Exam S

CASUALTY ACTUARIAL SOCIETY
AND THE
CANADIAN INSTITUTE OF ACTUARIES

Exam S

Statistics and Probabilistic Models

May 6, 2016

4 HOURS

INSTRUCTIONS TO CANDIDATES

1. This 90 point examination consists of 45 multiple choice questions 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 2016 and that the exam name is S.
   - 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. <u>No additional time will be allowed for this after the exam has ended.</u> 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.

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. <u>Writing will NOT be permitted during this time and you will not be permitted to hold pens or pencils.</u> <u>You will also not be allowed to use calculators.</u> The supervisor has additional exams for those candidates who have defective exam booklets.

CONTINUE TO NEXT PAGE OF INSTRUCTIONS

- Verify that you have a copy of "Tables for CAS Exam S" 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.
- Candidates should employ a non-parametric test unless otherwise specified in the problem, or when there is a standard distribution that is logically or commonly associated with the random variable in question. Examples of problems with a logical or commonly associated distribution would include exponential wait times for Poisson processes and applications of the Central Limit Theorem.

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.

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 23, 2016.

END OF INSTRUCTIONS

1.

Alice, Bob and Chris are hired by a firm to drive three vehicles – a bus, a taxi and a train (only one driver is needed for each). Each employee has different skills and requires different amounts of training on each vehicle. The cost to train each employee i for vehicle j, $C_{i,j}$, is an independent exponential random variable with mean of 200.

To minimize the total cost of training, the firm uses the following procedure to assign the employees to their vehicle:

- Alice is assigned to the vehicle which minimizes her training cost, $C_{Alice,j}$.
- Bob is then assigned one of the two remaining vehicles which minimizes $C_{Bob,j}$.
- Chris is then assigned the remaining vehicle.

Calculate the firm's expected total cost of training these three employees, if it uses this assignment algorithm.

A. Less than 200
B. At least 200, but less than 300
C. At least 300, but less than 400
D. At least 400, but less than 500
E. At least 500

Exam Continued on Next Page

2.

You are given the following information:

- Vehicles arrive at the drive-through bank at a Poisson rate of 30 per hour.
- On average, 10% of these vehicles require cash transactions.

Calculate the probability that at least 3 vehicles require cash transactions between 10 AM and noon.

A. Less than 0.915
B. At least 0.915, but less than 0.925
C. At least 0.925, but less than 0.935
D. At least 0.935, but less than 0.945
E. At least 0.945

Exam Continued on Next Page

3.

You are given:

- The number of claims, $N(t)$, follows a Poisson process with intensity:

$$\lambda(t) = \frac{1}{2}t,\ 0 < t < 5$$

$$\lambda(t) = \frac{1}{4}t,\ t \geq 5$$

- By time $t = 4$, 15 claims have occurred.

Calculate the probability that exactly 16 claims will have occurred by time $t = 6$.

A. Less than 0.075
B. At least 0.075, but less than 0.125
C. At least 0.125, but less than 0.175
D. At least 0.175, but less than 0.225
E. At least 0.225

Exam Continued on Next Page

4.

You are given the following information about the rate at which calls come into a customer service center:

- The customer service center is open 24 hours a day.
- Calls are received from four regions.
- The rate at which calls arrive is 10 calls per hour and follows a Poisson distribution.
- The probability that a call will be from a given region is independent of other calls and is shown in the table below.

Region	Probability
1	0.2
2	0.3
3	0.1
4	0.4

Calculate the probability that in a one-hour time period at least one call will have been received from each region.

A. Less than 0.25
B. At least 0.25, but less than 0.30
C. At least 0.30, but less than 0.35
D. At least 0.35, but less than 0.40
E. At least 0.40

Exam Continued on Next Page

5.

You are given the following information on a health insurance policy:

- Claims occur according to a Poisson process with mean $\lambda = 10$ per week.
- Claim amounts follow the Pareto distribution with probability density function

$$f(x) = \frac{4*1000^4}{(x+1000)^5} \quad , \; 0 < x$$

Calculate the probability that the total claim amount in 52 weeks exceeds 200,000 using the Normal approximation.

