2+4^2} = 10.8628$$

The absolute difference is $|22 - 10.8628| = \boxed{\mathbf{11.1372}}$. **(A)**

38. **[Section 56.2]**

1. PLS is a supervised method since it takes the response into account when determining the coefficients. ✓
2. In both dimension reduction methods we study, the selected directions are linear combinations of the original variables. ✓
3. PLS creates new variables that are functions of the original ones, so it does not select features. ✗

(D)

39. **[Section 55.2]** $C_p = \frac{1}{n}(\text{RSS} + 2d\hat{\sigma}^2)$, and we can ignore $1/n$. So we want

$$c_d + 2d(40) = c_{d+1} + 2(d+1)(40)$$

This implies

$$c_d - c_{d+1} = 2(40) = \boxed{\mathbf{80}} \qquad \textbf{(E)}$$

40. **[Lesson 58]** The sum of the 12 averages is $-5.3 - 8.5 + \cdots + 6.9 = -4.8$. 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 $102.8 - (-4.9) = \boxed{\mathbf{107.7}}$. **(D)**

41. **[Lesson 60]** All three statements are true. The random walk is not stationary, but its mean is always 0, so it is stationary in the mean. 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)**

42. **[Lesson 63]** The first autocorrelation is -0.4, and by equation (63.5) equals $\beta/(1+\beta^2)$.

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

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

43. **[Lesson 64]** 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)**

44. **[Lesson 62]** The difference process is $\alpha_1 + w_t - w_{t-1}$, an $\boxed{\textbf{MA(1)}}$ process. **(D)**

45. **[Lesson 63]**

$$-3.5 + 0.842(10) + 0.621(-4) + 0.200(6) = \boxed{\mathbf{3.636}} \qquad \textbf{(A)}$$

9 781635 887884

ASM Study Manual
for CAS Exam MAS-I