Exam MAS-I

Expertise. Insight. Solutions.

CASUALTY ACTUARIAL SOCIETY
AND THE
CANADIAN INSTITUTE OF ACTUARIES

Canadian Institute of Actuaries / Institut canadien des actuaires

MAS-I

Modern Actuarial Statistics I

April 23, 2019

4 HOURS

INSTRUCTIONS TO CANDIDATES

1. This 90 point examination consists of 45 multiple choice questions each worth 2 points.

2. To answer the multiple choice questions, use the short-answer card provided and a number 2 or HB pencil only.

   - Fill in that it is Spring 2019 and that the exam name is MAS-I.
   - Darken the spaces corresponding to your Candidate ID number. Five rows are available. If your Candidate ID number is fewer than 5 digits, include leading zeros. For example, if your Candidate ID number is 987, consider that your Candidate ID number is 00987, enter a zero on the first row, a zero on the second row, 9 on the third row, 8 on the fourth row, and 7 on the fifth [last] row. Write in your Candidate ID number next to the place where you darken the spaces for your Candidate ID number. Your name, or any other identifying mark, must not appear on the short-answer card.
   - Mark your short-answer card during the examination period. No additional time will be allowed for this after the exam has ended. Make your marks dark and fill in the spaces completely.
   - For each of the multiple choice questions, select the one best answer and fill in the corresponding letter. One quarter of the point value of the question will be subtracted for each incorrect answer. No points will be added or subtracted for responses left blank.

3. Do all problems until you reach the last page of the examination where "END OF EXAMINATION" is marked.

CONTINUE TO NEXT PAGE OF INSTRUCTIONS

4. Prior to the start of the exam you will have a **fifteen-minute reading period** in which you can silently read the questions and check the exam booklet for missing or defective pages. Writing will NOT be permitted during this time and you will not be permitted to hold pens or pencils. You will also not be allowed to use calculators. The supervisor has additional exams for those candidates who have defective exam booklets.

   - Verify that you have a copy of "Tables for CAS MAS-I" included in your exam packet.

   - Candidates should not interpolate in the tables unless explicitly instructed to do so in the problem, rather, a candidate should round the result that would be used to enter a given table to the same level of precision as shown in the table and use the result in the table that is nearest to that indicated by rounded result. For example, if a candidate is using the Tables of the Normal Distribution to find a significance level and has a Z value of 1.903, the candidate should round to 1.90 to find cumulative probability in the Normal table.

5. Your Examination Envelope is pre-labeled with your Candidate ID number, name, exam number, and test center. Do not remove this label. Keep a record of your Candidate ID number for future inquiries regarding this exam.

6. Candidates must remain in the examination center until two hours after the start of the examination. The examination starts after the reading period is complete. You may leave the examination room to use the restroom with permission from the supervisor. To avoid excessive noise during the end of the examination, candidates may not leave the exam room during the last fifteen minutes of the examination.

7. At the end of the examination, place the short-answer card in the Examination Envelope. Nothing written in the examination booklet will be graded. Only the short-answer card will be graded. Also place any included reference materials in the Examination Envelope. BEFORE YOU TURN THE EXAMINATION ENVELOPE IN TO THE SUPERVISOR, BE SURE TO SIGN IT IN THE SPACE PROVIDED ABOVE THE CUT-OUT WINDOW.

8. If you have brought a self-addressed, stamped envelope, you may put the examination booklet and scrap paper inside and submit it separately to the supervisor. It will be mailed to you. Do not put the self-addressed stamped envelope inside the Examination Envelope. Interoffice mail is not acceptable.

   If you do not have a self-addressed, stamped envelope, please place the examination booklet in the Examination Envelope and seal the envelope. You may not take it with you. Do not put scrap paper in the Examination Envelope. The supervisor will collect your scrap paper.

   Candidates may obtain a copy of the examination from the CAS Web Site.

   All extra answer sheets, scrap paper, etc. must be returned to the supervisor for disposal.

CONTINUE TO NEXT PAGE OF INSTRUCTIONS

9. Candidates must not give or receive assistance of any kind during the examination. Any cheating, any attempt to cheat, assisting others to cheat, or participating therein, or other improper conduct will result in the Casualty Actuarial Society and the Canadian Institute of Actuaries disqualifying the candidate's paper, and such other disciplinary action as may be deemed appropriate within the guidelines of the CAS Policy on Examination Discipline.

10. The exam survey is available on the CAS Web Site in the "Admissions/Exams" section. Please submit your survey by May 7, 2019.

END OF INSTRUCTIONS

1.

Cars arrive according to a Poisson process at a rate of two per hour.

Calculate the probability that in a given hour, exactly one car will arrive during the first ten minutes and no other cars will arrive during that hour.

A. Less than 0.03
B. At least 0.03 but less than 0.06
C. At least 0.06 but less than 0.09
D. At least 0.09 but less than 0.12
E. At least 0.12

CONTINUED ON NEXT PAGE

2.

You are given the following information about the waiting time until a certain number of events occur:

- The underlying events follow a homogenous Poisson process
- T_n is the time until the n[th] event occurs
- $E[T_2] = 2$

Calculate the variance of T_{10}.

A. Less than 4
B. At least 4, but less than 8
C. At least 8, but less than 12
D. At least 12, but less than 16
E. At least 16

CONTINUED ON NEXT PAGE

3.

You are given the following information:

- Accidents follow a compound Poisson process
- Accidents occur at the rate of $\lambda = 40$ per day
- Accident severity follows an exponential distribution with $\theta = 1{,}000$
- The insurance payment for each accident is subjected to a deductible
- V_1 is the variance of daily aggregate payments with a deductible of 100 per accident
- V_2 is the variance of daily aggregate payments with a deductible of 500 per accident

Calculate the ratio V_2 / V_1.

