Exam S

CASUALTY ACTUARIAL SOCIETY
AND THE
CANADIAN INSTITUTE OF ACTUARIES

Expertise. Insight. Solutions.

Exam S

Statistics and Probabilistic Models

October 31, 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 Fall 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. 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 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 November 17, 2016.

END OF INSTRUCTIONS

1.

You are given the following information:

- An insurance company pays claims according to a Poisson process at a rate of 5 per day.
- Claims are sub-divided into three categories: Minor, Major, and Severe, with claim amounts provided below:

Category	Claim Amount
Minor	1
Major	4
Severe	10

- It is known that the proportion of claims in the Severe category is 0.15.
- The total expected claim payment amount in one day is 17.

Calculate the proportion of Major claims.

A. Less than 0.25
B. At least 0.25, but less than 0.40
C. At least 0.40, but less than 0.55
D. At least 0.55, but less than 0.70
E. At least 0.70

2.

For an auto policy, claims occur according to a Poisson process with rate λ=10 per day. Once reported, claims are independently classified as Property Damage (PD) only, Bodily Injury (BI) only, or both PD and BI. You are given that 30 percent of claims are PD only, 20 percent are BI only, and the remaining 50 percent are both PD and BI.

Calculate the probability that in a given day there will be exactly 4 PD only claims, 2 BI only claims, and 6 both PD and BI claims.

A. Less than 0.63 percent
B. At least 0.63 percent, but less than 0.68 percent
C. At least 0.68 percent, but less than 0.73 percent
D. At least 0.73 percent, but less than 0.78 percent
E. At least 0.78 percent

Examination Continued On Next Page

3.

The time X to wait in line is an exponentially distributed random variable with mean 5 minutes.

Calculate the probability that the total waiting time will be longer than 30 minutes from the time that individual arrived in line, given that the wait has already been 20 minutes.

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

Examination Continued On Next Page

4.

Steve catches fish at a Poisson rate of 3 per hour. The price Steve gets at the market for each fish is randomly distributed as follows:

Price	Probability
$10	20%
$20	60%
$30	20%

Using the normal approximation without a continuity correction, calculate the probability that Steve will receive at least $300 for fish caught in a four hour-period.

A. Less than 0.19
B. At least 0.19, but less than 0.20
C. At least 0.20, but less than 0.21
D. At least 0.21, but less than 0.22
E. At least 0.22

5.

A call center currently has 2 representatives and 2 interns who can handle customer calls. If all representatives including interns are currently on a call, an incoming call will be placed on hold until a representative or intern is available.

You are given the following information:

- For each representative, the time taken to handle each call is given by an exponential distribution with a mean value equal to 1.
- For each intern, the time taken to handle each call is given by an exponential distribution with mean 2.
- Handle times are independent

A customer calls the call center and is placed on hold, and is the first person in line.

Calculate the expected time to complete the call (including both hold time and service).

A. Less than 1.2
B. At least 1.2, but less than 1.4
C. At least 1.4, but less than 1.6
D. At least 1.6, but less than 1.8
E. At least 1.8

6.

You are given the following information:

- At the time of installation a given machine has the following survival function:

  $S(x) = \left(1 - \frac{x}{100}\right)^{1/2}$, for $0 \leq x \leq 100; x$ in months

- After 16 months, that machine is still functioning.

Calculate the probability that the machine will stop working between months 36 and 51.

A. Less than 0.095
B. At least 0.095, but less than 0.105
C. At least 0.105, but less than 0.115
D. At least 0.115, but less than 0.125
E. At least 0.125

Examination Continued On Next Page

7.

You are given the following information about a watch with 6 different parts:

- There are 3 red wires with expected lifetimes of 50, 75, and 100.
- There are 3 yellow wires with expected lifetimes of 25, 50, and 75.
- The lifetimes of all wires are independent and exponentially distributed.

Calculate the probability that a red wire will break down before a yellow wire.

A. Less than 0.20
B. At least 0.20, but less than 0.25
C. At least 0.25, but less than 0.30
D. At least 0.30, but less than 0.35
E. At least 0.35

Examination Continued On Next Page

8.

For a parallel system with two independent machines, you are given the following information:

- The hazard rate for each machine is:

  $\mu_x = \frac{1}{100-x}$, for $0 \leq x < 100$; x in months

- One machine has worked for 40 months and the other machine has worked for 60 months.

Calculate the probability that the system will function for 20 more months.

