ACTEX Learning

Study Manual for Exam ASTAM

3rd Edition

Sam A. Broverman, PhD, ASA
Wenjun Jiang, PhD

An SOA Exam

ACTEX Learning

Study Manual for Exam ASTAM

3rd Edition

Sam A. Broverman, PhD,
ASA Wenjun Jiang, PhD

ACTEX Learning

Learn Today. Lead Tomorrow.

Actuarial & Financial Risk Resource Materials

Since 1972

Copyright © 2024, ACTEX Learning, a division of ArchiMedia Advantage Inc.

No portion of this ACTEX Study Manual may be reproduced or transmitted in any part or by any means without the permission of the publisher.

Welcome to Actuarial University

Actuarial University is a reimagined platform built around a more simplified way to study. It combines all the products you use to study into one interactive learning center.

You can find integrated topics using this network icon.

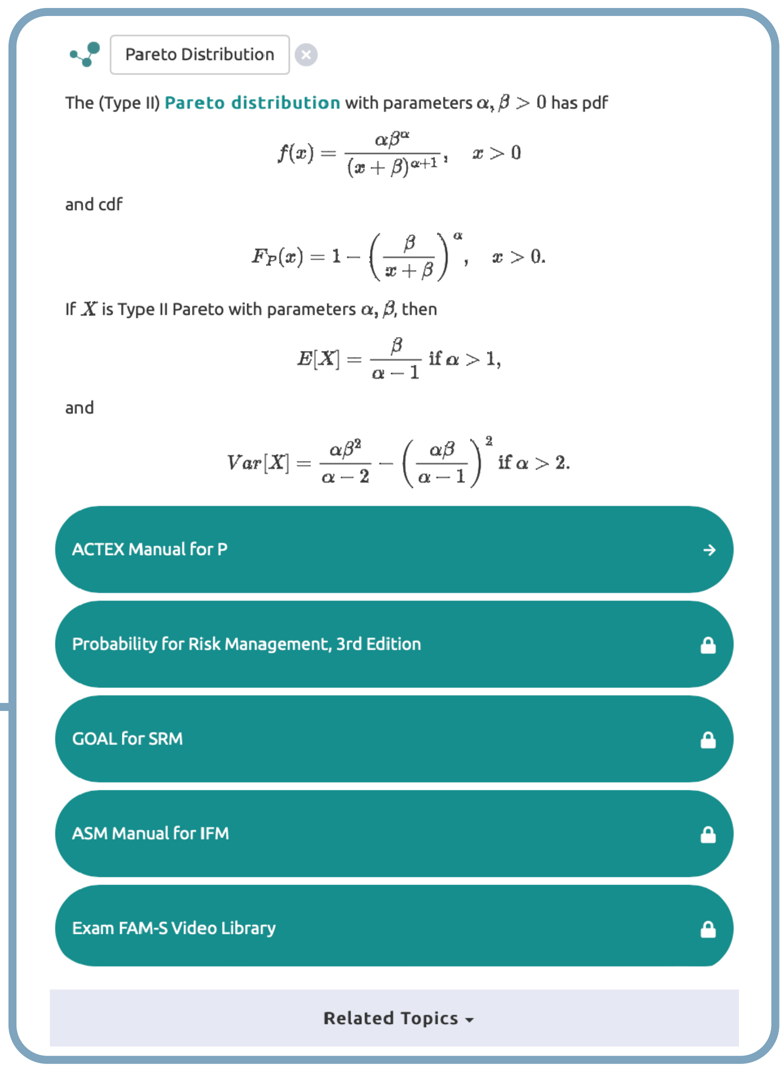
When this icon appears, it will be next to an important topic in the manual. Click the **link** in your digital manual, or search the underlined topic in your print manual.

1. Login to: www.actuarialuniversity.com

2. Locate the **Topic Search** on your exam dashboard and enter the word or phrase into the search field, selecting the best match.

3. A topic "**Hub**" will display a list of integrated products that offer more ways to study the material.

4. Here is an example of the topic **Pareto Distribution:**



Within the **Hub** there will be unlocked and locked products.

Unlocked Products are the products that you own.

ACTEX Manual for P →

Locked Products are products that you do not own, and are available for purchase.

Probability for Risk Management, 3rd Edition

Many of Actuarial University's features are already unlocked with your study program, including:

Instructional Videos*	**Planner**
Topic Search	**Formula & Review Sheet**

Make your study session more efficient with our Planner!

Planner

Template: ACTEX FM Study Manual - New 2022 syllabus

Begin Study: 07/01/2023 End Study: 11/14/2023

✔	7/1/2023 - 7/16/2023	Interest Rates and the Time Value of Money
✔	7/16/2023 - 8/12/2023	Annuities
✔	8/12/2023 - 8/27/2023	Loan Repayment
✔	8/27/2023 - 9/15/2023	Bonds
✔	9/15/2023 - 9/22/2023	Yield Rate of an Investment
✔	9/22/2023 - 10/11/2023	The Term Structure of Interest Rates
✔	10/11/2023 - 10/30/2023	Asset-Liability Management

**Available standalone, or included with the Study Manual Program Video Bundle*

GOAL

Practice. Quiz. Test. Pass!

- 16,000+ Exam-Style Problems
- Detailed Solutions
- Adaptive Quizzes
- 3 Learning Modes
- 3 Difficulty Modes

Free with your ACTEX or ASM Interactive Study Manual

Available for P, FM, FAM, FAM-L, FAM-S, ALTAM, ASTAM, MAS-I, MAS-II, CAS 5, CAS 6U & CAS 6C

Prepare for your exam confidently with GOAL custom Practice Sessions, Quizzes, & Simulated Exams

Actuarial University

QUESTION 19 OF 704 Question # Go! Prev Next

Question — Difficulty: Advanced

An airport purchases an insurance policy to offset costs associated with excessive amounts of snowfall. The insurer pays the airport 300 for every full ten inches of snow in excess of 40 inches, up to a policy maximum of 700.