A. Less than 0.5
B. At least 0.5, but less than 0.6
C. At least 0.6, but less than 0.7
D. At least 0.7, but less than 0.8
E. At least 0.8

CONTINUED ON NEXT PAGE

4.

You are given the following information:

- Computer lifetimes are exponentially distributed with mean of 10 months.
- Computer A has been functioning properly for 12 months.

Calculate the probability that Computer A will function properly for at least 4 more months.

A. Less than 0.64
B. At least 0.64, but less than 0.66
C. At least 0.66, but less than 0.68
D. At least 0.68, but less than 0.70
E. At least 0.70

CONTINUED ON NEXT PAGE

5.

You are given the following information about random variable X:

- The hazard rate function is:

$$r(x) = \begin{cases} \dfrac{k^2}{3x}, & \text{for } x \geq 2 \\ 0, & \text{otherwise} \end{cases}$$

- The value of the cumulative distribution function at $x = 5$ is:

$$F(5) = 0.936.$$

Calculate the absolute value of k.

A. Less than 1.5
B. At least 1.5, but less than 2.5
C. At least 2.5, but less than 3.5
D. At least 3.5, but less than 4.5
E. At least 4.5

CONTINUED ON NEXT PAGE

6.

You are given the following information:

- The severity of each theft loss in a Homeowners insurance policy independently follows an exponential distribution with mean 2,000.
- The insurance company only pays the amount exceeding the per-loss deductible.
- σ_j is the standard deviation of the amount the insurance company pays per theft with a deductible of j.

Calculate the absolute difference between σ_{1000} and σ_{500}.

A. Less than 90
B. At least 90, but less than 100
C. At least 100, but less than 110
D. At least 110, but less than 120
E. At least 120

CONTINUED ON NEXT PAGE

7.

You are given the following information:

- In a toolbox there are two types of components that all perform the same function:
    - There are 4 components of type A, each with reliability of 0.600
    - There are 20 components of type B, each with reliability of 0.300
    - All components are independent
- Using only the components in this toolbox, you want to construct a parallel system with a reliability of at least 0.995

Calculate the minimum number of components needed to create this system.

A. Fewer than 3
B. At least 3, but fewer than 5
C. At least 5, but fewer than 7
D. At least 7, but fewer than 9
E. At least 9

CONTINUED ON NEXT PAGE

8.

On any given day, Alan will play either one or two games of billiards.

- If he plays one game today, he'll play one game tomorrow with probability 0.6
- If he plays two games today, he'll play two games tomorrow with probability 0.7
- His probability of winning each game is 0.6

It is now Monday and Alan played two games today.

Calculate Alan's expected number of wins two days from now on Wednesday.

A. Less than 0.90
B. At least 0.90, but less than 0.95
C. At least 0.95, but less than 1.00
D. At least 1.00, but less than 1.05
E. At least 1.05

CONTINUED ON NEXT PAGE

9.

You are given the following information about a 40-state Markov chain:

- The states are numbered: $\{0,1,\ldots,39\}$
- $P_{m,n}$ is the one-step transition probability from State m to State n
- $P_{0,0} = \frac{1}{2}$
- $P_{0,i} = \frac{1}{78}$ for $i = 1,2,\ldots 39$
- $P_{i,0} = \frac{1}{39}$ for $i = 1,2,\ldots 39$
- $P_{i,i+1} = \frac{38}{39}$ for $i = 1,2,\ldots 38$
- $P_{39,1} = \frac{38}{39}$

Calculate the long run probability of being in State 0.

A. Less than 0.1
B. At least 0.1, but less than 0.2
C. At least 0.2, but less than 0.3
D. At least 0.3, but less than 0.4
E. At least 0.4

CONTINUED ON NEXT PAGE

10.

You are given the following information about two gamblers:

- The two gamblers start with a total fortune of 100 tokens between them
- They plan to go to two different and independent casinos to try their luck
- Each gambler has a 50% chance of winning 1 token and a 50% chance of losing 1 token at each play
- All plays are mutually independent
- The gamblers initially divide their 100 combined tokens to maximize the probability that both of their fortunes will eventually reach 100 tokens

Calculate the product of their initial fortunes, in tokens.

A. Less than 1200
B. At least 1200, but less than 1600
C. At least 1600, but less than 2000
D. At least 2000, but less than 2400
E. At least 2400

CONTINUED ON NEXT PAGE

11.

You are given the following information:

- An annuity-due is issued to a 30-year-old that will pay 1 each year until either she dies or reaches age 50, whichever comes first.
- Mortality follows the Illustrative Life Table.
- Annual interest rate $i = 6\%$.

Calculate the actuarial present value of this annuity.

A. Less than 10.0
B. At least 10.0, but less than 10.5
C. At least 10.5, but less than 11.0
D. At least 11.0, but less than 11.5
E. At least 11.5

CONTINUED ON NEXT PAGE

12.

You are given the following information for a whole life annuity-due of 10,000 on (x) that is payable annually:

- $q_x = 0.010$
- $q_{x+1} = 0.012$
- Interest rate $i = 0.06$
- $\ddot{a}_{x+1} = 10.905$

Calculate the absolute change in the actuarial present value of this annuity due if p_{x+1} is decreased by 0.01.

A. Less than 900
B. At least 900, but less than 925
C. At least 925, but less than 950
D. At least 950, but less than 975
E. At least 975

CONTINUED ON NEXT PAGE

13.