A. Less than 0.35
B. At least 0.35, but less than 0.55
C. At least 0.55, but less than 0.75
D. At least 0.75, but less than 0.95
E. At least 0.95

9.

You are given the following information:

- A system has two minimal path sets: $\{1, 2, 4\}$ and $\{1, 3, 4\}$.
- Reliability for components 1 and 2 is uniformly distributed from 0 to 1.
- Reliability for components 3 and 4 is uniformly distributed from 0 to 2.
- All components in the system are independent.
- You are starting at time 0.

Calculate the probability that the lifetime of the system will be less than 0.25.

A. Less than 0.30
B. At least 0.30, but less than 0.35
C. At least 0.35, but less than 0.40
D. At least 0.40, but less than 0.45
E. At least 0.45

Examination Continued On Next Page

10.

You are given the following information about a system:

- This is a 2-out-of-3 system.
- All components are independent.
- The probability of each component functioning is $p = 0.90$

Calculate the reliability of the system.

A. Less than 0.970
B. At least 0.970, but less than 0.975
C. At least 0.975, but less than 0.980
D. At least 0.980, but less than 0.985
E. At least 0.985

11.

A firm classifies its workers into one of four categories based on their employment status:

1. Full-Time
2. Part-Time
3. On-Leave
4. Retired

Workers independently transition between categories at the end of each year according to a Markov chain with the following transition probability matrix:

$$P = \begin{bmatrix} 0.6 & 0.1 & 0.1 & 0.2 \\ 0.4 & 0.4 & 0.1 & 0.1 \\ 0.1 & 0.1 & 0.7 & 0.1 \\ 0.0 & 0.0 & 0.0 & 1.0 \end{bmatrix}$$

Calculate the probability a current full-time employee will ever go on leave.

A. Less than 0.15
B. At least 0.15, but less than 0.30
C. At least 0.30, but less than 0.45
D. At least 0.45, but less than 0.60
E. At least 0.60

Examination Continued On Next Page

12.

You are given the following information:

- Stock XYZ pays a quarterly dividend based on the current state of the economy, as follows:

State	Quarterly Dividend per Share
1: Economic Peak	0.30
2: Recession	0.01
3: Economic Trough	0.00
4: Economic Expansion	0.20

- The economy changes among these four states each quarter according to the transition probability matrix:

$$P = \begin{bmatrix} 0.20 & 0.80 & 0.00 & 0.00 \\ 0.00 & 0.70 & 0.30 & 0.00 \\ 0.00 & 0.00 & 0.40 & 0.60 \\ 0.40 & 0.00 & 0.00 & 0.60 \end{bmatrix}$$

- Annual interest rate $i = 0\%$.
- Transition happens after the dividend is paid.

Calculate the long-term quarterly dividend per share for Stock XYZ.

A. Less than 0.08
B. At least 0.08, but less than 0.11
C. At least 0.11, but less than 0.14
D. At least 0.14, but less than 0.17
E. At least 0.17

13.

You are given the following information for ABC Insurance Company, which sells auto policies:

- Customers transition each policy period between three states. The average service cost per customer per policy period varies by state as follows:

State	**Average Service Cost per Customer**
1: Gold	2
2: Silver	1
3: Copper	15

- The transition probability matrix is:

$$P = \begin{bmatrix} 0.80 & 0.20 & 0.00 \\ 0.20 & 0.60 & 0.20 \\ 0.00 & 0.40 & 0.60 \end{bmatrix}$$

Calculate the long-term average service cost per customer per policy period.

A. Less than 3.0
B. At least 3.0, but less than 5.0
C. At least 5.0, but less than 7.0
D. At least 7.0, but less than 9.0
E. At least 9.0

14.

Two models are used to estimate the probability of at least one claim for an auto policy in a given year. The models are defined as follows:

- Model I: Independent events with probability of occurrence of at least one claim in a given year equal to 0.1.
- Model II: Markov Chain with conditional probability of at least one claim given no claims in prior year equal to 0.08, and conditional probability of at least one claim given at least one claim in prior year equal to 0.12.

There are no claims in Year 1.

The probability of at least one claim in Year 2 and no claims in Year 3 and at least one claim in Year 4 are denoted by P_I and P_{II}, for models I and II respectively.

Calculate $| P_I - P_{II} |$.

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

15.

You are given the following information:

- A new medical breakthrough is expected to reduce the probability of death between ages 70 and 75 by one-half.
- Prior to the breakthrough, mortality follows the Illustrative Life Table.

