ACTEX Study Manual for CAS Exam S
Statistics and Probabilistic Models

Ambrose Lo, Ph.D., FSA, CERA

Fall 2016 Edition

Preface

Exam S is a new exam which was offered for the first time in Fall 2015 and has formally replaced the previous Exam LC (Models for Life Contingencies), Exam ST (Models for Stochastic Processes and Statistics), the VEE Applied Statistics educational requirement since Fall 2016. The construction of this exam is a first step the CAS takes to incorporate more statistics into their curriculum, as the CAS finds a deep understanding of statistics to be more important for property and casualty actuaries than routine training in traditional actuarial subjects like life contingencies. This is partly driven by the recent popularization of *predictive analytics*, which has stressed the importance of actuary's literacy in statistics. You will considerably sharpen your statistics toolkit as a result of taking (and, in all likelihood, passing!) Exam S.

Syllabus

The syllabus of Exam S, available from `http://www.casact.org/admissions/syllabus/ExamS.pdf`, is extremely broad in scope, covering miscellaneous topics in applied probability, mathematical statistics, generalized linear models and time series analysis, many of which are new topics not tested in any past exams. As a rough estimate, you need at least *three months* of intensive study to master the material in this exam. The specific sections of the syllabus along with their approximate weights in the exam are shown below:

Section	**Range of Weight**
A. Probability Models (Stochastic Processes and Survival Models)	20–40%
B. Statistics	20–40%
C. Extended Linear Models	25–40%
D. Time Series with Constant Variance	5–10%

Section B on mathematical statistics is more or less taken intact from the syllabuses of Exam ST and its predecessor, Exam 3L. As a result, you can find lots of relevant past exam questions on this section. Part of Section A was also tested in Exam ST and Exam 3L, but an additional reading for Poisson process and several new topics are added. Sections C and D are, to a large extent, new topics.

Exam Format

Exam S is a four-hour multiple-choice exam consisting of approximately 40 to 50 multiple-choice questions. Before 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. However, writing is not permitted.

Each of the Fall 2015 and Spring 2016 exams has 45 questions, categorized into the four sections as follows:

Section	**Number of Questions**	
	2015F	2016S
A. Probability Models (Stochastic Processes and Survival Models)	13	13
B. Statistics	17	14
C. Extended Linear Models	11	14
D. Time Series with Constant Variance	4	3
Total	45	45

Such a distribution of questions is more or less consistent with the distribution shown on the last page. For a formal statistical test to justify such a claim, please try Practice Exam 2 Question 22 on page 1020. Most of the questions in the Fall 2015 exam are pretty straightforward, covering very fundamental concepts listed on the syllabus knowledge statements, as is often the case when a new exam is offered for the first time. The The pass mark for Fall 2015 is **52.50**, which means that candidates needed to answer about **26 out of 44 questions** correctly to earn a pass (each question carries 2 points). Spring 2016 exam is slightly harder. You should expect that the exam questions in later sittings will be more challenging.

Most answer choices are in the form of ranges, e.g.:

A. Less than 1%

B. At least 1%, but less than 2%

C. At least 2%, but less than 3%

D. At least 3%, but less than 4%

E. At least 4%

If your answer is *much* lower than the bound indicated by Answer A or *much* higher than that suggested by Answer E, do check your calculations. Chances are that you have made computational mistakes, but this is not definitely the case.

A guessing adjustment will be in place, so unless you can eliminate two or three of the answer choices, it will be wise of you not to answer questions which you are unsure of by pure guesswork.

What is Special about This Study Manual?

The objective of this study manual is to help you *grasp the material* in this brand-new Exam S effectively and efficiently, and *pass it with considerable ease.* Here are some of the valuable features of the study manual for achieving this all-important goal:

- Each chapter and section starts by explicitly stating which learning outcomes and knowledge statements of the exam syllabus we are going to cover, to let you that we are on track and hitting the right target.
- The knowledge statements of the syllabus are demystified by precise and concise expositions in the main text of the manual, helping you acquire a deep understanding of the subject matter.
- Intuitions and mnemonics are emphasized, so are highlights of important exam items and common mistakes committed by students.
- Formulae and results of utmost importance are $\boxed{\text{boxed}}$ for easy identification and memorization.
- The expositions are complemented by more than 550 in-text examples and 550 end-of-chapter problems, which are original or taken from required textbooks and relevant past exams, all with step-by-step solutions and problem-solving remarks, to give you a sense of what you can expect to see in the real exam. As you read this manual, skills are honed and confidence is built.
- Two full-length practice exam conclude this study manual giving you a holistic review of the syllabus material.

New to the Fall 2016 Edition

- Old SOA/CAS exam questions before 2000 are added as appropriate. Despite their seniority, these old exam questions, which are not easily available nowadays, do illustrate some less commonly tested concepts and have considerable value.

- The subtle connections between seemingly disparate topics in the syllabus are unraveled. For example, knowledge of order statistics (Section B of the syllabus) leads to significant shortcuts for calculating the expected lifetime of a system with uniformly distributed components (Section A of the syllabus).
- The material on the conditional distributions of waiting times of Poisson processes is refined and moved to the section on order statistics (Section 7.1), where you will acquire working knowledge of order statistics and better understand how the waiting times are distributed conditionally.
- The expositions of Section C (generalized linear models) are radically rewritten in view of the changes in the required readings. Redundant topics are deleted and new ones are added.
- All known typographical errors are fixed.

Exam Tables

In the exam, you will be supplied with a variety of tables, including:

- *Standard normal distribution table* (used throughout this study manual)

  You will need this table for values of the standard normal distribution function or standard normal quantiles, when you work with normally distributed random variables or perform normal approximation.

- *Illustrative Life Table* (used mostly in Chapter 4 of this study manual)

  You will need this when you are told that mortality of the underlying population follows the Illustrative Life Table.