You use the inversion method to simulate two random numbers X from a Weibull (θ, τ) distribution using two independent draws U from a uniform distribution on (0,1). The results of this simulation are shown below:

i	X_i	U_i
1	416.28	0.5
2	357.36	0.4

Calculate the value of θ used to perform this simulation.

A. Less than 425
B. At least 425, but less than 475
C. At least 475, but less than 525
D. At least 525, but less than 575
E. At least 575

14.

You are given the following simulation process to generate random variable X using the rejection method:

- X has a beta distribution, $f(x)$, with parameters $a = 2, b = 10, \theta = 1$, and is concentrated on the interval (0, 1).
- The rejection method uses $g(x) = 1$, for $0 < x < 1$.
- The rejection procedure is as follows:
  - Step 1: Generate independent random numbers Y and U, which are both uniform on (0,1).
  - Step 2: If $U \leq \frac{f(Y)}{cg(Y)}$ stop and set $X = Y$. Otherwise return to Step 1.

Calculate the minimum possible value for c.

A. Less than 3.0
B. At least 3.0, but less than 3.5
C. At least 3.5, but less than 4.0
D. At least 4.0, but less than 4.5
E. At least 4.5

CONTINUED ON NEXT PAGE

15.

You are given the following information:

- $x_1, x_2, \ldots, x_{500}$ is a random sample of 500 claims
- The distribution function is $F(x) = 1 - \left(\frac{200}{x}\right)^{\alpha}$; for $x > 200$ and $\alpha > 0$
- $\sum_{i=1}^{500} x_i = 117{,}378{,}845$ and $\sum_{i=1}^{500} \ln(x_i) = 3{,}646$
- The parameter α is estimated using the method of maximum likelihood

Calculate the estimated probability that a claim will exceed 400.

A. Less than 0.2
B. At least 0.2, but less than 0.4
C. At least 0.4, but less than 0.6
D. At least 0.6, but less than 0.8
E. At least 0.8

16.

Suppose that $X_1, \ldots, X_{10}$ is a random sample from the standard normal distribution, N(0, 1). An actuary is using $\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$ to estimate the population variance, where $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

Calculate the bias of this estimator.

A. Less than -0.2
B. At least -0.2, but less than 0.0
C. At least 0.0, but less than 0.2
D. At least 0.2, but less than 0.4
E. At least 0.4

CONTINUED ON NEXT PAGE

17.

You are given the following information:

- τ is a parameter of a distribution
- The true value of τ is 2
- $\hat{\alpha}$ and $\hat{\beta}$ are two uncorrelated estimators for τ
- $\mathrm{E}(\hat{\alpha}) = 2.0$
- $\mathrm{Var}(\hat{\alpha}) = 5.0$
- $\mathrm{E}(\hat{\beta}) = 3.0$
- $\mathrm{Var}(\hat{\beta}) = 1.0$

Consider the class of estimators of τ which are of the form: $w\hat{\alpha} + (1 - w)\hat{\beta}$.

Calculate the value of w that results in an estimator with the smallest mean squared error.

A. Less than 0.2
B. At least 0.2, but less than 0.4
C. At least 0.4, but less than 0.6
D. At least 0.6, but less than 0.8
E. At least 0.8

CONTINUED ON NEXT PAGE

18.

You are given a random sample from a uniform distribution $[0, \theta]$:

$$1.3,\ \ 2.7,\ \ 8.2,\ \ 5.3,\ \ X$$

You find that the estimate of θ using method of moments is equal to the estimate of θ using the maximum likelihood.

Calculate the largest possible value of X.

A. Less than 4
B. At least 4, but less than 6
C. At least 6, but less than 8
D. At least 8, but less than 10
E. At least 10

CONTINUED ON NEXT PAGE

19.

We observe a random sample, $\{X_1, X_2, \ldots X_n\}$ from a continuous uniform distribution $U(0, \beta)$.

$\hat{\beta} = \max(X_1, X_2, \ldots X_n)$ is chosen as an estimator of β.

I. $\hat{\beta}$ is both the maximum likelihood estimator and method of moments estimator of β
II. $\hat{\beta}$ is sufficient statistic for β
III. The minimum variance unbiased estimator of β is a function of $\hat{\beta}$

Determine which of the above statements are true.

A. None are true
B. I and II only
C. I and III only
D. II and III only
E. The answer is not given by (A), (B), (C), or (D)

CONTINUED ON NEXT PAGE

20.

A sample of losses has the following observations:

$$\{400, 600, 900, 1100, 1200\}$$

A rectangular kernel with bandwidth 300 is used to estimate the probability density function (pdf) of the losses at $x = 800$. The resulting estimate is $\tilde{f}_R(800)$.

A triangular kernel with bandwidth 300 is used to estimate the pdf of the losses at $x = 800$. The resulting estimate is $\tilde{f}_T(800)$.

The absolute difference in the estimates is denoted z, where $z = |\tilde{f}_R(800) - \tilde{f}_T(800)|$.

Calculate z.

A. Less than 0.0001
B. At least 0.0001, but less than 0.0002
C. At least 0.0002, but less than 0.0003
D. At least 0.0003, but less than 0.0004
E. At least 0.0004

CONTINUED ON NEXT PAGE

21.

$X_1, X_2, \dots, X_{100}$ is a random sample from a Bernoulli distribution with probability of success q.

You perform the hypothesis test:

$$H_0: q = 0.5 \quad \text{vs} \quad H_1: q > 0.5$$

The critical region is chosen to be $C = \{\sum_{i=1}^{100} x_i \geq 60\}$.

Calculate the probability of a Type I error.

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

CONTINUED ON NEXT PAGE

22.