The following table shows the probability function for the random variable X of annual (winter season) snowfall, in inches, at the airport.

Inches	[0,20)	[20,30)	[30,40)	[40,50)	[50,60)	[60,70)	[70,80)	[80,90)	[90,inf)
Probability	0.06	0.18	0.26	0.22	0.14	0.06	0.04	0.04	0.00

Calculate the standard deviation of the amount paid under the policy.

Possible Answers

A 134 | 235 | X 271 | D 313 | E 352

Help Me Start

Find the probabilities for the four possible payment amounts: 0, 300, 600, and 700.

Solution

With the amount of snowfall as X and the amount paid under the policy as Y, we have

y	$f_Y(y) = P(Y = y)$
0	$P(Y=0) = P(0 \le X < 50) = 0.72$
300	$P(Y=300) = P(50 \le X < 60) = 0.14$
600	$P(Y=600) = P(60 \le X < 70) = 0.06$
700	$P(Y=700) = P(Y \ge 70) = 0.08$

The standard deviation of Y is $\sqrt{E(Y^2) - [E(Y)]^2}$.

$$E(Y) = 0.14 \times 300 + 0.06 \times 600 + 0.08 \times 700 = 134$$

$$E\left(Y^2\right) = 0.14 \times 300^2 + 0.06 \times 600^2 + 0.08 \times 700^2 = 73400$$

$$\sqrt{E(Y^2) - [E(Y)]^2} = \sqrt{73400 - 134^2} = 235.465$$

Common Questions & Errors

Students shouldn't overthink the problem with fractional payments of 300. Also, account for probabilities in which payment cap of 700 is reached.

In these problems, we must distinguish between the REALT RV (how much snow falls) and the PAYMENT RV (when does the insurer pay)?. The problem states "The insurer pays the airport 300 for every full ten inches of snow in excess of 40 inches, up to a policy maximum of 700 ." So the insurer will not start paying UNTIL AFTER 10 full inches in excess of 40 inches of snow is reached (say at 50+ or 51). In other words, the insurer will pay nothing if X<50.

Rate this problem Excellent Needs Improvement Inadequate

Quickly access the Hub for additional learning.

Flag problems for review, record notes, and email your professor.

View difficulty level.

Helpful strategies to get you started.

Full solutions with detailed explanations to deepen your understanding.

Commonly encountered errors.

Rate a problem or give feedback.

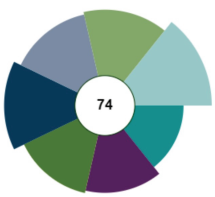
Track your exam readiness with GOAL Score!

GOAL

Available for P, FM, FAM, FAM-L, FAM-S, ALTAM, ASTAM, MAS-I, MAS-II, & CAS 5

GOAL Score tracks your performance through GOAL Practice Sessions, Quizzes, and Exams, resulting in an aggregate weighted score that gauges your exam preparedness.

By measuring both your performance, and the consistency of your performance, **GOAL Score** produces a reliable number that will give you confidence in your preparation before you sit for your exam.



Your GOAL Score
74

Target GOAL Score
70

If your GOAL Score is a 70 or higher, you are well-prepared to sit for your exam!

64
A. Multiple State Models
Weighting: 20%

71
B. Multiple Decrement Models
Weighting: 10%

73
C. Multiple Life Models
Weighting: 10%

82
D. Profit Analysis
Weighting: 15%

76
E. Universal Life Insurance
Weighting: 15%

78
F. Embedded Options and Equity-Linked Insurance
Weighting: 15%

87
G. Pension Plans and Retirement Benefits
Weighting: 15%

See key areas where you can improve.

A. Multiple State Models — A1. Multiple State Probabilities

Difficulty	Mean	Proportion
Core	38	22 / 48
Advanced	38	18 / 48
Mastery	38	8 / 48

Detailed performance tracking.

GOAL Session Activity Summary — Filter: Quizzes — Complete

Created	Last Accessed	Completed	Mode	Categories	Questions	Status	Action
03/24/2023	03/24/2023	03/24/2023	Quiz	G. Pension Plans and Retirement Benefits	15	Complete	Review
03/24/2023	03/24/2023	03/24/2023	Quiz	C1. Joint Life/Last Survivor Probabilities, C2. Contingent Probabilities, C3. Expectation and Variance, C6. ML Models with Transition Intensities, C4. ML Models: Premiums and Reserves (Discrete), C5. ML Models: Premiums and Reserves (Continuous)	15	Complete	Review
03/24/2023	03/24/2023	03/24/2023	Quiz	F. Embedded Options and Equity-Linked Insurance	20	Complete	Review
03/24/2023	03/24/2023	03/24/2023	Quiz	D. Profit Analysis	10	Complete	Review
09/19/2022	09/19/2022	09/19/2022	Quiz	G4. Pension Plans and Retirement Benefits (Written Answers)	2	Complete	Review

Quickly return to previous sessions.

Contents

Copyright © 2024 ACTEX Learning

Copyright © 2024 ACTEX Learning

INTRODUCTORY COMMENTS

This study guide is designed to help in the preparation for the Society of Actuaries ASTAM Examination.

This manual consists of a summary of notes, illustrative examples and problem sets with detailed solutions, four practice exams and a collection of exam-type questions. The level of difficulty of the practice exam questions has been designed to be similar to that of questions on the SOA exam.