- *A table of distributions for a number of common continuous and discrete distributions and the formulae for their moments and other probabilistic quantities* (used throughout Parts I and II of this study manual)

  This big table provides a great deal of information about some common as well as non-common distributions (e.g. inverse exponential, Pareto, etc.). When an exam question centers on these distributions and quantities such as their means or variances are needed, consult this table.

- *Quantiles of t-distribution, F-distribution, chi-square distribution* (used in Chapters 6, 9, 10 and 11 of this study manual)

  These quantiles will be of use when you perform parametric hypothesis tests.

- *Critical values of the signed-rank Wilcoxon test and Mann-Whitney U-test* (used in Chapter 7 of this study manual)

  These quantiles will be needed when you perform nonparametric hypothesis tests.

You should download these tables from `http://www.casact.org/admissions/syllabus/S_tables.pdf`, print out a copy and learn how to locate the relevant entries in these tables because they will be intensively used during your study as well as in the exam.

Acknowledgment

I would like to thank my colleague, Professor Elias S. W. Shiu, at the University of Iowa for sharing with me many old pre-2000 SOA/CAS exam papers. These hard-earned old exam papers have been invaluable.

Errata

While we go to great lengths to polish and proofread this manual, some mistakes will inevitably go unnoticed. For this reason, we would be extremely grateful if you could *send any errors you identify* or *any suggestions and comments you have about this manual to* ***ambrose-lo@uiowa.edu*** *and c.c.* ***support@actexmadriver.com***. The author will try his best to respond to any inquiries within 48 hours and an ongoing errata list will be maintained online. More importantly, *students who report errors will quality for a quarterly drawing for $100 in store credit.*

Ambrose Lo
July 2016

About the Author

Professor Ambrose Lo earned his B.S. in Actuarial Science (first class honors) and Ph.D. in Actuarial Science from The University of Hong Kong in 2010 and 2014 respectively. He joined the Department of Statistics and Actuarial Science at The University of Iowa in August 2014 as an assistant professor in actuarial science. He is a Fellow of the Society of Actuaries (FSA) and a Chartered Enterprise Risk Analyst (CERA). His research interests lie in dependence structures, quantitative risk management as well as optimal (re)insurance. His research papers have been published in top-tier actuarial journals, such as *Insurance: Mathematics and Economics* and the *Scandinavian Actuarial Journal*.

Besides dedicating himself to actuarial research, Ambrose attaches equal importance to teaching, through which he nurtures the next generation of actuaries and serves the actuarial profession. He has taught courses on financial derivatives, mathematical finance, life contingencies, credibility theory and advanced probability theory. His emphasis in teaching is always placed on concrete problem-solving skills complemented by a thorough understanding of the subject matter. As a result of his exceptional teaching performance, Ambrose has won numerous teaching awards ever since he was a graduate student (see `http://www.scifac.hku.hk/news/comm/award-excellence-teaching-assistant-2011-12-2` for example).

Contents

Part I

Probability Models (Stochastic Processes and Survival Models)

Chapter 1

Poisson Processes

Chapter overview: As a prospective P&C actuary, you would be interested in monitoring the number of insurance claims an insurance company receives as time goes by and how these claims can be appropriately analyzed by means of sound statistical analysis. In Exam S, we shall learn one way of modeling the flow of insurance claims – the Poisson process.

This part of the syllabus has two required readings:

(1) *A study note by J.W. Daniel*

The study note is precise and concise, introducing main results mostly without proof and supplementing its exposition with a few simple examples. It is suitable for a first-time introduction to Poisson processes.

(2) *The book entitled Introduction to Probability Models by S.M. Ross.*

This is a textbook used by a number of college courses on elementary applied probability. It balances rigor and intuition, and presents the theory of Poisson processes at a level that is much deeper than that in the study note by Daniel. In particular, it treats the conditional distribution of the arrival times as well as the interplay between two independent Poisson processes. A conspicuous feature of this book is its large number of sophisticated examples and exercises which require a large amount of ingenuity and cannot be done in a reasonable exam setting. This study manual improves the practicality of the book and rephrases these otherwise intractable examples in an exam tone.

You can download this book (Eleventh Edition) legally from *ScienceDirect* via the following link, chapter by chapter, if your university has subscribed to it:

`http://www.sciencedirect.com/science/book/9780124079489`

You should do so because the book has a number of good exercises whose solutions will be discussed in this study manual (the questions cannot be reproduced here because of copyright issues).

The study note has been on the syllabuses of Exams 3, 3L and ST, whereas *Introduction to Probability Models* has just entered the syllabus of Exam S. As a result of the addition of the latter reading, we expect more complex exam questions on Poisson processes in Exam S. In total, expect about 5 questions on the material of the entire chapter.

1.1 Fundamental Properties

LEARNING OBJECTIVES

1. Understand and apply the properties of Poisson processes:
   - For increments in the homogeneous case
   - For increments in the non-homogeneous case

KNOWLEDGE STATEMENTS

a. Poisson process

b. Non-homogeneous Poisson process

Definition. By definition, a *Poisson process* $\{N(t),\, t \geq 0\}$ with *rate function* (also known as *intensity function*) $\lambda(\cdot)$ [i] is a stochastic process, namely, a collection of random variables indexed by time t (in an appropriate unit, e.g. minute, hour, month, year, etc.), satisfying the following properties:

1. *(Counting)* $N(0) = 0$, $N(t)$ is non-decreasing in t and takes non-negative integer values only.

   Interpretation: $N(t)$ counts the number of claims which are submitted on or before time t. Thus $N(0)$ is 0 (we assume that should be no claims before the insurance company starts its business), $N(t)$ cannot decrease in time and must be integer-valued.

2. *(Distribution of increments are Poisson random variables)* For $s < t$, the increment $N(t) - N(s)$, which counts the number of events in the interval $(s, t]$, is a Poisson random variable with mean $\Lambda = \int_s^t \lambda(y)\, \mathrm{d}y$.