A. Less than 0.02
B. At least 0.02, but less than 0.03
C. At least 0.03, but less than 0.04
D. At least 0.04, but less than 0.05
E. At least 0.05

Exam Continued on Next Page

6.

You are given the following information:

- The lifetimes of all light bulbs follow the exponential distribution.
- A new incandescent light bulb has a hazard rate of 0.20 per year.
- The expected lifetime of an LED light bulb is twice the expected lifetime of a 2-year-old incandescent light bulb.

Calculate the hazard rate per year for an LED light bulb.

A. Less than 0.05
B. At least 0.05, but less than 0.07
C. At least 0.07, but less than 0.09
D. At least 0.09, but less than 0.11
E. At least 0.11

7.

You are given the following information:

- Consider a system of three independent components, each of which functions for an amount of time (in months) uniformly distributed over (0, 1).
- Under current design, the system will fail if any of the 3 components fail.
- A new design was made such that the system will fail if 2 or more components fail.

Calculate the increase in expected system life in months under the new system design.

A. Less than 0.195
B. At least 0.195, but less than 0.215
C. At least 0.215, but less than 0.235
D. At least 0.235, but less than 0.255
E. At least 0.255

Exam Continued on Next Page

8.

In an airport, you are given the following information:

- Taxis arrive in accordance with a Poisson process with rate of $\lambda_T = 2$ per minute.
- Customers arrive in accordance with a Poisson process with rate $\lambda_C = 3$ per minute.
- A taxi will wait no matter how many taxis are present.
- An arriving customer that does not find a taxi waiting will leave.

Calculate the proportion of arriving customers that find taxis.

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

Exam Continued on Next Page

9.

You are given the following information:

- A firm has modeled the annual movement of its workers using a Markov chain with the following three states:

  0. Fully Employed
  1. Temporarily Disabled
  2. Retired

- The associated transition matrix is:

$$P = \begin{bmatrix} 0.7 & 0.1 & 0.2 \\ 0.4 & 0.3 & 0.3 \\ 0.0 & 0.0 & 1.0 \end{bmatrix}$$

- Tom is Fully Employed with the firm this year.

Calculate the probability that Tom will be Fully Employed in two years.

A. Less than 0.52
B. At least 0.52, but less than 0.54
C. At least 0.54, but less than 0.56
D. At least 0.56, but less than 0.58
E. At least 0.58

Exam Continued on Next Page

10.

You are given the following information:

- A taxi driver provides service for two cities, C1 and C2, only.
- If the taxi driver is in the city C1, the probability that he has to drive the customer to city C2 is 0.35.
- If the taxi driver is in the city C2, the probability that he has to drive the customer to city C1 is 0.75.
- The expected profit for each trip is:
  - Within C1: \$6
  - Within C2: \$5
  - Between C1 and C2: \$10

Calculate the long-term expected profit per trip.

A. Less than \$7.50
B. At least \$7.50, but less than \$8.00
C. At least \$8.00, but less than \$8.50
D. At least \$8.50, but less than \$9.00
E. At least \$9.00

Exam Continued on Next Page

11.

You are given the following information:

A three-state Markov Chain, with the following transition probability matrix, is used to model the movement of policyholders between three States:

States	0	1	2
0	0.5	0.5	0.0
1	0.4	0.4	0.2
2	0.3	0.4	0.3

Calculate the stationary percentage of policyholders in State 2.

A. Less than 13%
B. At least 13%, but less than 23%
C. At least 23%, but less than 33%
D. At least 33%, but less than 43%
E. At least 43%

12.

You are given the following information:

- A 3-year term insurance policy on a life age (80) provides for a death benefit, payable at the end of the year of death.
- The death benefit is $1000t,\ t \in \{1, 2, 3\}$, where t is the year of death.
- This policy is purchased by a single premium, P, at time 0.
- If (80) lives to age 83, the single premium is returned without interest.
- Mortality rates follow the Illustrative Life Table.
- $i = 0.10$

Calculate P using the equivalence principle.

A. Less than 830
B. At least 830, but less than 850
C. At least 850, but less than 870
D. At least 870, but less than 890
E. At least 890

Exam Continued on Next Page

13.

Let $Y_1, \dots, Y_n$ denote a random sample from the uniform distribution on the interval: $(\theta, \theta + 1)$.