We have attempted to be thorough in the coverage of the topics upon which the exam is based, and consistent with the notation can content of the official exam references.

Because of the time constraint on the exam, a crucial aspect of exam taking is the ability to work quickly. We believe that working through many problems and examples is a good way to build up the speed at which you work. It can also be worthwhile to work through problems that have been done before, as this helps to reinforce familiarity, understanding and confidence. Working many problems will also help in being able to more quickly identify topic and question types. We have attempted, wherever possible, to emphasize shortcuts and efficient and systematic ways of setting up solutions. There are also occasional comments on interpretation of the language used in some exam questions. While the focus of the study guide is on exam preparation, from time to time there will be comments on underlying theory in places that I feel those comments may provide useful insight into a topic.

The notes and examples are divided into 35 sections of varying lengths, with some suggested time frames for covering the material. There are about 140 examples in the notes and over 770 exercises in the problem sets, all with detailed solutions. Each of the 4 practice exams has 60 points assigned like the SOA exams, also with detailed solutions. Some of the examples and exercises are taken from previous SOA exams. Some of the in the problem sets that have come from previous SOA exams. Some of the problem set exercises are more in depth than actual exam questions, but the practice exam questions have been created in an attempt to replicate the level of depth and difficulty of actual exam questions. In total, there are over 1000 examples/problems/sample exam questions with detailed solutions. ACTEX gratefully acknowledges the SOA for allowing the use of its exam problems in this study guide.

We suggest that you work through the study guide by studying a section of notes and then attempting the problems in the problem set that follows that section. The order of the sections of notes is the order that we recommend in covering the material. The order of topics in this manual close to but not exactly the same as the order presented on the exam syllabus.

It has been our intention to make this study guide self-contained and comprehensive for the ASTAM Exam topics, It is advisable to refer to original reference material on all topics.

While the ability to derive formulas used on the exam is usually not the focus of an exam question, it is useful in enhancing the understanding of the material and may be helpful in memorizing formulas. There may be an occasional reference in the review notes to a derivation, but you are encouraged to review the official reference material for more details on formula derivations. It is assumed that you are familiar with the material covered in the FAM-S exam syllabus, as some of the material is prerequisite to material covered in the ASTAM exam.

Of the various calculators that are allowed for use on the exam, we are most familiar with the BA II PLUS. It has several easily accessible memories. The TI-30X IIS has the advantage of a multi-line display. Both have the functionality needed for the exam.

There is a set of tables that has been provided with the exam in past sittings. These tables consist of some detailed description of a number of probability distributions along with tables for the standard normal and chi-squared distributions. The tables can be downloaded from the SOA website www.soa.org.

If you have any questions, comments, criticisms or compliments regarding this study guide, please contact the publisher ACTEX, or you may contact one of us directly at the addresses below. We apologize in advance for any errors, typographical or otherwise, that you might find, and it would be greatly appreciated if you would bring them to our attention. ACTEX will maintain a website for errata that can be accessed from www.actexmadriver.com.

It is our sincere hope that you find this study guide helpful and useful in your preparation for the exam. We wish you the best of luck on the exam.

February 2024

Samuel Broverman

Department of Statistical Sciences www.sambroverman.com

University of Toronto E-mail: sam.broverman@utoronto.ca or 2brove@rogers.com

Wenjun Jiang

Department of Mathematics ands Statistics science.ucalgary.ca/mathematics-statistics

University of Calgary E-mail: wenjun.jiang@ucalgary.ca

Section 2

Preliminary Review - Random Variables I

Probability, Density and Distribution Functions

This section covers more preliminary topics needed for the later ASTAM syllabus material and Section 2.7 relates to Section 4.2.4 in the "Loss Models" book. The suggested time frame for covering this section is two hours. A brief review of some basic calculus relationships is presented first.

2.1 Calculus Review

Natural logarithm and exponential functions
$\ln(x) = \log(x)$ is the logarithm to the base e;

$$\begin{array}{lll} \ln(e) = 1, & \ln(1) = 0, & e^0 = 1, \\ \ln(e^y) = y, & e^{\ln(x)} = x, & \ln(a^y) = y \times \ln(a), \\ \ln(y \times z) = \ln(y) + \ln(z), & \ln\left(\frac{y}{z}\right) = \ln(y) - \ln(z), & e^x e^z = e^{x+z}, \\ (e^x)^w = e^{xw} & & \end{array} \tag{2.1}$$

Differentiation
For the function $f(x)$, $f'(x) = \dfrac{df}{dx} = \lim\limits_{h \to 0} \dfrac{f(x+h) - f(x)}{h}$ (2.2)

Product rule:

$$\frac{d}{dx}[g(x) \times h(x)] = g'(x) \times h(x) + g(x) \times h'(x) \tag{2.3}$$

Quotient rule:

$$\frac{d}{dx}\left[\frac{g(x)}{h(x)}\right] = \frac{h(x) \times g'(x) - g(x) \times h'(x)}{[h(x)]^2} \tag{2.4}$$

Chain rule:

$$\begin{array}{l} \dfrac{d}{dx} f(g(x)) = f'(g(x)) \times g'(x), \quad \dfrac{d}{dx} \ln[g(x)] = \dfrac{g'(x)}{g(x)} \\ \dfrac{d}{dx}[g(x)]^n = n \times [g(x)]^{n-1} \times g'(x), \quad \dfrac{d}{dx} a^x = a^x \times \ln(a) \end{array} \tag{2.5}$$