   Interpretation: Increments of a Poisson process, as its name suggests, are Poisson random variables with mean computed by integrating the rate function over the same interval. In this regard, we can see that the rate function of a Poisson process completely specifies the distribution of each increment.

3. *(Increments are independent)* If $(s_1, t_1]$ and $(s_2, t_2]$ are non-overlapping intervals, then $N(t_1) - N(s_1)$ and $N(t_2) - N(s_2)$ are independent random variables.

[i] The study note by Daniel simply writes a Poisson process as N in short. While this is a perfectly correct way of writing, some students may confuse that with a Poisson random variable N. Also, here we write $\lambda(\cdot)$ with a parenthesis containing the argument of the function instead of λ to emphasize that $\lambda(\cdot)$ is a function.

Interpretation: This is the most amazing property of a Poisson process. Its increments not only follow Poisson distribution, but also are independent on disjoint intervals (e.g. $(0, 1)$ and $(2, 5)$ are disjoint intervals, so are $(3, 4]$ and $(4, 5]$). This means that, in this model, the frequency of claims you received last month has nothing to do with the frequency this month.

In the context of insurance applications, we interpret $N(t)$ as the number of claims that occur on or before time t. The same interpretation can easily carry over to more general contexts where we are interested in counting a particular type of event, e.g. the number of customers that enter a store, the number of cars passing through an intersection, the number of lucky candidates passing Exam S, etc.

A special Poisson process whose rate function is constant, say $\lambda(t) = \lambda$ for all $t \geq 0$, is called a *homogeneous* Poisson process. In addition to having independent increments, a homogeneous Poisson process also possesses *stationary increments*, meaning that the distribution of $N(t + s) - N(s)$ depends on the length of the interval, which is t in this case, but not on s.

Probability calculations. The second and third properties of a Poisson process allow us to calculate many probabilistic quantities, such as the probability of a certain number of events, as well as the expected and variance of the number of events in a particular time interval. These two properties will be intensively used in exam questions. The following string of past exam questions serves as excellent illustrations.

RECALL

Just in case you forgot:

1. The probability mass function of a Poisson random variable X with parameter λ (a scalar, not a function) is given by

$$\Pr(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0, 1, 2, \ldots$$

   The mean and variance of X are both equal to λ.

2. If $X_1, X_2, \ldots, X_n$ are independent Poisson random variables with respective means $\lambda_1, \lambda_2, \ldots, \lambda_n$, then $X_1 + X_2 + \cdots + X_n$ is also a Poisson random variable with a mean of $\lambda_1 + \lambda_2 + \cdots + \lambda_n$. In other words, the sum of independent Poisson random variables is also a Poisson random variable whose mean is the sum of the individual Poisson means.

Example 1.1.1. (SOA/CAS Exam P/1 Sample Question 173: Warm-up question) In a given region, the number of tornadoes in a one-week period is modeled by a Poisson distribution with mean 2. The numbers of tornadoes in different weeks are mutually independent.

Calculate the probability that fewer than four tornadoes occur in a three-week period.

A. 0.13

B. 0.15

C. 0.29

D. 0.43

E. 0.86

Ambrose's comments: This is not even a Poisson process question. It simply reminds you of how probabilities for a Poisson random variable are typically calculated.

Solution. We are interested in $\Pr(N_1 + N_2 + N_3 < 4)$, where N_i is the number of tornadoes in the i^{th} week for $i = 1, 2, 3$. As $N_1 + N_2 + N_3$ is also a Poisson random variable with a mean of $3(2) = 6$, we have

$$\begin{aligned}\Pr(N_1 + N_2 + N_3 < 4) &= \sum_{i=0}^{3} \Pr(N_1 + N_2 + N_3 = i) \\ &= \mathrm{e}^{-6}\left(1 + 6 + \frac{6^2}{2!} + \frac{6^3}{3!}\right) \\ &= \boxed{0.1512}. \quad \textbf{(Answer: B)}\end{aligned}$$

□

Example 1.1.2. (CAS Exam 3L Fall 2013 Question 9: Probability I) You are given that claim counts follow a non-homogeneous Poisson Process with $\lambda(t) = 30t^2 + t^3$.

Calculate the probability of at least two claims between time 0.2 and 0.3.

A. Less than 1%

B. At least 1%, but less than 2%

C. At least 2%, but less than 3%

D. At least 3%, but less than 4%

E. At least 4%

Solution. The number of claims between time 0.2 and 0.3 is a Poisson random variable with parameter

$$\int_{0.2}^{0.3} (30t^2 + t^3)\,\mathrm{d}t = 10t^3 + \frac{t^4}{4}\bigg|_{0.2}^{0.3} = 0.191625.$$

Hence the probability of at least two claims between time 0.2 and 0.3 is the complement of the probability of having 0 or 1 claim:

$$1 - \Pr(0 \text{ claim}) - \Pr(1 \text{ claim}) = 1 - \mathrm{e}^{-0.191625}(1 + 0.191625) = \boxed{0.0162}.$$ **(Answer: B)**

□

Example 1.1.3. (CAS Exam 3L Spring 2010 Question 12: Probability II) Downloads of a song on a musician's Web site follow a heterogeneous Poisson process with the following Poisson rate function:

$$\lambda(t) = \mathrm{e}^{-0.25t}$$

Calculate the probability that there will be more than two downloads of this song between times $t = 1$ and $t = 5$.