You are given:

- There are two unbiased estimators for θ, $\hat{\theta}_1$ and $\hat{\theta}_2$

$$\hat{\theta}_1 = \bar{Y} - \frac{1}{2}$$

$$\hat{\theta}_2 = Y_{(n)} - \frac{n}{n+1}, \text{ where } Y_{(n)} = max(Y_1, \dots, Y_n)$$

- The efficiency of $\hat{\theta}_1$ relative to $\hat{\theta}_2$ is given by $eff(\hat{\theta}_1, \hat{\theta}_2) = \frac{12n^2}{(n+2)(n+1)^2}$
- s is the minimum sample size required to make $\hat{\theta}_2$ more efficient than $\hat{\theta}_1$
- s is less than 10

Calculate $Var(\hat{\theta}_2 | n = s)$.

A. Less than 0.007
B. At least 0.007, but less than 0.009
C. At least 0.009, but less than 0.011
D. At least 0.011, but less than 0.013
E. At least 0.013

Exam Continued on Next Page

14.

You are given the following information:

- Loss severity follows an exponential distribution with mean θ.
- Small claims are handled directly by the insured, and no information about losses less than 10,000 is available to you.
- You observe the following four claim payments net of the 10,000 deductible:

11,000 10,500 12,000 15,000

Calculate the Maximum Likelihood Estimate of θ.

A. Less than 2,000
B. At least 2,000, but less than 2,050
C. At least 2,050, but less than 2,100
D. At least 2,100, but less than 2,150
E. At least 2,150

Exam Continued on Next Page

15.

You are given three normal distributions with the same mean and varying variances,

$\frac{1}{\theta}, \frac{1}{2\theta}, \frac{1}{4\theta}$.

You select one random draw from each of the distributions respectively:

0, 4, 8

The probability density of the normal distribution is:

$$f(x \mid \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Determine the Maximum Likelihood Estimate of θ.

A. Less than 0.04
B. At least 0.04, but less than 0.06
C. At least 0.06, but less than 0.08
D. At least 0.08, but less than 0.10
E. At least 0.10

Exam Continued on Next Page

16.

You are given:

- A density function, $f(x) = 5\theta x^3 e^{-\theta x^4}$
- A random sample of size n from this distribution

Calculate the Rao-Cramer lower bound for the variance of an unbiased estimator of θ.

A. $\frac{\theta^2}{n}$

B. $\frac{1}{\theta n}$

C. $\frac{\theta^2}{4n}$

D. $\frac{5\theta^2}{n-1}$

E. $\frac{n}{\theta^2}$

Exam Continued on Next Page

17.

Determine the Fisher Information of n independent samples from a geometric distribution with mean β.

A. $1/\beta$

B. n/β

C. $n/\beta - n/(\beta + 1)$

D. $n\beta/(\beta + 1)$

E. $\beta(\beta + 1)/n$

Exam Continued on Next Page

18.

You are testing whether the number of claims in year 2 is independent of the number of claims in year 1.

You are given the following contingency table:

		Claims in Year 2	
		0	1+
Claims in Year 1	0	20	21
	1+	30	21

Calculate the Chi-Square test statistic.

a) Less than 0.5
b) At least 0.5, but less than 1.0
c) At least 1.0, but less than 1.5
d) At least 1.5, but less than 2.0
e) At least 2.0

Exam Continued on Next Page

19.

To test whether taking a driver's education course improves driving skills, you set up an experiment. Participants are given a pretest of their driving skills and are tested on their skills again after the course. The data is contained in the table below.

User ID	Pretest	Posttest	Difference
1	62	60	-2
2	93	94	1
3	97	100	3
4	77	80	3
5	69	75	6
6	72	78	6
7	99	100	1
8	55	57	2
9	90	91	1
10	66	63	-3
11	76	83	7
12	51	49	-2
13	74	84	10
14	52	61	9
15	97	99	2
16	54	52	-2
Sum	1184	1226	42
Sum of Squares	91940	98616	352

Assume the scores follow a Normal distribution.

Calculate the statistic to test if there is a significant difference in the two scores.

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

20.