Calculate the probability that a life (50) lives to age 75 after the breakthrough.

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

Examination Continued On Next Page

16.

You are given the following information:

- There are three independent lives (50), (60), and (70).
- Mortality follows the Illustrative Life Table.

Calculate the probability that (70) will be the only person alive after 20 years.

A. Less than 0.0200
B. At least 0.0200, but less than 0.0220
C. At least 0.0220, but less than 0.0240
D. At least 0.0240, but less than 0.0260
E. At least 0.0260

17.

You are given an independent random sample of size n, $x_1, x_2, \ldots, x_n$, from an exponential distribution with mean θ.

Calculate the Rao-Cramer lower bound for $\hat{\theta}$, an unbiased estimator of θ.

A. θ
B. $\bar{X}$
C. $1/\theta^2$
D. θ^2/n
E. $1/\hat{\theta}^2$

18.

You are given an independent random sample of size 50, $y_1, y_2, \ldots, y_{50}$, from a population with mean μ and variance σ^2.

Two estimators for μ are:

- $\hat{\mu}_1$= the sample mean
- $\hat{\mu}_2$= the weighted average sample mean, where the first 10 observations receive twice as much weight as the other observations.

Calculate the efficiency of $\hat{\mu}_1$relative to $\hat{\mu}_2$.

A. Less than 1.025
B. At least 1.025, but less than 1.050
C. At least 1.050, but less than 1.075
D. At least 1.075, but less than 1.100
E. At least 1.100

19.

You are given:

I. If $\hat{\theta}_n$ is an unbiased estimator of θ and the variance of $\hat{\theta}_n = 1/n$, it must be a consistent estimator of θ.
II. If $\hat{\theta}_n$ is a consistent estimator of θ, it must be an unbiased estimator of θ.
III. If $\hat{\theta}_n$ is a biased estimator of θ, it is not a consistent estimator of θ.

Determine which of the following is correct.

A. I is true; II is false; III is true.
B. I is false; II is false; III is true.
C. I is true; II is false; III is false.
D. I is true; II is true; III is false.
E. I is false; II is true; III is true.

Examination Continued On Next Page

20.

You are given an independent random sample from a normal distribution X with unknown mean, μ. You are given the following information:

- $\sigma^2(X) = 400$
- $H_0: \mu = 0$
- $H_1: \mu = 10$
- You will reject the null hypothesis if $\bar{X} > \alpha$ for some value of α.

Calculate the minimum sample size needed so that the probability of a Type I error and the probability of a Type II error are each no more than 10%.

A. Less than 10
B. At least 10, but less than 15
C. At least 15, but less than 20
D. At least 20, but less than 25
E. At least 25

21.

You are given:

- A coin has two sides, Heads and Tails
- This coin is tossed 400 times
- H_0: Probability of Heads = 0.50
- H_1: Probability of Heads $\neq$ 0.50
- H_0 is not rejected if the number of heads in the 400 tosses is between 184 and 216 inclusive, and rejected otherwise
- The normal approximation to the binomial distribution is used

Calculate the probability of a Type I error.

A. Less than 0.01
B. At least 0.01, but less than 0.03
C. At least 0.03, but less than 0.05
D. At least 0.05, but less than 0.07
E. At least 0.07

Examination Continued On Next Page

22.

You are given:

- A coin has two sides, Heads and Tails
- This coin is tossed 150 times
- H_0: Probability of Heads = 0.50
- H_1: Probability of Heads = 0.60
- H_0 is not rejected if the number of heads in the 150 tosses is less than 79, and rejected otherwise
- The normal approximation to the binomial distribution is used

Calculate the probability of a Type II error.

A. Less than 0.01
B. At least 0.01, but less than 0.03
C. At least 0.03, but less than 0.05
D. At least 0.05, but less than 0.06
E. At least 0.06

Examination Continued On Next Page

23.

You are given:

- H_0: $\theta = \theta_0$
- H_1: $\theta = \theta_1$
- An independent random sample is drawn from a distribution with probability mass function $f(x|\theta)$, for $x = 1, 2, \ldots, 7$. The values of the likelihood function at θ_0 and θ_1 are given in the table below.

x	1	2	3	4	5	6	7
$f(x, \theta_0)$	0.01	0.01	0.01	0.01	0.01	0.01	0.94
$f(x, \theta_1)$	0.06	0.05	0.04	0.03	0.02	0.01	0.79

The Neyman–Pearson lemma is used to find the most powerful test for H_0 versus H_1 with significance level $\alpha = 0.04$.