A. Less than 29%

B. At least 29%, but less than 30%

C. At least 30%, but less than 31%

D. At least 31%, but less than 32%

E. At least 32%

Solution. Because $N(5) - N(1)$ is a Poisson random variable with parameter

$$\int_1^5 \lambda(t)\,\mathrm{d}t = \int_1^5 \mathrm{e}^{-0.25t}\,\mathrm{d}t = \frac{\mathrm{e}^{-0.25(1)} - \mathrm{e}^{-0.25(5)}}{0.25} = 1.969184,$$

the required probability equals

$$\begin{aligned} & \Pr(N(5) - N(1) > 2) \\ = \; & 1 - \Pr(N(5) - N(1) = 0) - \Pr(N(5) - N(1) = 1) - \Pr(N(5) - N(1) = 2) \\ = \; & 1 - e^{-1.969184}\left(1 + 1.969184 + \frac{1.969184^2}{2}\right) \\ = \; & \boxed{0.3150}. \quad \textbf{(Answer: D)} \end{aligned}$$

□

Example 1.1.4. (CAS Exam S Spring 2016 Question 3: Probability III) You are given:

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

$$\begin{aligned} \lambda(t) &= \frac{1}{2}t, \quad 0 < t < 5 \\ \lambda(t) &= \frac{1}{4}t, \quad t \geq 5 \end{aligned}$$

- 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

Solution. The number of claims between $t = 4$ and $t = 6$ is a Poisson random variable with mean

$$\int_4^5 \frac{1}{2}t \, dt + \int_5^6 \frac{1}{4}t \, dt = \frac{5^2 - 4^2}{2(2)} + \frac{6^2 - 5^2}{4(2)} = 3.625.$$

The probability of having exactly one claim between $t = 4$ and $t = 6$ is

$$3.625e^{-3.625} = \boxed{0.0966}. \quad \textbf{(Answer: B)}$$

□

Example 1.1.5. (CAS Exam 3L Spring 2012 Question 9: Expected value) Claims reported for a group of policies follow a non-homogeneous Poisson process with rate function:

$$\lambda(t) = 100/(1+t)^3, \text{ where } t \text{ is the time (in years) after January 1, 2011.}$$

Calculate the expected number of claims reported after January 1, 2011 for this group of policies.

A. Less than 45

B. At least 45, but less than 55

C. At least 55, but less than 65

D. At least 65, but less than 75

E. At least 75

Solution. We are interested in $N(\infty) = \lim_{t\to\infty} N(t)$, which is a Poisson random variable with mean

$$\int_0^\infty \lambda(t)\,\mathrm{d}t = \int_0^\infty \frac{100}{(1+t)^3}\,\mathrm{d}t = 100\left[-\frac{1}{2(1+t)^2}\right]_0^\infty = \boxed{50}. \quad \textbf{(Answer: B)}$$

□

Example 1.1.6. (CAS Exam 3L Spring 2013 Question 9: Variance) You are given the following:

- An actuary takes a vacation where he will not have access to email for eight days.
- While he is away, emails arrive in the actuary's inbox following a non-homogeneous Poisson process where

$$\lambda(t) = 8t - t^2 \text{ for } 0 \le t \le 8. \,(t \text{ is in days})$$

Calculate the variance of the number of emails received by the actuary during this trip.

A. Less than 60

B. At least 60, but less than 70

C. At least 70, but less than 80

D. At least 80, but less than 90

E. At least 90

Solution. The trip of the actuary lasts for 8 days, during which the number of emails is a Poisson random variable with variance (same as the Poisson parameter)

$$\int_0^8 (8t - t^2)\,\mathrm{d}t = 4t^2 - \frac{t^3}{3}\bigg|_0^8 = \boxed{85.3333}. \quad \textbf{(Answer: D)}$$

□

Example 1.1.7. (CAS Exam ST Fall 2015 Question 1: Calculation of homogeneous Poisson intensity) For two Poisson processes, N_1 and N_2, you are given:

- N_1 has intensity function $\lambda_1(t) = \begin{cases} 2t & \text{for } 0 < t \leq 1 \\ t^3 & \text{for } t > 1 \end{cases}$
- N_2 is a homogeneous Poisson process.
- $\mathrm{Var}[N_1(3)] = 4\mathrm{Var}[N_2(3)]$

Calculate the intensity of N_2 at $t = 3$.

A. Less than 1

B. At least 1, but less than 3

C. At least 3, but less than 5

D. At least 5, but less than 7

E. At least 7

Solution. Note that $N_1(3)$ has a mean and variance equal to

$$\int_0^3 \lambda_1(t)\,\mathrm{d}t = \int_0^1 2t\,\mathrm{d}t + \int_1^3 t^3\,\mathrm{d}t = [t^2]_0^1 + \left[\frac{t^4}{4}\right]_1^3 = 1 + \frac{3^4 - 1^4}{4} = 21,$$

while $N_2(3)$ has a mean and variance equal to $3\lambda_2$, where λ_2 is the constant intensity of N_2. As $\mathrm{Var}[N_1(3)] = 4\mathrm{Var}[N_2(3)]$, we have $21 = 4(3\lambda_2)$, so $\lambda_2 = 21/12 = \boxed{1.75}$. **(Answer: B)** □

Probabilities involving overlapping intervals. A harder exam question may ask that you determine probabilities for increments on overlapping intervals. The key step to calculate these probabilities lies in rewriting the events in terms of increments on *non-overlapping* intervals, which are independent according to the definition of a Poisson process.

Example 1.1.8. (Probability for overlapping increments I) The number of calls received in a telephone exchange follow a homogeneous Poisson process with a rate of 30 per hour.

Calculate the probability that there are exactly 2 calls in the first ten minutes and exactly 5 calls in the first twenty minutes.

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

Solution. When time is measured in hours, the required probability is

$$\Pr(N(1/6) = 2, N(1/3) = 5) = \Pr(N(1/6) = 2, N(1/3) - N(1/6) = 3),$$

which can be factored, because of independence, into

$$\begin{aligned} \Pr(N(1/6) = 2)\Pr(N(1/3) - N(1/6) = 3) &= \frac{\mathrm{e}^{-30/6}(30/6)^2}{2!} \times \frac{\mathrm{e}^{-30/6}(30/6)^3}{3!} \\ &= \boxed{0.0118}. \quad \textbf{(Answer: B)} \end{aligned}$$

□

Example 1.1.9. (Probability for overlapping increments II) Customers arrive at a post office in accordance with a Poisson process with a rate of 5 per hour. The post office opens at 9:00 am.