You are given the following information:

- The Underwriting Department tells you that 40% of homeowners who stated they have alarm systems have cancelled their alarm service.
- A survey of 1,000 homeowners who stated they have alarm systems shows that 428 homeowners have cancelled their service.
- You test the following hypotheses:
  - H_0: The Underwriting Department's claim is accurate.
  - H_1: The Underwriting Department's claim is inaccurate.

Calculate the minimum significance level at which you would reject the null hypothesis.

A. Less than 0.005
B. At least 0.005, but less than 0.010
C. At least 0.010, but less than 0.025
D. At least 0.025, but less than 0.050
E. At least 0.050

Exam Continued on Next Page

21.

Given

- $X_1, X_2, \ldots, X_{10}$ are random samples from a Normal distribution with mean μ_x and variance 7
- $Y_1, Y_2, \ldots, Y_8$ are random samples from a Normal distribution with mean μ_y and variance 9
- X and Y are independent
- $W = \bar{X} - \bar{Y}$ is a Normal distribution with mean $\mu_x - \mu_y$
- $H_o: \mu_x = \mu_y$
- $H_1: \mu_x > \mu_y$

For W = 2, what is the result of the hypothesis test using the Normal distribution?

A. Reject H_o at the 0.025 level.
B. Reject H_o at the 0.050 level, but not at the 0.025 level.
C. Reject H_o at the 0.075 level, but not at the 0.050 level.
D. Reject H_o at the 0.100 level, but not at the 0.075 level.
E. Do not reject H_o at the 0.100 level.

Exam Continued on Next Page

22.

You are given the following:

- The time it takes an insurance company to adjust a claim follows a Normal distribution with $\mu = 90$ and $\sigma = 30$. (Time is in minutes)
- The insurance company is considering purchasing software from a vendor that claims the software will reduce adjusting time by ten minutes per claim but without any impact on variation. Before purchasing the software, the insurance company will test the software on a sample of 100 claims and measure the time to adjust claims.
- The insurance company will purchase the software only if the improvement is shown to be at least ten minutes and the sample result is subject to no more than a 5% Type I error.

Calculate the probability of a Type II error if the vendor's claim is true.

A. At least 1%, but less than 2%
B. At least 2%, but less than 3%
C. At least 3%, but less than 4%
D. At least 4%, but less than 5%
E. At least 5%

Exam Continued on Next Page

23.

To test if the distribution of the grades of two classes which have the same teacher are the same, 7 students from Class A and 9 from Class B are randomly selected. The sum of the ranks for Class A is 75.

Calculate the Mann–Whitney U test statistic.

A. Less than 17
B. 17
C. 18
D. 19
E. At least 20

Exam Continued on Next Page

24.

You are given:

- Y_1, Y_2, and Y_3 denotes a random sample from an exponential distribution

$$f(y) = \frac{1}{5}e^{-\frac{y}{5}} \qquad 0 < y < \infty$$

- The second-order statistic is $Y_{(2)}$

Calculate $E[Y_{(2)}]$.

A. Less than 3.0
B. At least 3.0, but less than 4.0
C. At least 4.0, but less than 5.0
D. At least 5.0, but less than 6.0
E. At least 6.0

Exam Continued on Next Page

25.

Two models are used to determine annual losses for a group of policies.

- Ten policies are selected at random

Policy	Model Y	Model Z
A	500	400
B	500	525
C	500	650
D	750	925
E	750	950
F	1000	1225
G	1250	1125
H	1250	1300
I	1250	1500
J	1500	1575

- H_0: $Median_Y = Median_Z$
- H_1: $Median_Y < Median_Z$

Calculate the smallest value for significance level for which H_0 can be rejected using the Signed-Rank Wilcoxon test for a matched pairs experiment.

A. Less than 0.010
B. At least 0.010, but less than 0.025
C. At least 0.025, but less than 0.050
D. At least 0.050, but less than 0.100
E. At least 0.100

Exam Continued on Next Page

26.

You are given:

- $X_1, X_2, \ldots, X_n$ is a random sample from a $N(\mu, \sigma^2)$
- $\sigma^2 = 200$
- The prior distribution is a normal $N(\mu_0, \sigma_0^2)$ distribution
- $\mu_0 = 70$ and $\sigma_0^2 = 10$
- There are 5 observations from this distribution, X_i:

53 22 61 86 70

Calculate the lower bound of the 95% Bayesian credibility interval for μ.