Copyright © 2024 ACTEX Learning

Integration:

$$\int x^n \, dx = \frac{x^{n+1}}{n+1} + c, \; \int a^x \, dx = \frac{a^x}{\ln(a)} + c, \; \int \frac{1}{a+bx} \, dx = \frac{1}{b} \times \ln[a+bx] + c \tag{2.6}$$

Integration by parts:

$$\int_a^b u(t) \; dv(t) = u(b) \times v(b) - u(a) \times v(a) - \int_a^b v(t) \; du(t) \tag{2.7}$$

for **definite integrals**, and

$$\int u \; dv = uv - \int v \; du$$

for indefinite integrals (this is derived by integrating both sides of the product rule); note that

$$dv(t) = v'(t) \; dt \quad \text{and} \quad du(t) = u'(t) \; dt$$

$$\frac{d}{dx} \int_a^x g(t) \, dt = g(x), \qquad \frac{d}{dx} \int_x^b g(t) \, dt = -g(x) \tag{2.8}$$

$$\frac{d}{dx} \int_{h(x)}^{j(x)} g(t) dt = g(j(x)) \times j'(x) - g(h(x)) \times h'(x) \tag{2.9}$$

$$\int_0^\infty x^n e^{-kx} \; dx = \frac{n!}{k^{n+1}} \quad \text{if} \quad k > 0 \quad \text{and } n \text{ is an integer } \geq 0 \tag{2.10}$$

The word "model" used in the context of a loss model, usually refers to the distribution of a loss random variable. Random variables are the basic components used in actuarial modeling. In this section we review the definitions and illustrate the variety of random variables that we will encounter in the ASTAM Exam material.

A **random variable** is a numerical quantity that is related to the outcome of some random experiment on a probability space. For the most part, the random variables we will encounter are the numerical outcomes of some loss related event such as the dollar amount of claims in one year from an auto insurance policy, or the number of tornados that touch down in Kansas in a one year period.

2.2 Discrete Random Variable

The random variable X is **discrete** and is said to have a **discrete distribution** if it can take on values only from a finite or countably infinite sequence (usually the integers or some subset of the integers). As an example, consider the following two random variables related to successive tosses of a coin:

$X = 1$ if the first head occurs on an even-numbered toss, $X = 0$ if the first head occurs on an odd-numbered toss;

$Y = n$, where n is the number of the toss on which the first head occurs.

Both X and Y are discrete random variables, where X can take on only the values 0 or 1, and Y can take on any positive integer value.

Probability function of a discrete random variable

The **probability function** (pf) of a discrete random variable is usually denoted $p(x)$ (or $f(x)$), and is equal to $P[X = x]$. As its name suggests, the probability function describes the probability of individual outcomes occurring.

The probability function must satisfy the following two conditions:

$$(i) 0 \le p(x) \le 1 \quad \text{for all } x, \qquad \text{and} \qquad (ii) \sum_{all\ x} p(x) = 1 \tag{2.11}$$

For the random variable X above, the probability function is $p(0) = \frac{2}{3}$, $p(1) = \frac{1}{3}$,

and for Y it is $p(k) = \frac{1}{2^k}$ for $k = 1, 2, 3, \ldots$.

An event A is a subset of the set of all possible outcomes of X, and the probability of event A occurring is $P[A] = \sum_{x \in A} p(x)$.

For Y above, $P[Y \text{ is even}] = P[Y = 2,4,6,...] = \frac{1}{2^2} + \frac{1}{2^4} + \frac{1}{2^6} + \cdots = \frac{1}{3}$,

and this is also equal to $P[X = 1]$.

Some specific discrete random variables will be considered in some detail a little later in this study guide, but the following is a brief description of a few important discrete distributions.

Discrete uniform distribution

If N is an integer ≥ 1, the **discrete uniform distribution** on the integers from 1 to N has probability function $P[X = k] = p(k) = \frac{1}{N}$ for $k = 1, \ldots, N$ and $p(k) = 0$ otherwise. The discrete uniform distribution with $N = 6$ would apply to the outcome of the toss of a fair die.

Binomial distribution

The **binomial distribution** with parameters m (integer ≥ 1) and number q $(0 < q < 1)$ has probability function

$$P[X = k] = p(k) = \binom{m}{k} q^k (1-q)^{m-k}, \qquad k = 0, 1, \ldots, m$$

where $\binom{m}{k} = \frac{m!}{k! \times (m-k)!}$ is the "binomial coefficient", and $p(k) = 0$ otherwise. The binomial distribution describes the number of "successful outcomes" out of m trials of a random "experiment" in which trials are mutually independent and each trial results in a successful outcome with probability q or unsuccessful outcome with probability $1 - q$.

Poisson distribution

The **Poisson distribution** with parameter λ has probability function

$$P[X = k] = \frac{e^{-\lambda} \lambda^k}{k!} \text{ for } k \ge 0,$$

where k is an integer. The Poisson distribution is a very important distribution in actuarial applications to the modeling of the number of events occurring in a specified period of time.

Copyright ©2024 ACTEX Learning

2.3 Continuous Random Variable

A **continuous random variable** usually can assume numerical values from an interval of real numbers, perhaps the entire set of real numbers. As an example, the length of time between successive streetcar arrivals at a particular (in service) streetcar stop could be regarded as a continuous random variable (assuming that time measurement can be made perfectly accurate).

Probability density function

A continuous random variable X has a **probability density function (pdf)** denoted $f(x)$ or $f_X(x)$ (or sometimes denoted $p(x)$), which is a continuous function (except possibly at a finite or countably infinite number of points). For a continuous random variable, we do not describe probability at single points. We describe probability in terms of intervals. In the streetcar example, we would not define the probability that the next street car will arrive in exactly 1.23 minutes, but rather we would define a probability such as the probability that the streetcar will arrive between 1 and 1.5 minutes from now.