Calculate the probability that only one customer arrives before 9:20 am and ten customers arrive before 11:20 am.

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

Solution. The required probability is

$$\begin{aligned}\Pr(N(1/3)=1, N(7/3)=10) &= \Pr(N(1/3)=1, N(7/3)-N(1/3)=9)\\ &= \Pr(N(1/3)=1)\Pr(N(7/3)-N(1/3)=9)\\ &= e^{-5/3}\left(\frac{5}{3}\right)\times e^{-5(2)}\frac{[5(2)]^9}{9!}\\ &= \boxed{0.0394}. \quad \textbf{(Answer: D)}\end{aligned}$$

□

Conditional distribution of $N(t)$ given $N(s)$ for $s \le t$. Suppose that we know the value of the Poisson process at one time point s with $N(s)=m$, and we wish to study the probabilistic behavior of the Poisson process at a later time point t with $s \le t$. Then $N(t)$ turns out to be a *translated Poisson* random variable in the sense that it has the same distribution as the sum of a Poisson random variable and a constant. To see this, let's write $N(t)$ as

$$N(t) = [N(t)-N(s)] + N(s).$$

The second term $N(s)$ is known to be m, while the first term, owing to the property of independent increments of a Poisson process, is a Poisson random variable, say M, whose distribution does not depend on the value of m. Therefore, we have the distributional representation

$$\boxed{[N(t)|N(s)=m] \overset{d}{=} M+m, \quad s\le t,}$$

where "$\overset{d}{=}$" means equality in distribution. This result allows us to answer questions about many probabilistic quantities associated with $N(t)$ when the value of $N(s)$ is given.

Example 1.1.10. (CAS Exam 3L Fall 2010 Question 11: Conditional variance) You are given the following information:

- A Poisson process N has a rate function $\lambda(t)=3t^2$.
- You have observed 50 events by time $t=2.1$.

Calculate $\mathrm{Var}[N(3) \mid N(2.1)=50]$.

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

Solution. Conditional on $N(2.1) = 50$, $N(3)$ has the same distribution as $M + 50$, where M is a Poisson random variable with mean and variance

$$\int_{2.1}^{3} \lambda(t)\,\mathrm{d}t = \int_{2.1}^{3} 3t^2\,\mathrm{d}t = t^3\Big|_{2.1}^{3} = 17.739.$$

Hence

$$\mathrm{Var}[N(3) \mid N(2.1) = 50] = \mathrm{Var}(M + 50) = \mathrm{Var}(M) = \boxed{17.739} \quad \textbf{(Answer: B)}.$$

□

Normal approximation: For more cumbersome probabilities such as $\Pr(N(t) > c)$ with c being a large number, exact calculations can be tedious and normal approximation may be used. That is, we approximate $N(t)$ by a normal random variable with the same mean and variance, and instead calculate the probability of the same event for this normal random variable using the standard normal distribution table you have in the exam. Because the distribution of $N(t)$ is discrete, and a continuous distribution (i.e. normal) is used to approximate this discrete distribution, a continuity correction should be made.

Recall - Continuity correction

Let X be a random variable taking values in the set of integers $\{0, \pm 1, \pm 2, \ldots\}$ and N is a normal random variable having the same mean and variance as X. The following shows how various probabilities are approximated using the normal approximation *with* continuity correction: (c is an integer)

Probability of Interest		Approximant
$\Pr(X \leq c)$	$\approx$	$\Pr(N \leq c + 0.5)$
$\Pr(X < c)$	$\approx$	$\Pr(N \leq c - 0.5)$
$\Pr(X \geq c)$	$\approx$	$\Pr(N \geq c - 0.5)$
$\Pr(X > c)$	$\approx$	$\Pr(N > c + 0.5)$

In the second column, it does not matter whether we take sharp or strict inequalities because N is a continuous random variable. In other words, we may replace "$\leq$" by "$<$" and "$\geq$" by "$>$".

Throughout this study manual, we denote the distribution function of the standard normal distribution by Φ.

Example 1.1.11. (CAS Exam 3L Spring 2008 Question 11: Normal approximation) A customer service call center operates from 9:00 AM to 5:00 PM. The number of calls received by the call center follows a Poisson process whose rate function varies according to the time of day, as follows:

Time of Day	Call Rate (per hour)
9:00 AM to 12:00 PM	30
12:00 PM to 1 :00 PM	10
1:00 PM to 3:00 PM	25
3:00 PM to 5:00 PM	30

Using a normal approximation, what is the probability that the number of calls received from 9:00AM to 1:00PM exceeds the number of calls received from 1:00PM to 5:00PM?

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%

Solution. The number of calls received from 9:00AM to 1:00PM is a Poisson random variable N^1 with parameter $30(3) + 10(1) = 100$, while the number of calls received from 1:00PM to 5:00PM is a Poisson random variable N^2 with parameter $25(2) + 30(2) = 110$. Because N^1 and N^2 are independent,

$$\mathrm{E}[N^1 - N^2] = \mathrm{E}[N^1] - \mathrm{E}[N^2] = 100 - 110 = -10,$$

and

$$\mathrm{Var}(N^1 - N^2) = \mathrm{Var}(N^1) + \mathrm{Var}(N^2) = 100 + 110 = 210.$$

Using the normal approximation with continuity correction, we have

$$\Pr(N^1 > N^2) = \Pr(N^1 - N^2 > 0) \approx \Pr\left(\underbrace{\mathrm{N}(-10, 210)}_{\substack{\text{a normal r.v. with mean} \\ -10 \text{ and variance } 210}} > 0.5\right),$$

which, upon standardization, equals

$$\Pr\left(\mathrm{N}(0,1) > \frac{0.5 - (-10)}{\sqrt{210}}\right) = 1 - \Phi(0.72) = 1 - 0.7642 = \boxed{0.2358}. \quad \textbf{(Answer: C)}$$

□

Remark. If you do not use continuity correction, you will calculate

$$\begin{aligned} \Pr(N^1 > N^2) &\approx \Pr\left(\mathrm{N}(-10, 210) > 0\right) \\ &= 1 - \Phi\left(\frac{0 - (-10)}{\sqrt{210}}\right) \\ &= 1 - \Phi(0.69) \\ &= 1 - 0.7549 \\ &= 0.2451, \end{aligned}$$

in which case you will also end up with Answer C.