A. Less than 55
B. At least 55, but less than 60
C. At least 60, but less than 65
D. At least 65, but less than 70
E. At least 70

Exam Continued on Next Page

27.

For ABC Insurance Company, you are given the following information:

- All insureds come from one of three population classes.
- The claim frequency of each insured follows a Poisson process.
- The claims information is as follows:

Population Class	Expected time between claims	Probability of being in class	Claim cost
I	12 months	0.25	\$5,000
II	18 months	0.25	\$5,000
III	36 months	0.50	\$5,000

Calculate the expected loss in year 2 for an insured that had no claims in year 1.

A. Less than 2,400
B. At least 2,400, but less than 2,500
C. At least 2,500, but less than 2,600
D. At least 2,600, but less than 2,700
E. At least 2,700

Exam Continued on Next Page

28.

You are given the following information:

- There are five observations, y_1, y_2, y_3, y_4, y_5, from a gamma distribution:

6.85	10.3	20.89	15.06	6.26

- $\mathrm{Var}(Y) = 37.67$

- The gamma distribution $G(\mu, \nu)$ is a member of the exponential family with:

$$a(\theta) = -\ln(-\theta)$$

$$E(Y) = \mu$$

$$\mathrm{Var}(Y) = \frac{\mu^2}{\nu}$$

Calculate the maximum likelihood estimate of the dispersion parameter, φ, for this distribution.

A. Less than 0.00
B. At least 0.00, but less than 1.00
C. At least 1.00, but less than 2.00
D. At least 2.00, but less than 3.00
E. At least 3.00

Exam Continued on Next Page

29.

You are given the following information for a fitted GLM:

Response variable	Occurrence of Accidents
Response distribution	Binomial
Link	Logit

Parameter	df	$\hat{\beta}$
Intercept	1	x
Driver's Age	2	
1	1	0.288
2	1	0.064
3	0	0
Area	2	
A	1	-0.036
B	1	0.053
C	0	0
Vehicle Body	2	
Bus	1	1.136
Other	1	-0.371
Sedan	0	0

The probability of a driver in age group 2, from area C and with vehicle body type Other, having an accident is 0.22.

Calculate the odds ratio of the driver in age group 3, from area C and with vehicle body type Sedan having an accident.

A. Less than 0.200
B. At least 0.200, but less than 0.250
C. At least 0.250, but less than 0.300
D. At least 0.300, but less than 0.350
E. At least 0.350

Exam Continued on Next Page

30.

You are given the following information for a fitted GLM:

Response variable	Occurrence of Accidents		
Response distribution	Binomial		
Link	Logit		

Parameter	df	$\hat{\beta}$	se
Intercept	1	-2.358	0.048
Area	2		
Suburban	0	0.000	
Urban	1	0.905	0.062
Rural	1	-1.129	0.151

Calculate the modeled probability of an Urban driver having an accident.

A. Less than 0.01
B. At least 0.01, but less than 0.05
C. At least 0.05, but less than 0.10
D. At least 0.10, but less than 0.20
E. At least 0.20

31.

You are given the following information for a fitted GLM:

Response variable		Claim size
Response distribution		Gamma
Link		Log
Dispersion parameter		1
Parameter	**df**	$\hat{\beta}$
Intercept	1	2.100
Zone	4	
1	1	7.678
2	1	4.227
3	1	1.336
4	0	0.000
5	1	1.734
Vehicle Class	6	
Convertible	1	1.200
Coupe	1	1.300
Sedan	0	0.000
Truck	1	1.406
Minivan	1	1.875
Stationwagon	1	2.000
Utility	1	2.500
Driver Age	2	
Youth	1	2.000
Middle age	0	0.000
Old	1	1.800

Calculate the predicted claim size for an observation from Zone 3, with Vehicle Class Truck and Driver Age Old.

A. Less than 650
B. At least 650, but less than 700
C. At least 700, but less than 750
D. At least 750, but less than 800
E. At least 800

Exam Continued on Next Page

32.