Probabilities related to X are found by integrating the density function over an interval.

$P[X \in (a,b)] = P[a < X < b]$ is defined to be equal to $\int_a^b f(x)\ dx$.

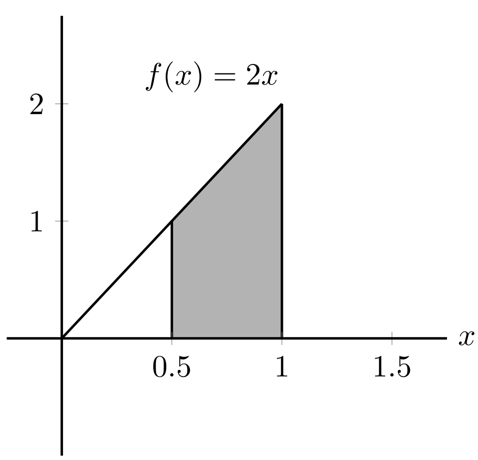
A pdf $f(x)$ must satisfy $\quad (i) f(x) \geq 0$ for all x and $\quad (ii) \int_{-\infty}^{\infty} f(x)\ dx = 1 \qquad (2.12)$

Often, the region of non-zero density is a finite interval, and $f(x) = 0$ outside that interval. If $f(x)$ is continuous except at a finite number of points, then probabilities are defined and calculated as if $f(x)$ was continuous everywhere (the discontinuities are ignored).

For example, suppose that X has density function $f(x) = \begin{cases} 2x & \text{for } \quad 0 < x < 1 \\ 0 & \text{elsewhere} \end{cases}$.

Then f satisfies the requirements for a density function, since $\int_{-\infty}^{\infty} f(x)\ dx = \int_0^1 2x\ dx = 1$.

Then, for example $P[.5 < X < 1] = \int_{0.5}^{1} 2x\,dx = x^2\Big|_{0.5}^{1} = 0.75$. This is illustrated in the shaded area in the graph below.



For a continuous random variable X,

$$P[a < X < b] = P[a \leq X < b] = P[a < X \leq b] = P[a \leq X \leq b],$$

so that when calculating the probability for a continuous random variable on an interval, it is irrelevant whether or not the endpoints are included. For a continuous random variable, $P[X = a] = 0$ for any point a; non-zero probabilities only exist over an interval, not at a single point.

Some specific continuous distributions will be considered in some detail later in this study guide, but the following is a brief description of a few important continuous distributions.

Continuous uniform distribution
If a and b are real numbers with $a < b$, the **continuous uniform distribution** on the interval (a,b) has pdf $f(x) = \frac{1}{b-a}$ for $a < x < b$, and $f(x) = 0$, otherwise.

Exponential distribution
If $\lambda > 0$ is a real number, then the **exponential distribution** with parameter λ has pdf $f(x) = \frac{e^{-\lambda x}}{\lambda}$ for $x > 0$, and $f(x) = 0$, otherwise. The exponential distribution and generalizations of it are very important in actuarial modelling.

Another very important distribution, central to probability and statistics, is the **normal distribution**. This distribution will be considered in some detail a little later in this section.

2.4 Mixed Distribution

A random variable may have some points with non-zero **probability mass** and with a continuous pdf elsewhere. Such a distribution may be referred to as a **mixed distribution**, but the more general notion of mixtures of distributions will be covered later. The sum of the probabilities at the discrete points of probability plus the integral of the density function on the continuous region for X must be 1. For example, suppose that X has probability of 0.5 at $X = 0$, and X is a continuous random variable on the interval (0,1) with density function $f(x) = x$ for $0 < x < 1$, and X has no density or probability elsewhere. This satisfies the requirements for a random variable since the total probability is

$$P[X = 0] + \int_0^1 f(x)\, dx = 0.5 + \int_0^1 x\, dx = 0.5 + 0.5 = 1.$$

Then,

$$P[0 < X < 0.5] = \int_0^{.5} x\, dx = 0.125,$$

and

$$P[0 \leq X < 0.5] = P[X = 0] + P[0 < X < 0.5] = 0.5 + 0.125 = 0.625.$$

Notice that for this random variable $P[0 < X < 0.5] \neq P[0 \leq X < 0.5]$ because there is a probability mass at $X = 0$.

Copyright © 2024 ACTEX Learning

2.5 Cumulative Distribution, Survival and Hazard Functions

Given a random variable X, the **cumulative distribution function** of X (also called the **distribution function**, or cdf) is $F(x) = P[X \le x]$ (also denoted $F_X(x)$).

The cdf $F(x)$ is the "left-tail" probability, or the probability to the left of and including x.

The **survival function** is the complement of the distribution function,

$$S(x) = 1 - F(x) = P[X > x]. \tag{2.13}$$

The event $X > x$ is referred to as a "tail" or right tail of the distribution.

For any cdf $P[a < X \le b] = F(b) - F(a), \quad \lim_{x \to \infty} F(x) = 1, \quad \lim_{x \to -\infty} F(x) = 0.$ (2.14)

For a discrete random variable with probability function $p(x)$, $F(x) = \sum_{w \le x} p(w)$, and in this case $F(x)$ is a "step function" (see Example 2.1 below); it has a jump (or step increase) at each point that has non-zero probability, while remaining constant until the next jump. Note that for a discrete random variable, $F(x)$ includes the probability at the point x as well as the sum of the probabilities of all the points to the left of x.