Calculate the probability of Type II error for this test.

A. Less than 0.02
B. At least 0.02, but less than 0.32
C. At least 0.32, but less than 0.62
D. At least 0.62, but less than 0.92
E. At least 0.92

Examination Continued On Next Page

24.

You are given an independent random sample of 10 draws from an exponential distribution with mean θ. The lowest value is excluded, and Y is the expected average value of the remaining 9 draws.

Calculate Y/θ.

A. Less than 1.02
B. At least 1.02, but less than 1.07
C. At least 1.07, but less than 1.12
D. At least 1.12, but less than 1.17
E. At least 1.17

25.

You are given an independent random sample of size 40, $x_1, x_2, \dots, x_{40}$, from an exponential distribution with mean 150.

Calculate the probability that the second smallest observation (the observation with one observation smaller and 38 observations larger in the dataset) is greater than 20.

A. Less than 0.02
B. At least 0.02, but less than 0.04
C. At least 0.04, but less than 0.06
D. At least 0.06, but less than 0.08
E. At least 0.08

Examination Continued On Next Page

26.

You are testing the hypothesis that the median values from two populations, m_1 and m_2 are the same.

- $H_0: m_1 = m_2$
- $H_1: m_1 \neq m_2$
- Sample size $n_1 = 6$
- Sample size $n_2 = 8$
- Wilcoxon rank sum of sample 1: $W_1 = 34$

Calculate the p-value of this test.

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

27.

You are given the results of a matched pairs experiment by policy.

Policy	Sample Loss	Modeled Loss
S	0	5,000
T	0	10,000
U	0	15,000
V	0	20,000
W	0	25,000
X	0	30,000
Y	80,000	35,000
Z	0	50,000

- H_0: $Median_{Sample} = Median_{Modeled}$
- H_1: $Median_{Sample} > Median_{Modeled}$

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

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

28.

You are given the following random sample, $x_1, x_2, \dots, x_7$, of size 7:

4	2	0	0	1	1	a

from the following distributions:

$$f(x|\lambda) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x \in \{0, 1, 2, \dots\}$$

$$f(\lambda) = \frac{\lambda^2}{2}e^{-\lambda}, \quad \lambda > 0, \quad \theta > 0$$

Calculate the value of a such that the Bayesian point estimate for $\lambda = 3$.

A. Less than 2
B. At least 2, but less than 6
C. At least 6, but less than 10
D. At least 10, but less than 14
E. At least 14

Examination Continued On Next Page

29.

You are given the following information:

- Policies are divided into two classes, Safe and Unsafe
- Claims per year per policy follow a Poisson distribution

Classes	Percent of Population	Mean Claims Per Year Per Policy
Safe	40%	0.4
Unsafe	60%	0.8

- A randomly selected policy had 1 claim in the first year.

Calculate the probability that this randomly selected policy is an Unsafe policy.

A. Less than 0.66
B. At least 0.66, but less than 0.68
C. At least 0.68, but less than 0.70
D. At least 0.70, but less than 0.72
E. At least 0.72

30.

For a large portfolio of insurance contracts, ε represents the proportion of policies for which there were claims made within one year.

The random variable ε has a beta prior distribution with parameters a = 10, b = 5, and θ = 1.

In a given year, for 100 randomly selected policies from the portfolio, there were 22 claims.

Calculate the posterior mean of ε.

A. Less than 0.15
B. At least 0.15, but less than 0.25
C. At least 0.25, but less than 0.35
D. At least 0.35, but less than 0.45
E. At least 0.45

Examination Continued On Next Page

31.

Within the context of Generalized Linear Models, suppose that y has an exponential distribution with probability density function expressed as:

$$f(y) = \frac{1}{\mu}\exp\left(-\frac{y}{\mu}\right); \quad \text{for } y > 0$$

Determine the variance of y in terms of μ.

A. $1/\mu$
B. $\sqrt{\mu}$
C. μ
D. μ^2
E. Cannot be determined from the given information

Examination Continued On Next Page

32.

You are given the following GLM output:

Response variable	Pure Premium	
Response distribution	Gamma	
Link	log	
Parameter	**df**	$\hat{\beta}$
Intercept	1	4.78
Risk Group	2	
Group 1	0	0.00
Group 2	1	-0.20
Group 3	1	-0.35
Vehicle Symbol	1	
Symbol 1	0	0.00
Symbol 2	1	0.42

Calculate the predicted pure premium for an insured in Risk Group 2 with Vehicle Symbol 2.