Let X be a Bernoulli random variable with probability of success q. You want to perform the following hypothesis test:

- $H_0: q = 0.75$
- $H_1: q > 0.75$

We will reject H_0 in favor of H_1 if $\sum_{i=1}^{30} X_i \geq k$.

Calculate the smallest k such that significance level of this test will be at most 0.02.

A. Fewer than 27
B. 27
C. 28
D. 29
E. At least 30

CONTINUED ON NEXT PAGE

23.

Two independent random samples of sizes 10 and 12 are taken from two separate normal distributions with unknown variances σ_1^2 and σ_2^2, respectively.

- S_1^2 and S_2^2 are the unbiased sample variances
- $S_1^2 = 4$
- You want to perform the following hypothesis test:
  - $H_0: 2\sigma_1^2 = \sigma_2^2$
  - $H_1: 2\sigma_1^2 < \sigma_2^2$
- The p-value of this test is 1.0%

Calculate S_2^2.

A. Less than 10
B. At least 10, but less than 20
C. At least 20, but less than 30
D. At least 30, but less than 40
E. At least 40

24.

You are given the following information:

- Random variable X has an exponential distribution
- $\text{Var}(X) = \frac{1}{9}$

Calculate the median of the distribution.

A. Less than 0.15
B. At least 0.15, but less than 0.20
C. At least 0.20, but less than 0.25
D. At least 0.25, but less than 0.30
E. At least 0.30

CONTINUED ON NEXT PAGE

25.

A type of machine has a lifetime following an exponential distribution with mean μ. All machine lifetimes are independent.

To estimate the expected lifetime of this machine, the factory measures lifetimes of 4 sample machines and takes the average excluding the lowest and highest values. Denote this estimate $\hat{\mu}$.

Calculate the ratio $E(\hat{\mu})/\mu$.

A. Less than 0.8
B. At least 0.8, but less than 1.0
C. At least 1.0, but less than 1.2
D. At least 1.2, but less than 1.4
E. At least 1.4

CONTINUED ON NEXT PAGE

26.

For a general liability policy, loss amounts, Y, follow the Weibull distribution with probability density function:

$$f(y) = e^{-\left(\frac{y}{\theta}\right)^2}\left(\frac{2}{y}\right)\left(\frac{y}{\theta}\right)^2, \qquad y > 0$$

For reinsurance purposes we are interested in the distribution of the largest loss amount in a random sample of size 10, denoted by $Y_{(10)}$.

The 90^{th} percentile of $Y_{(10)}$ is equal to $k\theta$ for some constant k.

Calculate k.

A. Less than 1.2
B. At least 1.2, but less than 1.6
C. At least 1.6, but less than 2.0
D. At least 2.0, but less than 2.4
E. At least 2.4

CONTINUED ON NEXT PAGE

27.

A statistician uses a logistic model to predict the probability of success, π, of a binomial random variable.

You are given the following information:

- There is one predictor variable, X, and an intercept in the model
- The estimates of π at $x = 4$ and 6 are 0.88877 and 0.96562, respectively

Calculate the estimated intercept coefficient, b_0, in the logistic model that produced the above probability estimates.

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

CONTINUED ON NEXT PAGE

28.

Determine which one of the following statements about ridge regression is false.

A. As the tuning parameter $\lambda \rightarrow \infty$, the coefficients tend to zero.

B. The ridge regression coefficients can be calculated by determining the coefficients $\hat{\beta}_1^R, \hat{\beta}_2^R, \ldots, \hat{\beta}_p^R$ that minimize: $\sum_{i=1}^{n}\left(y_i - \beta_0 - \sum_{j=1}^{n} \beta_j x_{ij}\right)^2 + \lambda \sum_{i=1}^{p} \beta_j^2$

C. Unlike standard least squares coefficients, ridge regression coefficients are not scale equivariant.

D. The shrinkage penalty is applied to all the coefficient estimates except for the intercept.

E. Ridge regression shrinks the coefficient estimates, which has the benefit of reducing the bias.

CONTINUED ON NEXT PAGE

29.

Tim uses an ordinary least squares regression model to predict salary based on Experience and Gender. Gender is a qualitative variable and is coded as follows:

$$\text{Gender} = \begin{cases} 1 & \text{if Male} \\ 0 & \text{if Female} \end{cases}$$

His analysis results in the following output:

Coefficients	Estimate	Std. Error	t-value	Pr(>|t|)
Intercept	18169.300	212.2080	85.62027	2.05E-14
Experience	1110.233	59.8224	18.55881	1.75E-08
Gender	169.550	162.9177	10.38285	2.62E-06

Abby uses the same data set but codes the Gender as follows:

$$\text{Gender} = \begin{cases} 1 & \text{if Female} \\ 0 & \text{if Male} \end{cases}$$

Calculate the value of the Intercept in Abby's model.

A. At most 18,169.3
B. Greater than 18,169.3, but at most 18,400.0
C. Greater than 18,400.0, but at most 18,600.0
D. Greater than 18,600.0
E. The answer cannot be computed from the information given

CONTINUED ON NEXT PAGE

30.

You are given the following information about a linear model:

- $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$

Observed Y's	Estimated Y's
2.441	1.827
3.627	3.816
5.126	5.806
7.266	7.796
10.570	9.785

- Residual Sum of Squares = 1.772

Calculate the R^2 of this model.

A. Less than 0.6
B. At least 0.6, but less than 0.7
C. At least 0.7, but less than 0.8
D. At least 0.8, but less than 0.9
E. At least 0.9

CONTINUED ON NEXT PAGE

31.