If X has a continuous distribution with density function $f(x)$, then

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(t)\ dt \quad \text{and} \quad S(x) = P(X > x) = \int_{x}^{\infty} f(t)\ dt \tag{2.15}$$

and $F(x)$ is a continuous, differentiable, non-decreasing function such that

$$\frac{d}{dx}\ F(x) = F'(x) = -S'(x) = f(x).$$

Also, for a continuous random variable, the **hazard rate** or **failure rate** is

$$h(x) = \frac{f(x)}{1 - F(x)} = \frac{f(x)}{S(x)} = -\frac{d}{dx}\ \ln S(x). \tag{2.16}$$

The **cumulative hazard function** is

$$H(x) = \int_0^x h(t)\ dt \tag{2.17}$$

If X is continuous and $X \ge 0$, then the survival function satisfies

$$S(0) = 1 \text{ and } S(x) = e^{-\int_0^x h(t)\ dt} = e^{-H(x)}.$$

If X has a mixed distribution with some discrete points and some continuous regions, then $F(x)$ is continuous except at the points of non-zero probability mass, where $F(x)$ will have a jump.

The region of positive probability of a random variable is called the **support** of the random variable.

Copyright ©2024 ACTEX Learning

2.6 Examples of Distribution Functions

The following examples illustrate the variety of distribution functions that can arise from random variables. The **support of a random variable** is the set of points over which there is positive probability or density.

Example 2.1.

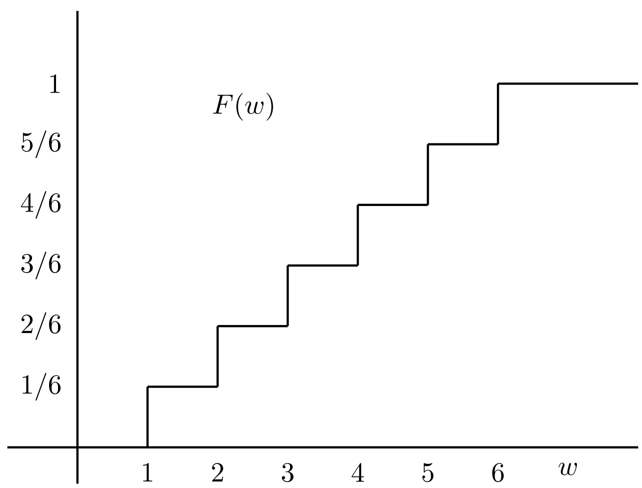
Finite Discrete Random Variable (finite support)

W = number turning up when tossing one fair die. W has probability function

$p_W(w) = P[W = w] = \dfrac{1}{6}$ for $w = 1, 2, 3, 4, 5, 6$.

$$F_W(w) = P[W \leq w] = \begin{cases} 0 & \text{if} \quad w < 1 \\ \frac{1}{6} & \text{if} \quad 1 \leq w < 2 \\ \frac{2}{6} & \text{if} \quad 2 \leq w < 3 \\ \frac{3}{6} & \text{if} \quad 3 \leq w < 4 \\ \frac{4}{6} & \text{if} \quad 4 \leq w < 5 \\ \frac{5}{6} & \text{if} \quad 5 \leq w < 6 \\ 1 & \text{if} \quad w \geq 6 \end{cases}$$

The graph of the cdf is a step-function that increases at each point of probability by the amount of probability at that point (all 6 points have probability $\frac{1}{6}$ in this example). Since the support of W is finite (the support is the set of integers from 1 to 6), $F_W(w)$ reaches 1 at the largest point $W = 6$ (and stays at 1 for all $w \geq 6$).



□

Copyright © 2024 ACTEX Learning

Example 2.2.

Infinite Discrete Random Variable (infinite support)

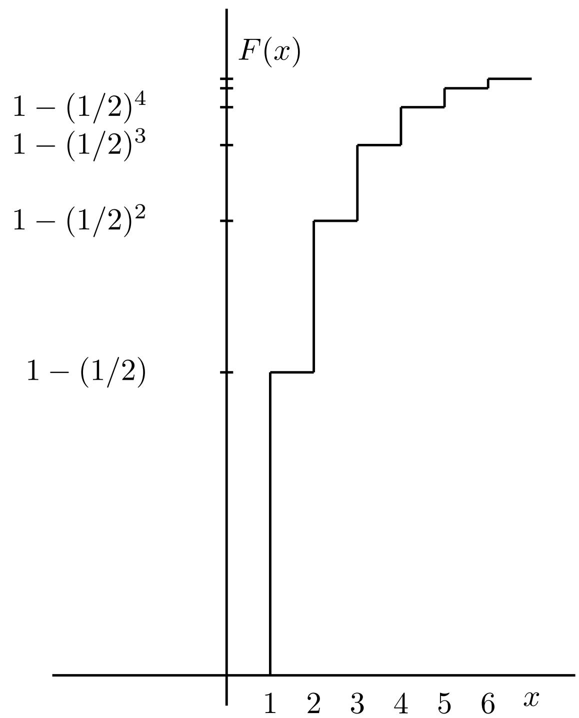
X = the toss number of successive independent tosses of a fair coin on which the first head turns up.

X can be any integer ≥ 1, and the probability function of X is $p_X(x) = \dfrac{1}{2^x}$.