You are given the following information for a fitted GLM:

Response variable		Claim size
Response distribution		Gamma
Link		Log
Dispersion parameter		1
Parameter	**df**	$\hat{\beta}$
Intercept	1	2.100
Zone	4	
1	1	7.678
2	1	4.227
3	1	1.336
4	0	0.000
5	1	1.734
Vehicle Class	6	
Convertible	1	1.200
Coupe	1	1.300
Sedan	0	0.000
Truck	1	1.406
Minivan	1	1.875
Stationwagon	1	2.000
Utility	1	2.500
Driver Age	2	
Youth	1	2.000
Middle age	0	0.000
Old	1	1.800

Calculate the variance of a claim size for an observation from Zone 4, with Vehicle Class Sedan and Driver Age Middle age

A. Less than 55
B. At least 55, but less than 60
C. At least 60, but less than 65
D. At least 65, but less than 70
E. At least 70

Exam Continued on Next Page

33.

You are given the following information for a GLM of customer retention:

Response variable		Retention
Response distribution		Binomial
Link		Logit
Parameter	df	$\hat{\beta}$
Intercept	1	1.530
Number of Drivers	1	
1	0	0.000
> 1	1	0.735
Last Rate Change	2	
< 0%	0	0.000
0%-10%	1	-0.031
> 10%	1	-0.372

Calculate the probability of retention for a policy with 3 drivers and a prior rate change of 5%.

A. Less than 0.850
B. At least 0.850, but less than 0.870
C. At least 0.870, but less than 0.890
D. At least 0.890, but less than 0.910
E. At least 0.910

Exam Continued on Next Page

34.

Determine which of the following statements are true.

I. The deviance is useful for testing the significance of explanatory variables in nested models.

II. The deviance for normal distributions is proportional to the residual sum of squares.

III. The deviance is defined as a measure of distance between the saturated and fitted model.

A. I only
B. II only
C. III only
D. All but III
E. All

Exam Continued on Next Page

35.

You are given the following information about three candidates for a Poisson frequency GLM on a group of condominium policies:

Model	Variables in the Model	DF	Log Likelihood	AIC	BIC
1	Risk Class	5	-47,704	95,418	95,473.61182
2	Risk Class + Region		-47,495		
3	Risk Class + Region + Claim Indicator	10	-47,365	94,750	

- Insureds are from one of five Risk Class: A, B, C, D, E
- Condominium policies are located in several regions
- Claim Indicator is either Yes or No
- All models are built on the same data

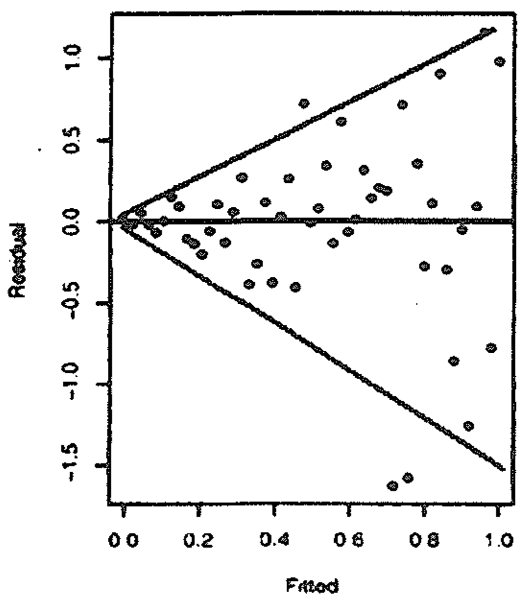
Calculate the absolute difference between the AIC and the BIC for Model 2.

A. Less than 85
B. At least 85, but less than 95
C. At least 95, but less than 105
D. At least 105, but less than 115
E. At least 115

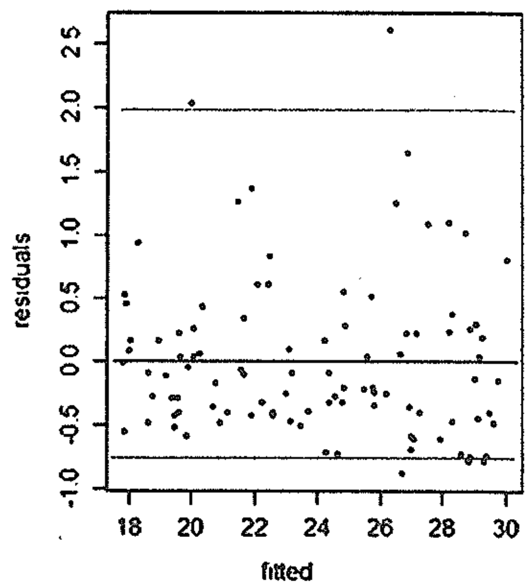
36.