You are fitting a linear regression model of the form:

$$y = X\beta + e; \quad e_i \sim N(0, \sigma^2)$$

and are given the following values used in this model:

$$X = \begin{bmatrix} 1 & 0 & 1 & 9 \\ 1 & 1 & 1 & 15 \\ 1 & 1 & 1 & 8 \\ 0 & 1 & 1 & 7 \\ 0 & 1 & 1 & 6 \\ 0 & 0 & 1 & 6 \end{bmatrix}; \quad y = \begin{bmatrix} 21 \\ 32 \\ 19 \\ 17 \\ 15 \\ 15 \end{bmatrix}; \quad X^T X = \begin{bmatrix} 3 & 2 & 3 & 32 \\ 2 & 4 & 4 & 36 \\ 3 & 4 & 6 & 51 \\ 32 & 36 & 51 & 491 \end{bmatrix}$$

$$(X^T X)^{-1} = \begin{bmatrix} 1.38 & 0.25 & 0.54 & -0.16 \\ 0.25 & 0.84 & -0.20 & -0.06 \\ 0.54 & -0.20 & 1.75 & -0.20 \\ -0.16 & -0.06 & -0.20 & 0.04 \end{bmatrix}$$

$$H = X(X^T X)^{-1} X^T = \begin{bmatrix} 0.684 & 0.070 & 0.247 & -0.171 & -0.146 & 0.316 \\ 0.070 & 0.975 & -0.044 & 0.108 & -0.038 & -0.070 \\ 0.247 & -0.044 & 0.797 & 0.063 & 0.184 & -0.247 \\ -0.171 & 0.108 & 0.063 & 0.418 & 0.411 & 0.171 \\ -0.146 & -0.038 & 0.184 & 0.411 & 0.443 & 0.146 \\ 0.316 & -0.070 & -0.247 & 0.171 & 0.146 & 0.684 \end{bmatrix}$$

$$(X^T X)^{-1} X^T y = \begin{bmatrix} 0.297 \\ -0.032 \\ 3.943 \\ 1.854 \end{bmatrix}; \quad X(X^T X)^{-1} X^T y = \begin{bmatrix} 20.93 \\ 32.03 \\ 19.04 \\ 16.89 \\ 15.04 \\ 15.07 \end{bmatrix}; \quad \sigma^2 = 0.012657$$

Calculate how many observations are influential, using a unity threshold for Cook's distance.

A. 0
B. 1
C. 2
D. 3
E. 4

CONTINUED ON NEXT PAGE

32.

You are fitting the following linear regression model with an intercept:

$$y = X\beta + e; \quad e_i \sim N(0, \sigma^2)$$

and are given the following values used in the model:

$$X = \begin{bmatrix} 1 & 0 & 1 & 9 \\ 1 & 1 & 1 & 15 \\ 1 & 1 & 1 & 8 \\ 0 & 1 & 1 & 7 \\ 0 & 1 & 1 & 6 \\ 0 & 0 & 1 & 6 \end{bmatrix}; \quad y = \begin{bmatrix} 21 \\ 32 \\ 19 \\ 17 \\ 15 \\ 15 \end{bmatrix}; \quad X^TX = \begin{bmatrix} 3 & 2 & 3 & 32 \\ 2 & 4 & 4 & 36 \\ 3 & 4 & 6 & 51 \\ 32 & 36 & 51 & 491 \end{bmatrix}$$

$$(X^TX)^{-1} = \begin{bmatrix} 1.38 & 0.25 & 0.54 & -0.16 \\ 0.25 & 0.84 & -0.20 & -0.06 \\ 0.54 & -0.20 & 1.75 & -0.20 \\ -0.16 & -0.06 & -0.20 & 0.04 \end{bmatrix}$$

$$H = X(X^TX)^{-1}X^T = \begin{bmatrix} 0.684 & 0.070 & 0.247 & -0.171 & -0.146 & 0.316 \\ 0.070 & 0.975 & -0.044 & 0.108 & -0.038 & -0.070 \\ 0.247 & -0.044 & 0.797 & 0.063 & 0.184 & -0.247 \\ -0.171 & 0.108 & 0.063 & 0.418 & 0.411 & 0.171 \\ -0.146 & -0.038 & 0.184 & 0.411 & 0.443 & 0.146 \\ 0.316 & -0.070 & -0.247 & 0.171 & 0.146 & 0.684 \end{bmatrix}$$

$$(X^TX)^{-1}X^Ty = \begin{bmatrix} 0.297 \\ -0.032 \\ 3.943 \\ 1.854 \end{bmatrix}; \quad X(X^TX)^{-1}X^Ty = \begin{bmatrix} 20.93 \\ 32.03 \\ 19.04 \\ 16.89 \\ 15.04 \\ 15.07 \end{bmatrix}; \quad \sigma^2 = 0.012657$$

Calculate the modeled estimate of the intercept parameter.

A. Less than 0
B. At least 0, but less than 1
C. At least 1, but less than 2
D. At least 2, but less than 3
E. At least 3

CONTINUED ON NEXT PAGE

33.

You are given the following information regarding a GLM that was used to predict claim severity of Homeowners insurance losses:

- The available predictors are:
  - Number of Stories of home ("1" or "2+")
  - Square Footage of the home ("Under 1500", "1501 to 2500", or "Over 2500")
- The partial ANOVA results, which are available to evaluate whether the interaction term "Number of Stories x Square Footage" is predictive, are given below:

Response variable	Insured Loss	
Response distribution	Normal	
Link	Identity	
Parameters	Degrees of freedom	Sum of squares
Number of Stories, Square Footage, and intercept	4	1,128.19
Improvement from Interaction: Number of Stories x Square Footage	2	1.25
Residuals	6	

- The scaled deviance of the model without the interaction terms is 2.64

Calculate the F-statistic that is used to evaluate whether the interaction term is significant.