The cdf is

$$F_X(x) = \sum_{k=1}^{x} \frac{1}{2^k} = 1 - \frac{1}{2^x} \text{ for } x = 1, 2, 3, \ldots.$$

The graph of the cdf is a step-function that increases at each point of probability by the amount of probability at that point. Since the support of X is infinite (the support is the set of integers ≥ 1) $F_X(x)$ never reaches 1, but approaches 1 as a limit as $x \to \infty$. The graph of $F_X(x)$ is



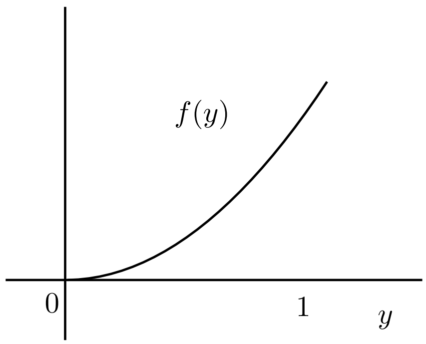
□

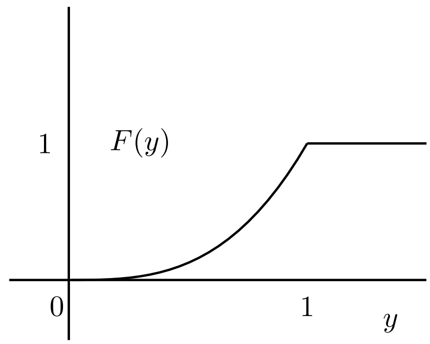
Example 2.3.

<u>Continuous Random Variable on a Finite Interval</u>

Y is a continuous random variable on the interval (0,1) with density function

$$f_Y(y) = \begin{cases} 3y^2 & \text{for} \quad 0 < y < 1 \\ 0 & \text{elsewhere} \end{cases}. \quad \text{Then} \quad F_Y(y) = \begin{cases} 0 & \text{if} \quad y < 0 \\ y^3 & \text{if} \quad 0 \le y < 1 \\ 1 & \text{if} \quad y \ge 1 \end{cases}.$$





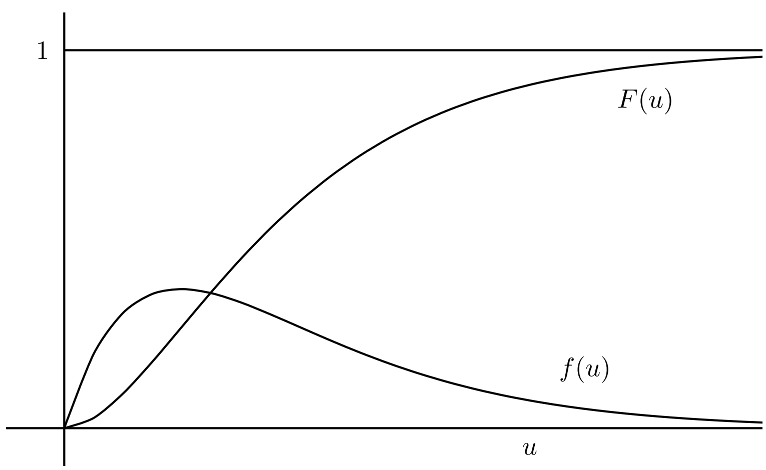
□

Example 2.4.

<u>Continuous Random Variable on an Infinite Interval</u>

U is a continuous random variable on the interval $(0,\infty)$ with density function

$$f_U(u) = \begin{cases} ue^{-u} & \text{for} \quad u > 0 \\ 0 & \text{for} \quad u \le 0 \end{cases}. \quad \text{Then} \quad F_U(u) = \begin{cases} 0 & \text{for} \quad u \le 0 \\ 1-(1+u)e^{-u} & \text{for} \quad u > 0 \end{cases}.$$



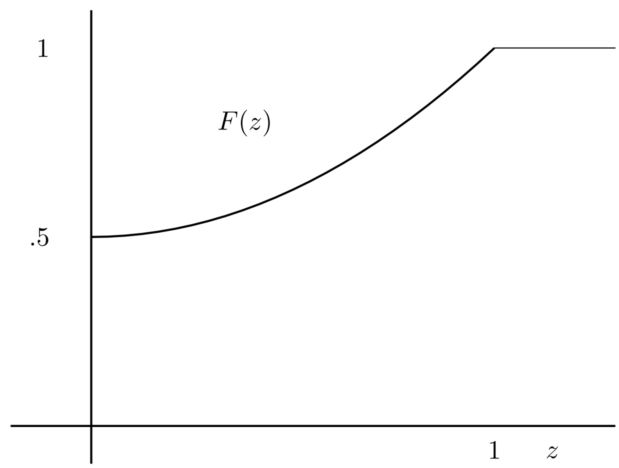
□

Example 2.5.

Mixed Random Variable

Z has a mixed distribution on the interval [0,1). Z has probability of 0.5 at $Z = 0$, and Z has density function $f_Z(z) = z$ for $0 < z < 1$, and Z has no density or probability elsewhere. Then,

$$F_Z(z) = \begin{cases} 0 & \text{if} \quad z < 0 \\ 0.5 & \text{if} \quad z = 0 \\ 0.5 + \frac{1}{2}z^2 & \text{if} \quad 0 < z < 1 \\ 1 & \text{if} \quad z \geq 1 \end{cases}.$$



□

2.7 The Empirical Distribution

The **empirical distribution** is a discrete random variable constructed from a random sample. Suppose that the random sample consists of n observations, say $x_1, x_2, ..., x_n$. If the data is from a loss distribution, then the x_i's are loss amounts, and if the data is from a survival distribution, they are times of death or failure. Either way, knowing the exact value of each outcome is what is referred to as complete data.

The empirical distribution assigns a probability of $\frac{1}{n}$ to each x_j. There may be some repeated numerical values of the observations, so let us suppose that there are k distinct numerical values that have been observed (some possibly repeated). Let us assume that these k values have been ordered from smallest to largest as $y_1 < y_2 < \cdots < y_k$, with s_j =number of observations equal to y_j (so that $s_1 + s_2 + \cdots + s_k = n$, the total number of observed values).