[Harder!] Conditional distribution of $N(s)$ given $N(t)$ with $s \le t$. We know that conditional on $N(s)$, the distribution of $N(t)$, where $0 \le s \le t$, is a translated Poisson random variable. What about the conditional distribution of $N(s)$ given $N(t)$? We answer this question for a *homogeneous* Poisson process, whose rate function is constant at λ, and consider, for $0 \leq k \leq n$,

$$\begin{aligned}\Pr(N(s)=k|N(t)=n) &= \frac{\Pr(N(s)=k, N(t)=n)}{\Pr(N(t)=n)} \\ &= \frac{\Pr(N(s)=k, N(t)-N(s)=n-k)}{\Pr(N(t)=n)}.\end{aligned}$$

Because a Poisson process possesses independent increments, we further have

$$\begin{aligned}\Pr(N(s)=k|N(t)=n) &= \frac{\Pr(N(s)=k)\Pr(N(t)-N(s)=n-k)}{\Pr(N(t)=n)} \\ &= \frac{\frac{e^{-\lambda s}(\lambda s)^k}{k!} \times \frac{e^{-\lambda(t-s)}[\lambda(t-s)]^{n-k}}{(n-k)!}}{\frac{e^{-\lambda t}(\lambda t)^n}{n!}} \\ &= \frac{n!}{k!(n-k)!}\left(\frac{s}{t}\right)^k\left(1-\frac{s}{t}\right)^{n-k} \\ &= \binom{n}{k}\left(\frac{s}{t}\right)^k\left(1-\frac{s}{t}\right)^{n-k}.\end{aligned}$$

In other words, given $N(t)=n$, $N(s)$ is a $\boxed{\text{binomial}}$ random variable with parameters n and s/t.

Example 1.1.12. (CAS Exam ST Spring 2016 Question 1: Probability for $N(s)$ given $N(t)$) You are given that $N(t)$ follows the Poisson process with rate $\lambda = 2$. Calculate $\Pr[N(2)=3|N(5)=7]$.

A. Less than 0.25

B. At least 0.25, but less than 0.35

C. At least 0.35, but less than 0.45

D. At least 0.45, but less than 0.55

E. At least 0.55

Solution. Conditional on $N(5)=7$, $N(2)$ is a binomial random variable with parameters 7 and 2/5, so

$$\Pr[N(2)=3|N(5)=7] = \binom{7}{3}\left(\frac{2}{5}\right)^3\left(1-\frac{2}{5}\right)^4 = \boxed{0.290304}. \quad \textbf{(Answer: B)}$$

□

Remark. We do not need the value of λ.

Problems

Problem 1.1.1. (CAS Exam 3 Fall 2006 Question 26: True-of-false questions) Which of the following is/are true?

1. A counting process is said to possess independent increments if the number of events that occur between time s and t is independent of the number of events that occur between time s and $t+u$ for all $u > 0$.
2. All Poisson processes have stationary and independent increments.
3. The assumption of stationary and independent increments is essentially equivalent to asserting that at any point in time the process probabilistically restarts itself.

A. 1 only

B. 2 only

C. 3 only

D. 1 and 2 only

E. 2 and 3 only

Solution. Only 3. is correct. **(Answer: C)**

1. This would be true if the second s is changed to t.
2. A non-homogeneous Poisson process does not have stationary increments in general.

□

Problem 1.1.2. (CAS Exam 3 Fall 2006 Question 28: Piecewise linear intensity function) Customers arrive to buy lemonade according to a Poisson distribution with $\lambda(t)$, where t is time in hours, as follows:

$$\lambda(t) = \begin{cases} 2+6t & 0 \le t \le 3 \\ 20 & 3 < t \le 4 \\ 36-4t & 4 < t \le 8 \end{cases}$$

At 9:00 a.m., t is 0.

Calculate the number of customers expected to arrive between 10:00 a.m. and 2:00 p.m.

A. Less than 63

B. At least 63, but less than 65

C. At least 65, but less than 67

D. At least 67, but less than 69

E. At least 69

Solution. The expected number of customers arriving between 10:00 a.m. ($t = 1$) and 2:00 p.m. ($t = 5$) is

$$\begin{aligned}\int_1^5 \lambda(t)\,\mathrm{d}t &= \int_1^3 (2+6t)\,\mathrm{d}t + \int_3^4 20\,\mathrm{d}t + \int_4^5 (36-4t)\,\mathrm{d}t \\ &= [2t+3t^2]_1^3 + 20 + [36t - 2t^2]_4^5 \\ &= \boxed{66}. \quad \textbf{(Answer: C)}\end{aligned}$$

□

Remark. Because the intensity function is piecewise linear, integrating it is the same as calculating the areas of trapeziums.

Problem 1.1.3. (CAS Exam 3L Fall 2008 Question 1: Expected value) The number of accidents on a highway from 3:00 PM to 7:00 PM follows a nonhomogeneous Poisson process with rate function

$$\lambda = 4 - (t-2)^2, \text{ where } t \text{ is the number of hours since 3:00 PM.}$$

How many more accidents are expected from 4:00 PM to 5:00 PM than from 3:00 PM to 4:00PM?