A. Less than 2.00
B. At least 2.00, but less than 2.25
C. At least 2.25, but less than 2.50
D. At least 2.50, but less than 2.75
E. At least 2.75

CONTINUED ON NEXT PAGE

34.

A statistician has a dataset with n = 50 observations and p = 22 independent predictors. He is using 10-fold cross validation to select from a variety of available models.

Calculate the number of times that the first observation will be included in the training dataset as part of this procedure.

A. 0
B. At least 1, but less than 5
C. At least 5, but less than 10
D. At least 10, but less than 20
E. At least 20

CONTINUED ON NEXT PAGE

35.

You are given the following information about a linear model:

- $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$
- Three bootstrap samples were drawn from the data

Bootstrap Sample	Estimate for β_0
1	10.525
2	2.499
3	16.456

Calculate the standard error of these bootstrap estimates of β_0.

A. Less than 6
B. At least 6, but less than 8
C. At least 8, but less than 10
D. At least 10, but less than 12
E. At least 12

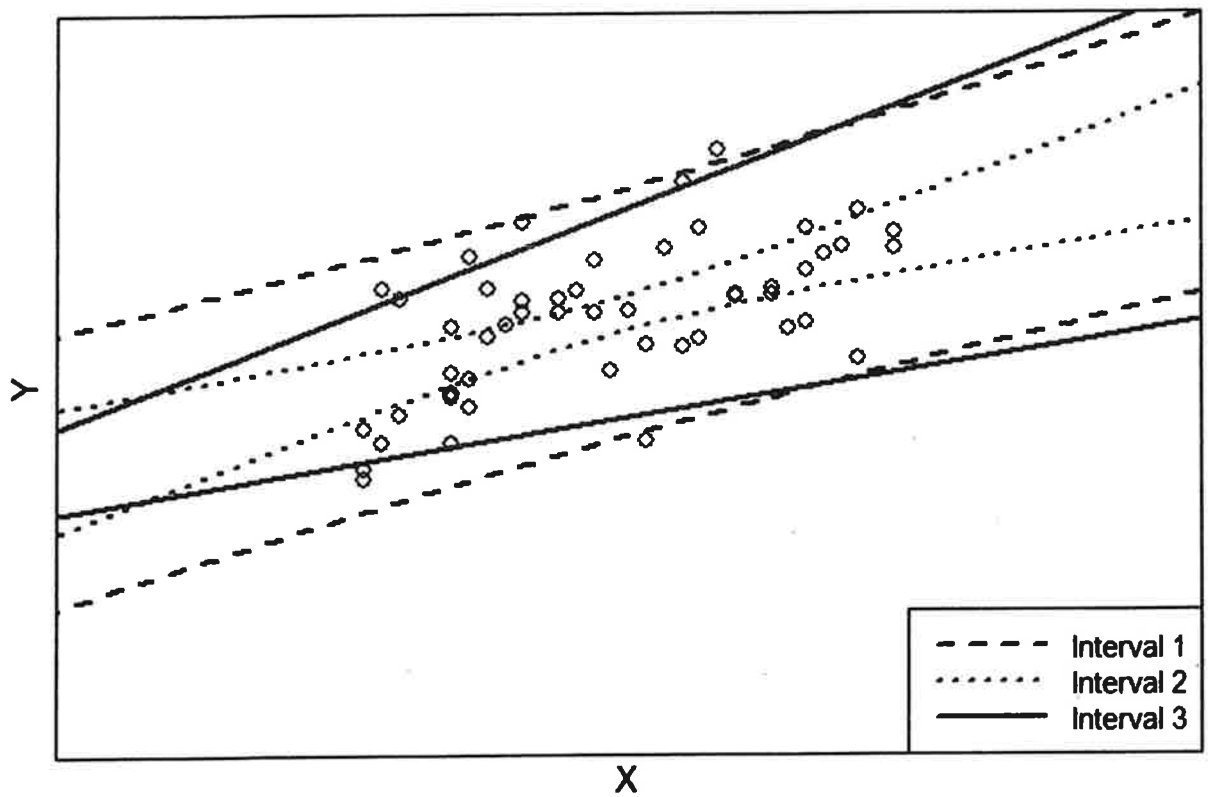
CONTINUED ON NEXT PAGE

36.

An ordinary least squares regression model is fit with the following model form:

$$E(Y_i) = \beta_0 + \beta_1 X_i$$

After fitting the model, the following plot with the original data (points) and three sets of 95% intervals are provided:



Let "CI" be the 95% confidence interval for $E(Y_i)$, and let "PI" be the 95% prediction interval for Y_i.

Determine which one of the following best describes the intervals shown above.

A. Interval 1 = CI Interval 2 = PI
B. Interval 1 = PI Interval 2 = CI
C. Interval 1 = CI Interval 3 = PI
D. Interval 1 = PI Interval 3 = CI
E. None of (A), (B), (C), or (D) are correct

CONTINUED ON NEXT PAGE

37.

You want to perform a regression of Y onto predictors $X_1, X_2, \dots, X_p$, using a large number of observations, and are considering the following modelling techniques:

- Lasso Regression
- Partial Least Squares
- Principal Component Analysis
- Ridge Regression

Determine how many of the above modelling procedures perform variable selection.

A. 0
B. 1
C. 2
D. 3
E. 4

CONTINUED ON NEXT PAGE

38.

Determine which one of the following statements regarding Generalized Additive Models (GAMs) is false.