A. Less than 135
B. At least 135, but less than 140
C. At least 140, but less than 145
D. At least 145, but less than 150
E. At least 150

33.

You are given the following two probability density functions:

(i) $f(y;\theta) = \theta y^{-\theta-1}$; for $y > \theta$

(ii) $f(y;\theta) = \theta e^{-y\theta}$; for $y > 0$ and $\theta > 0$

Determine which of the following statements is true.

A. (i) doesn't belong to the exponential family
B. (i) and (ii) both belong to the exponential family
C. Only (i) is in canonical form
D. (i) and (ii) both are in canonical form
E. None of above are true

34.

You are given the following information for a fitted GLM:

Response variable		Status (1 or 0)
Response distribution		Bernoulli
Link		Logit
AIC		28.971
Parameter	**df**	$\hat{\beta}$
Intercept	1	-12.7921
x1	1	1.9104
x2	1	0.1558

- The deviance residual value for the first observation is 0.4902.
- The corresponding hat matrix diagonal value is 0.0505.

Calculate the standardized deviance residual for the first observation.

A. Less than 0
B. At least 0, but less than 0.40
C. At least 0.40, but less than 0.80
D. At least 0.80, but less than 1.20
E. At least 1.20

35.

You are given, $y_1, y_2, \dots, y_n$, independent and Poisson distributed random variables with respective means μ_i for $i = 1,2, \dots, n$.

A Poisson GLM was fitted to the data with a log-link function expressed as:

$$E(y_i) = e^{\beta_0+\beta_1 x_i}$$

where x_i refers to the predictor variable.

Analysis of a set of data provided the following output:

x_i	y_i	$\hat{y}_i$	$y_i \log(y_i/\hat{y}_i)$
0	7	6.0	1.0791
0	9	6.0	3.6492
0	2	6.0	-2.1972
1	3	6.6	-2.3654
1	10	6.6	4.1552
1	8	6.6	1.5390
1	5	6.6	-1.3882
1	7	6.6	0.4119

Calculate the observed deviance for testing the adequacy of the model.

A. Less than 4.0
B. At least 4.0, but less than 6.0
C. At least 6.0, but less than 8.0
D. At least 8.0, but less than 10.0
E. At least 10.0

36.

You are given the following information for a fitted model:

Response variable	Rating
Response distribution	Normal
Link	Identity
Residual Std. Error	7.139

Parameter	df	$\hat{\beta}$
Intercept	1	11.011
Complaints	1	0.692
Privileges	1	-0.104
Learning	1	0.249
Raises	1	-0.033
Critical	1	0.015

- The first record in the data has the following values:

ID	Rating	Complaints	Privileges	Learning	Raises	Critical
1	43	51	30	39	61	92

- The corresponding hat matrix diagonal value is 0.3234.

Calculate the DFITS for the observation above.

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

37.

You are given the outputs from two GLMs fitted to the same data from a trial of a new drug.

Model 1		
Response variable	Number	
Response distribution	Poisson	
Link	Log	
AIC	273.877	
Parameter	$\hat{\beta}$	s. e. $(\hat{\beta})$
Intercept	4.529	0.147
Treatdrug		
Placebo	0.000	0.000
Drug	-1.359	0.118
Age	-0.039	0.006

Model 2		
Response variable	Number	
Response distribution	Negative binomial	
Link	Log	
AIC	164.880	
Parameter	$\hat{\beta}$	s. e. $(\hat{\beta})$
Intercept	4.526	0.595
Treatdrug		
Placebo	0.000	0.000
Drug	-1.368	0.369
Age	-0.039	0.021

Determine which of the following statements is false using the Wald test.

A. Under Model 1, the *Treatdrug* coefficient has a p-value less than 0.01.
B. Under Model 1, the *Age* coefficient has a p-value less than 0.01.
C. Under Model 2, the *Treatdrug* coefficient has a p-value less than 0.01.
D. Under Model 2, the *Age* coefficient has a p-value less than 0.01.
E. Under both models, the intercept coefficient has a p-value of less than 0.01.

38.