You are given the following two graphs comparing the fitted values to the residuals of two different linear models:

Graph 1



Graph 2



Determine which of the following statements are true.

I. Graph 1 indicates the data is homoscedastic
II. Graph 1 indicates the data is heteroskedastic
III. Graph 2 indicates the data is non-normal

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

Exam Continued on Next Page

37.

Determine which of the following GLM selection considerations is true.

A. The model with the largest AIC is always the best model in model selection process.
B. The model with the largest BIC is always the best model in model selection process.
C. The model with the largest deviance is always the best model in model selection process.
D. Other things equal, when the number of observations > 1000, AIC penalizes more for the number of parameters used in the model than BIC.
E. Other things equal, when number of observations > 1000, BIC penalizes more for the number of parameters used in the model than AIC.

Exam Continued on Next Page

38.

You are testing the addition of a new categorical variable into an existing GLM, and are given the following information:

- A is the change in AIC and B is the change in BIC after adding the new variable.
- $B > A + 25$
- There are 1500 observations in the model.

Calculate the minimum possible number of levels in the new categorical variable.

A. Less than 3
B. 3
C. 4
D. 5
E. More than 5

Exam Continued on Next Page

39.

Determine which of the following statements about GLMs are true.

I. The saturated model has the highest possible deviance.

II. Deviance follows a χ^2distribution for all models in the exponential family.

III. Deviance is a useful measure of goodness of fit for all models in the exponential family.

A. I only
B. II only
C. III only
D. I, II and III
E. None of the above

Exam Continued on Next Page

40.

You are given the following results from a fitted GLM on the frequency of accidents:

Parameter	df	$\hat{\beta}$	se
Intercept	1	-11.2141	0.1826
Location	1		
Rural	0	0.0000	
City	1	1.0874	0.3162

Calculate the Wald statistic for testing the null hypothesis of $\beta_{city} = 0$.

A. Less than 10.5
B. At least 10.5, but less than 11.5
C. At least 11.5, but less than 12.5
D. At least 12.5, but less than 13.5
E. At least 13.5

Exam Continued on Next Page

41.

A Poisson regression model with log link is used to estimate the number of diabetes deaths. The parameter estimates for the model are:

Response variable	Number of Diabetes Deaths		
Response distribution	Poisson		
Link	Log		
Parameter	df	$\hat{\beta}$	p-value
Intercept	1	-15.000	<0.0001
Gender: Female	1	-1.200	<0.0001
Gender: Male	0	0.000	
Age	1	0.150	<0.0001
Age^2	1	0.004	<0.0001
Age×Gender: Female	1	0.012	<0.0001
Age×Gender: Male	0	0.000	

Calculate the expected number of deaths for a population of 100,000 females age 25.

A. Less than 3
B. At least 3, but less than 5
C. At least 5, but less than 7
D. At least 7, but less than 9
E. At least 9

42.

You are building a model of claim counts, and in order to evaluate the risk of overdispersion, you calculate the following:

- The log-likelihood of the Poisson regression is $l_P = 350$.
- The log-likelihood of the Negative Binomial regression is l_{NB}.

At a significance level of 0.5%, you reject the hypothesis that a Poisson response is more appropriate than a Negative Binomial response.

Calculate the smallest possible value of l_{NB}.

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

43.

You are given the following information:

- The time series $\{x_t\}$ is stationary in mean and variance.
- The first 8 observed values of x are as follow:

t	x_t
1	13
2	10
3	12
4	9
5	8
6	8
7	7
8	5

Calculate the sample auto covariance function, c_k, for lag $k = 2$.