A. Natural Splines, Regression Splines, Smoothing Splines, Local Regression, Polynomial Regression, and Step Functions are all types of models that can be used as building blocks for GAMs.
B. A limitation of GAMs is that interactions cannot be added to the model.
C. The smoothness of a continuous predictor variable can be summarized via degrees of freedom.
D. GAMs are a useful representation if we are interested in inference, since you can examine the effect of the predictor variables on the response while holding all of the other predictor variables constant.
E. GAMs allow for non-linear relationships between each predictor variable and the response.

CONTINUED ON NEXT PAGE

39.

Three actuaries were given a dataset and asked to build a model to predict claim frequency using any of 5 independent predictors {1, 2, 3, 4, 5} as well as an intercept {I}.

- Actuary A chooses the model using Best Subset Selection
- Actuary B chooses the model using Forward Stepwise Regression
- Actuary C chooses the model using Backwards Stepwise Regression
- When evaluating the models they all used R-squared to compare models with the same number of parameters, and AIC to compare models with different numbers of parameters.

Below are the results for all possible models:

Model	# of Non Intercept Parameters	Parameters	R^2	AIC
1	0	I	0	1.9
2	1	I, 1	0.56	1.4
3	1	I, 2	0.57	1.2
4	1	I, 3	0.55	1.6
5	1	I, 4	0.52	1.7
6	1	I, 5	0.51	1.8
7	2	I, 1, 2	0.61	1.0
8	2	I, 1, 3	0.64	0.5
9	2	I, 1, 4	0.63	0.8
10	2	I, 1, 5	0.69	0.0
11	2	I, 2, 3	0.61	1.0
12	2	I, 2, 4	0.62	0.9
13	2	I, 2, 5	0.68	0.2
14	2	I, 3, 4	0.66	0.4
15	2	I, 3, 5	0.64	0.5
16	2	I, 4, 5	0.60	1.1

Model	# of Non Intercept Parameters	Parameters	R^2	AIC
17	3	I, 1, 2, 3	0.73	1.3
18	3	I, 1, 2, 4	0.71	1.5
19	3	I, 1, 2, 5	0.72	1.4
20	3	I, 1, 3, 4	0.75	1.0
21	3	I, 1, 3, 5	0.76	0.8
22	3	I, 1, 4, 5	0.79	0.2
23	3	I, 2, 3, 4	0.78	0.6
24	3	I, 2, 3, 5	0.74	1.2
25	3	I, 2, 4, 5	0.75	1.1
26	3	I, 3, 4, 5	0.73	1.3
27	4	I, 1, 2, 3, 4	0.88	1.6
28	4	I, 1, 2, 3, 5	0.80	2.1
29	4	I, 1, 2, 4, 5	0.87	1.8
30	4	I, 1, 3, 4, 5	0.83	2.0
31	4	I, 2, 3, 4, 5	0.85	1.9
32	5	I, 1, 2, 3, 4, 5	0.90	3.5

- AIC_j is the AIC of the model chosen by Actuary j

Determine the correct ordering of the AIC values of the three selected models.

A. $AIC_A < AIC_B < AIC_C$
B. $AIC_A = AIC_B < AIC_C$
C. $AIC_A < AIC_C < AIC_B$
D. $AIC_A = AIC_C < AIC_B$
E. The answer is not given by (A), (B), (C) or (D)

CONTINUED ON NEXT PAGE

40.

You are fitting a model to predict the height of an adult male using shoe size as a predictor. You have a sample size of 200.

Determine which one of the following model forms will both have a discontinuity in the fitted curve and most likely overfit the data.

A. Piecewise cubic with two knots
B. Piecewise linear with two knots
C. Cubic spline with one knot
D. Natural cubic spline with two knots
E. Smoothing spline with 10 degrees of freedom

CONTINUED ON NEXT PAGE

41.

The time series below shows study manuals sold by month by company XYZ:

t	Month	Year	Study Manuals Sold
1	Jan	2016	52
2	Feb	2016	18
3	Mar	2016	16
4	Apr	2016	14
5	May	2016	16
6	Jun	2016	16
7	Jul	2016	18
8	Aug	2016	50
9	Sep	2016	20
10	Oct	2016	16
11	Nov	2016	18
12	Dec	2016	20

t	Month	Year	Study Manuals Sold
13	Jan	2017	60
14	Feb	2017	22
15	Mar	2017	17
16	Apr	2017	19
17	May	2017	21
18	Jun	2017	23
19	Jul	2017	25
20	Aug	2017	60
21	Sep	2017	28
22	Oct	2017	22
23	Nov	2017	24
24	Dec	2017	28

Calculate the monthly additive effect at time t = 13 using the centering approach.

A. Less than 25
B. At least 25, but less than 30
C. At least 30, but less than 35
D. At least 35, but less than 40
E. At least 40

CONTINUED ON NEXT PAGE

42.

Consider the following time-series data for price of a stock on January 1 for the last 5 years:

Date	Jan 1, 2013	Jan 1, 2014	Jan 1, 2015	Jan 1, 2016	Jan 1, 2017
Price	63.18	81.89	103.43	123.90	133.53

Calculate the sample autocorrelation at lag 1 for this data.