A. Less than 0.75

B. At least 0.75, but less than 1.25

C. At least 1.25, but less than 1.75

D. At least 1.75, but less than 2.25

E. At least 2.25

Solution.

- The expected number of accidents from 3:00 PM to 4:00PM is

$$\int_0^1 [4-(t-2)^2]\,\mathrm{d}t = \left[4t - \frac{(t-2)^3}{3}\right]_0^1 = \frac{5}{3}.$$

- The expected number of accidents from 4:00 PM to 5:00 PM is

$$\int_1^2 [4-(t-2)^2]\,\mathrm{d}t = \left[4t - \frac{(t-2)^3}{3}\right]_1^2 = \frac{11}{3}.$$

The difference is $\boxed{2}$. **(Answer: D)** □

Problem 1.1.4. (CAS Exam 3L Fall 2008 Question 2: Probability, homogeneous) You are given the following:

- Hurricanes occur at a Poisson rate of 1/4 per week during the hurricane season.
- The hurricane season lasts for exactly 15 weeks.

Prior to the next hurricane season, a weather forecaster makes the statement, "There will be at least three and no more than five hurricanes in the upcoming hurricane season."
Calculate the probability that this statement will be correct.

A. Less than 54%

B. At least 54%, but less than 56%

C. At least 56%, but less than 58%

D. At least 58%, but less than 60%

E. At least 60%

Solution. Note that $N(15)$, the number of hurricanes during the 15-week hurricane season, is a Poisson random variable with a mean of $15/4 = 3.75$. The probability that the statement will be correct is

$$\Pr(3 \le N(15) \le 5) = \mathrm{e}^{-3.75}\left(\frac{3.75^3}{3!} + \frac{3.75^4}{4!} + \frac{3.75^5}{5!}\right) = \boxed{0.5458}. \quad \textbf{(Answer: B)}$$

□

Problem 1.1.5. (CAS Exam 3L Spring 2008 Question 10: Probability, nonhomogeneous) Car accidents follow a Poisson process, as described below:

- On Monday and Friday, the expected number of accidents per day is 3.
- On Tuesday, Wednesday, and Thursday, the expected number of accidents per day is 4.
- On Saturday and Sunday, the expected number of accidents per day is 1.

Calculate the probability that exactly 18 accidents occur in a week.

A. Less than .06

B. At least .06 but less than .07

C. At least .07 but less than .08

D. At least .08 but less than .09

E. At least .09

Solution. The total number of accidents in a week is a Poisson random variable with a mean of $3(2) + 4(3) + 1(2) = 20$, so the probability of having exactly 18 accidents in a week is

$$\frac{e^{-20}20^{18}}{18!} = \boxed{0.0844}. \quad \textbf{(Answer: D)}$$

□

Problem 1.1.6. (CAS Exam 3 Spring 2006 Question 33: Probability, non-homogeneous) While on vacation, an actuarial student sets out to photograph a Jackalope and a Snipe, two animals common to the local area. A tourist information booth informs the student that daily sightings of Jackalopes and Snipes follow independent Poisson processes with intensity parameters:

$$\begin{aligned} \lambda_J(t) &= \frac{t^{1/3}}{5} \qquad \text{for Jackalopes} \\ \lambda_S(t) &= \frac{t^{1/2}}{10} \qquad \text{for Snipes} \end{aligned}$$

where: $0 \le t \le 24$ and t is the number of hours past midnight

If the student takes photographs between 1 pm and 5 pm, calculate the probability that he will take at least 1 photograph of each animal.

A. Less than 0.45

B. At least 0.45, but less than 0.60

C. At least 0.60, but less than 0.75

D. At least 0.75, but less than 0.90

E. At least 0.90

Solution. The number of Jackalopes and Snipes between 1 pm and 5 pm are Poisson random variables with respective means

$$\frac{1}{5}\int_{13}^{17} t^{1/3}\,dt = \frac{3}{4(5)}(17^{4/3} - 13^{4/3}) = 1.971665$$

and

$$\frac{1}{10}\int_{13}^{17} t^{1/2}\,dt = \frac{2}{3(10)}(17^{3/2} - 13^{3/2}) = 1.548042.$$

Because the two Poisson processes are independent (note: here we are not using the property of independent increments),

$$\begin{aligned}
& \Pr\left(N_J(17)-N_J(13)\geq 1, N_S(17)-N_S(13)\geq 1\right)\\
= & \Pr\left(N_J(17)-N_J(13)\geq 1\right)\Pr\left(N_S(17)-N_S(13)\geq 1\right)\\
= & \left[1-\Pr\left(N_J(17)-N_J(13)=0\right)\right]\left[1-\Pr\left(N_S(17)-N_S(13)=0\right)\right]\\
= & (1-\mathrm{e}^{-1.971665})(1-\mathrm{e}^{-1.548042})\\
= & \boxed{0.6777}. \quad \textbf{(Answer: C)}
\end{aligned}$$

□

Problem 1.1.7. (CAS Exam 3 Fall 2005 Question 26: Probability, non-homogeneous) The number of reindeer injuries on December 24 follows a Poisson process with intensity function:

$$\lambda(t)=(t/12)^{1/2} \qquad 0\leq t\leq 24, \text{ where } t \text{ is measured in hours}$$

Calculate the probability that no reindeer will be injured during the last hour of the day.

A. Less than 30%

B. At least 30%, but less than 40%

C. At least 40%, but less than 50%

D. At least 50%, but less than 60%

E. At least 60%

Solution. We need

$$\begin{aligned}
\Pr\left(N(24)-N(23)=0\right) &= \exp\left[-\int_{23}^{24}(t/12)^{1/2}\,\mathrm{d}t\right]\\
&= \exp\left[-\frac{2}{3(12)^{1/2}}(24^{3/2}-23^{3/2})\right]\\
&= \boxed{0.24675}. \quad \textbf{(Answer: A)}
\end{aligned}$$

□

Problem 1.1.8. (Rate function mimics the normal density function) You are given that claim counts follow a non-homogeneous Poisson process with intensity function $\lambda(t)=\mathrm{e}^{-t^2/4}$.