You are given the following probability density function for a single random variable, X:

$$f(x|\theta) = \left(\frac{\theta}{2\pi x^3}\right)^{\frac{1}{2}} \exp\left(-\frac{\theta(x-1)^2}{2x}\right)$$

Consider the following statements:

I. $f(x)$ is a member of the exponential family of distributions

II. The score function, $U(\theta)$, is:

$$U(\theta) = \frac{1}{2\theta} - \frac{(x-1)^2}{2x}$$

III. The Fisher Information, $I(\theta)$, is:

$$I(\theta) = 2\theta^2$$

Determine which of the above statements are true.

A. I only
B. II only
C. III only
D. I, II and III
E. The correct answer isn't given by A, B, C or D

Examination Continued On Next Page

39.

You are performing a two-factor analysis of variance with interactions, and are given the following ANOVA table:

Source of variation	Degrees of freedom	Sum of squares
Levels of Factor 1	4	1514.94
Levels of Factor 2	1	240.25
Interactions	4	190.30
Residuals	90	722.30

Calculate the F-statistic for testing the significance of the interaction terms.

A. Less than 2.0
B. At least 2.0, but less than 4.0
C. At least 4.0, but less than 6.0
D. At least 6.0, but less than 8.0
E. At least 8.0

40.

You are given the following information to find the evidence of interaction between two variables, α and β. Multiple models are fitted to a dataset and their results are shown in the following table:

Model	Scaled deviance	Degrees of Freedom
μ	26	9
$\mu + \alpha_j + \beta_k$	6.07	6
$\mu + \alpha_j + \beta_k + (\alpha\beta)_{jk}$	5	4

Calculate the F-statistic for testing the significance of the interaction terms.

A. Less than 0.5
B. At least 0.5, but less than 2.5
C. At least 2.5, but less than 3.0
D. At least 3.0, but less than 3.5
E. At least 3.5

Examination Continued On Next Page

41.

A modeler is considering revising a GLM for claim counts with *age* as an explanatory variable. It is currently being included in the model as a continuous variable with no interactions.

Determine which of the following statements is false.

A. Including *age* as a categorical variable with more than two levels would increase the model degrees of freedom.
B. Including a polynomial term age^2 may decrease model deviance.
C. A plot of *age* against the residuals may be used to assess if a transformation of *age* is necessary.
D. One way of assessing multicollinearity is by performing a regression of *age* against all other explanatory variables.
E. Including several interaction terms involving *age* may make the model more parsimonious.

Examination Continued On Next Page

42.

You are given the following ordered sample of size 6 from a time series:

1	1.5	1.6	1.4	1.5	1.7

Calculate the sample lag 2 autocorrelation.

A. Less than -0.6
B. At least -0.6, but less than -0.3
C. At least -0.3, but less than 0.0
D. At least 0.0, but less than 0.3
E. At least 0.3

43.

An ARMA(3,0) model is fit to the following quarterly time series:

Year	Quarter 1	Quarter 2	Quarter 3	Quarter 4
2013	3.53	1.33	1.85	0.61
2014	0.98	3.61	3.44	3.38
2015	2.91	2.12	4.62	2.93

The estimated coefficients are:

ar1	ar2	ar3	Intercept
0.252	0.061	-0.202	2.637

Forecast the value for Quarter 1 of 2016.

A. Less than 3.00
B. At least 3.00, but less than 3.25
C. At least 3.25, but less than 3.50
D. At least 3.50, but less than 3.75
E. At least 3.75

44.

You are given the AR(3) model below for z_t, a company's revenue for year t:

$$z_t = 5 + 0.85z_{t-1} - 0.02z_{t-3} + w_t$$

where w_t is white noise, with:

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

The revenues for the last 4 years are as follows:

Year	Revenue
2012	20
2013	15
2014	22
2015	19

Forecast the expected revenue for 2017.

A. Less than 19
B. At least 19, but less than 20
C. At least 20, but less than 21
D. At least 21, but less than 22
E. At least 22

45.

You are given the following information:

- x_t is modeled as an ARIMA(1, 0, 0) given by:

$$x_t - 300.53 = 0.95(x_{t-1} - 300.53) + w_t$$

  where w_t is white noise, with:
  - $E(w_t) = 0$
  - $Var(w_t) = \sigma^2$

- $x_{2015} = 310.12$

Calculate the forecast for the time period 2027, x_{2027}.

A. Less than 306.0
B. At least 306.0, but less than 306.4
C. At least 306.4, but less than 306.8
D. At least 306.8, but less than 307.2
E. At least 307.2

End Of Examination

Exam S Fall 2016 Final Answer Key

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