A. Less than 0.2
B. At least 0.2, but less than 0.4
C. At least 0.4, but less than 0.6
D. At least 0.6, but less than 0.8
E. At least 0.8

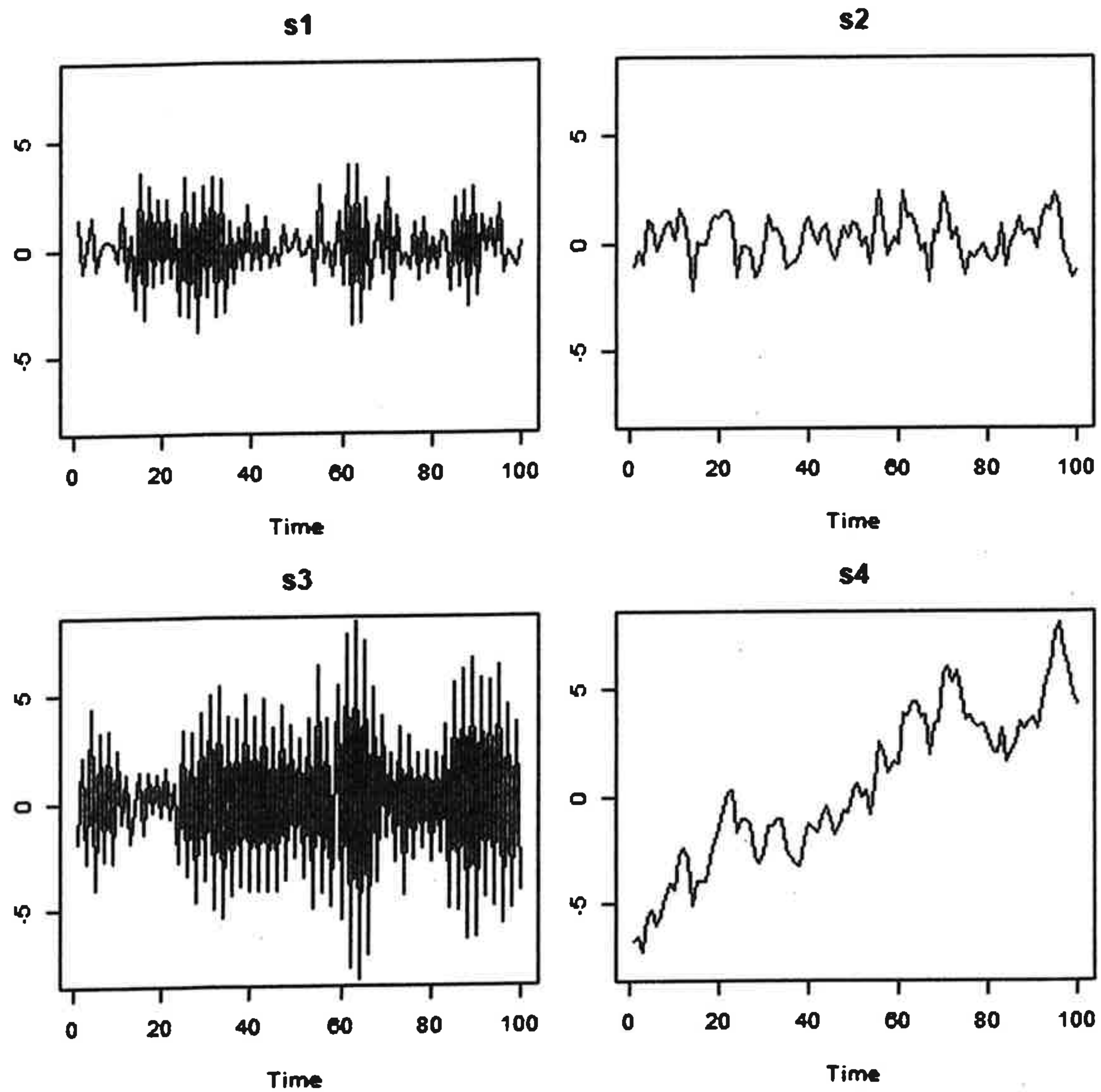
CONTINUED ON NEXT PAGE

43.

You are given an autoregressive time series of order 1:

$$x_t = \alpha_1 x_{t-1} + w_t$$

and the following time series graphs:



Determine the most likely coefficient α_1 for each graph, s1-s4.

A. s1 = 0.90, s2 = 0.50, s3 = 0.99, s4 = 0.99
B. s1 = 0.99, s2 = -0.50, s3 = 0.90, s4 = -0.99
C. s1 = -0.90, s2 = 0.50, s3 = -0.99, s4 = 0.99
D. s1 = -0.99, s2 = 0.50, s3 = -0.90, s4 = 0.99
E. s1 = -0.90, s2 = 0.99, s3 = -0.99, s4 = 0.50

CONTINUED ON NEXT PAGE

44.

You are given the following statements regarding deterministic and stochastic trends:

I. Deterministic trends can be modeled using regression.
II. Stochastic trends usually have a plausible, physical explanations such as an increase in population.
III. Short term extrapolation of deterministic trends can be justified by claiming that underlying trends will usually change slowly in comparison with the forecast lead time.

Determine which of the above statements are true.

A. None are true
B. I and II only
C. I and III only
D. II and III only
E. The answer is not given by (A), (B), (C) or (D)

45.

A time series, $\{x_t\}$ is modeled as an AR(3) process given by:

$$x_t = 4 + 0.45x_{t-1} + 0.25x_{t-2} + 0.05x_{t-3} + w_t$$

where w_t is white noise, with:

- $E(w_t) = 0$
- $Var(w_t) = \sigma^2$

The most recent three observed values in this series are:

t	x_t
2016	30
2017	23
2018	21

Calculate the ten-step-ahead forecast value, x_{2028}.

A. Less than 12
B. At least 12, but less than 16
C. At least 16, but less than 20
D. At least 20, but less than 24
E. At least 24

END OF EXAMINATION

Spring 2019 MAS-I Final Answer Key

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