Calculate the probability of at least two claims between time 1 and time 2.

A. Less than 0.10

B. At least 0.10, but less than 0.15

C. At least 0.15, but less than 0.20

D. At least 0.20, but less than 0.25

E. At least 0.25

Solution. The number of claims between time 1 and time 2 is a Poisson random variable with mean

$$\int_1^2 \lambda(t)\,\mathrm{d}t = \int_1^2 \mathrm{e}^{-t^2/4}\,\mathrm{d}t.$$

Note that the integrand resembles the density function of a normal distribution with mean 0 and variance 2, except that the normalizing constant $1/\sqrt{2\pi(2)}$ is missing. Hence

$$\begin{aligned}\int_1^2 \mathrm{e}^{-t^2/4}\,\mathrm{d}t &= \sqrt{2\pi(2)}\int_1^2 \frac{1}{\sqrt{2\pi(2)}}\mathrm{e}^{-t^2/4}\,\mathrm{d}t \\ &= 2\sqrt{\pi}\,\Pr(1 < \mathrm{N}(0,2) < 2) \\ &= 2\sqrt{\pi}\left[\Phi\left(\frac{2}{\sqrt{2}}\right) - \Phi\left(\frac{1}{\sqrt{2}}\right)\right] \\ &= 2\sqrt{\pi}[\Phi(1.41) - \Phi(0.71)] \\ &= 2\sqrt{\pi}(0.9207 - 0.7611) \\ &= 0.565767.\end{aligned}$$

Finally, the required probability is

$$\begin{aligned}\Pr(N(2) - N(1) \geq 2) &= 1 - \Pr(N(2) - N(1) \leq 1) \\ &= 1 - \mathrm{e}^{-0.565767}(1 + 0.565767) \\ &= \boxed{0.1108}. \quad \textbf{(Answer: B)}\end{aligned}$$

□

Problem 1.1.9. (Probability for $N(s)$ given $N(t)$) Customers arrive at a post office in accordance with a Poisson process with a rate of 5 per hour. The post office opens at 9:00 am.

Ten customers have arrived before 11:00 am.

Calculate the probability that only two customers have arrived before 9:30 am.

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

Solution. Note that $N(0.5)|N(2) = 10$ has a binomial distribution with parameters 10 and $0.5/2 = 0.25$. The conditional probability that $N(0.5) = 2$ equals

$$\binom{10}{2} 0.25^2 (1 - 0.25)^8 = \boxed{0.2816}. \quad \textbf{(Answer: D)}$$

□

Part V

Practice Examinations

Prelude. It is now time to test your understanding of the whole syllabus by working out two comprehensive practice exams. Each of these two exams has 45 multiple-choice questions, sorted and distributed in line with the distribution of questions in the Fall 2015 and Spring 2016 Exam S. Detailed solutions are provided and the relevant chapters and sections in this manual identified following each practice exam. The difficulty of these questions is slightly higher than that of a typical CAS exam question, so if you do well in these two practice exams (say, you answer more than 35 questions correctly in each exam), you should be able to pass Exam S with ease!

An abridged version of the CAS exam instructions is reproduced overleaf for your information.

Practice Examination 1

1. People arrive at a train station in accordance with a Poisson process with rate λ.

   At time 0, the train station is empty. At time 10, the bus departs.

   Five people get on the train when the bus departs.

   Calculate the expected amount of waiting time of the person who first arrives at the bus stop.

   A. Less than 2.5
   B. At least 2.5, but less than 5.0
   C. At least 5.0, but less than 7.5
   D. At least 7.5, but less than 10.0
   E. There is not enough information to determine the answer.

2. You are given the following information about the arrival of vehicles:

   - Taxis arrivals follow the Poisson process with rate $\lambda = 1$ per 10 minutes.
   - Bus arrivals follow the Poisson process with rate $\lambda = 4$ per 30 minutes.
   - Streetcar arrivals follow the Poisson process with rate $\lambda = 2$ per hour.

   Calculate the probability that the second vehicle arrives within 10 minutes.

   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

Exam Continued on Next Page

3. Male and female customers arrive independently at a store according to Poisson processes of rate 2 per minute and 3 per minute, respectively.

   Twelve females have arrived the store in ten minutes.

   Calculate the expected number of customers having arrived in ten minutes.

   A. Less than 30
   B. At least 30, but less than 35
   C. At least 35, but less than 40
   D. At least 40, but less than 45
   E. At least 45

4. At at tunnel, each arrival of a car is 4 times as likely to be a truck. The interarrival time of each vehicle follows an exponential distribution.

   Determine the probability that the sixth vehicle that arrives at the tunnel is also the fourth car arriving.

   A. Less than 0.2
   B. At least 0.2, but less than 0.3
   C. At least 0.3, but less than 0.4
   D. At least 0.4, but less than 0.5
   E. At least 0.5

Exam Continued on Next Page

5. Customers enter a store in accordance with a Poisson process with rate 2 per hour. The amount of money spent by a customer is uniformly distributed over the interval $[0, 120]$.

   The store operates for 10 hours per day.

   Calculate the variance of the daily amount of money that the store receives.

   A. Less than 97,000

   B. At least 97,000, but less than 98,000

   C. At least 98,000, but less than 99,000

   D. At least 99,000, but less than 100,000

   E. At least 100,000

6. The lifetime of a newly purchased electronic product is modeled by a hazard rate function given by

$$r(t) = \frac{t}{2}, \quad t > 0.$$

   Calculate the expected lifetime of a newly purchased electronic product.

   A. Less than 1.5

   B. At least 1.5, but less than 2.0

   C. At least 2.0, but less than 2.5

   D. At least 2.5, but less than 3.0

   E. At least 3.0

Exam Continued on Next Page