A. Less than 1.80
B. At least 1.80, but less than 1.90
C. At least 1.90, but less than 2.00
D. At least 2.00, but less than 2.10
E. At least 2.10

Exam Continued on Next Page

44.

You are given the following $ARMA(p,q)$ model:

$$x_t = \frac{5}{2}x_{t-1} - 2x_{t-2} + \frac{1}{2}x_{t-3} + w_t - \frac{1}{2}w_{t-1}$$

The model is parameter redundant.

Which of the following statements is false?

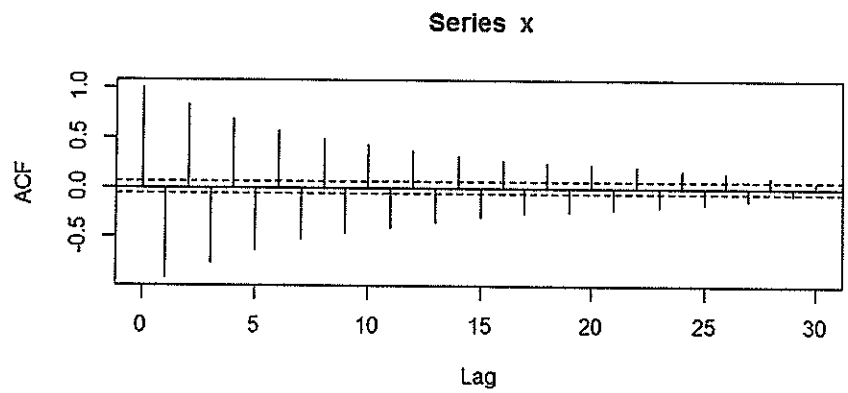
A. The model is non-stationary.
B. The model is $ARMA(3,1)$.
C. The model is invertible.
D. The model can be restated as $ARMA(2,0)$.
E. None of the above.

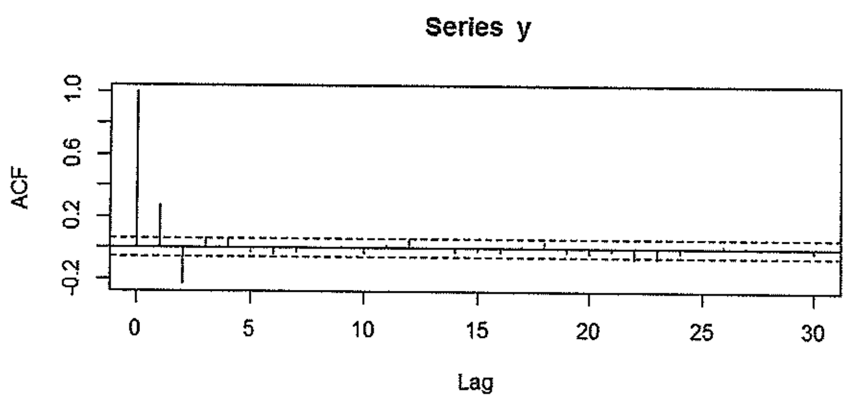
Exam Continued on Next Page

45.

You are given the following information:

- x and y are two time series that are stationary.
- The graphs below are generated using the ACF (auto correlation function) function of R.
- The dashed lines above and below zero indicate the range within which the ACF results are considered not significantly different than zero.
- The symbol w_t is white noise with zero mean





Which of the following statements displays the model structure that best describes series x and series y?

A. $x_t = 0.9x_{t-1} + w_t$ and $y_t = 0.9y_{t-1} + w_t + 0.6w_{t-1} - 0.3w_{t-2}$
B. $x_t = -0.9x_{t-1} + w_t$ and $y_t = w_t + 0.6w_{t-1} - 0.3w_{t-2}$
C. $x_t = 0.9x_{t-1} + w_t$ and $y_t = w_t + 0.6w_{t-1}$
D. $x_t = -0.9x_{t-1} + w_t$ and $y_t = w_t + 0.6w_{t-1} - 0.3w_{t-2} + 0.4w_{t-3}$
E. $x_t = 0.9x_{t-1} + w_t$ and $y_t = w_t + 0.6w_{t-1} - 0.3w_{t-2}$

End of Examination

Spring 2016 Exam S Answer Key

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