For instance, if we have a sample of $n = 8$ points, say $x_1,\ldots,x_8$ that are 7, 2, 4, 4, 6, 2, 1, 9, then we have $k = 6$ distinct values (in numerical order) $y_1 = 1$, $y_2 = 2, y_3 = 4, y_4 = 6, y_5 = 7$, $y_6 = 9$, with $s_1 = 1$, $s_2 = 2, s_3 = 2, s_4 = 1, s_5 = 1$, $s_6 = 1$. This empirical distribution is a 6-point discrete random variable based on the numerical values of the y's, and it has all the properties of a discrete random variable.

The **empirical distribution probability function** is defined to be

$$p_n(y_j) = \frac{\text{number of } x_i\text{'s that are equal to } y_j}{n} = \frac{s_j}{n} \tag{2.18}$$

(a probability of $\frac{1}{n}$ is assigned to each of the n observations, and s_j denotes the number of observations equal to y_j). In the example above, $p_8(1) = \frac{1}{8}, p_8(2) = \frac{2}{8}$, $p_8(4) = \frac{2}{8}$, $p_8(6) = \frac{1}{8}$, $p_8(7) = \frac{1}{8}, p_8(9) = \frac{1}{8}$.

The **empirical distribution function** is the distribution function of the empirical random variable that we have just defined: $\boldsymbol{F_n(t) = \dfrac{\textbf{number of } x_i\textbf{'s} \leq t}{n}}$ (2.19)

In the example above, $F_8(4) = \frac{5}{8}$. The subscript "8" in F_8 just indicates the total number of data points in the sample.

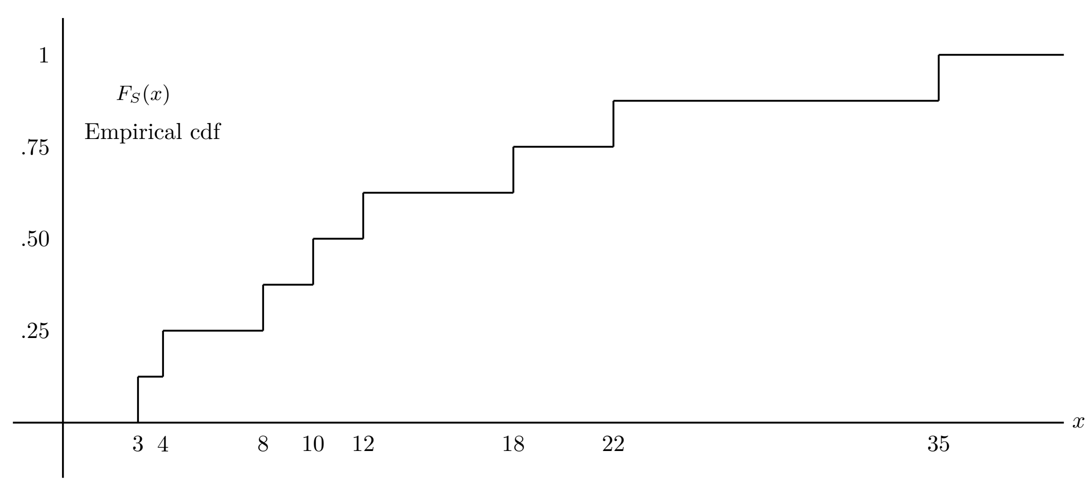
Example 2.6. A random sample of $n = 8$ values from distribution of X is given: $3, 4, 8, 10, 12, 18, 22, 35$
Formulate the empirical distribution function $F_8(x)$ and draw the graph of $F_{10}(x)$.

Solution.

There are no repeated points. $F_8(t) = \dfrac{\text{number of } x_i\text{'s} \leq t}{8}$. The empirical distribution function has values

$$F_8(3) = .125,\ F_8(4) = .25,\ F_8(8) = .375,\ \ldots,\ F_8(22) = .875,\ F_8(35) = 1.0.$$

The graph of the empirical distribution function $F_8(x)$ is a step function, rising by .125 at each of the sample x-values. The following is the graph of the empirical distribution function.



□

2.8 Gamma Function and Related Functions

Many of the continuous distributions described in the ASTAM Exam Tables make reference to the **gamma function** and the **incomplete gamma function**. The definitions of these functions are

- **gamma function**: $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1}e^{-t}\, dt$ for $\alpha > 0$
- **incomplete gamma function**: $\Gamma(\alpha; x) = \dfrac{1}{\Gamma(\alpha)} \times \int_0^x t^{\alpha-1}e^{-t}\, dt$ for $\alpha > 0,\ x > 0$ (2.20)

Some important points to note about these functions are the following:

- if n is an integer and $n \geq 1$, then $\Gamma(n) = (n-1)!$ (2.21)
- $\Gamma(\alpha+1) = \alpha \times \Gamma(\alpha)$ and $\Gamma(\alpha+k) = (\alpha+k-1) \times (\alpha+k-2) \times \cdots \times \alpha \times \Gamma(\alpha)$ (2.22)

  for any $\alpha > 0$ and integer $k \geq 1$.
- $\int_0^\infty x^k e^{-cx}\, dx = \dfrac{\Gamma(k+1)}{c^{k+1}}$ for $k \geq 0$ and $c > 0$ (use substitution $u = cx$) (2.23)
- $\int_0^\infty \dfrac{1}{x^k}\, e^{-c/x}\, dx = \dfrac{\Gamma(k-1)}{c^{k-1}}$ for $k > 1$ and $c > 0$ (use substitution $u = \frac{c}{x}$) (2.24)

Some of the table distributions make reference to the **incomplete beta function**:

$$\beta(a,b;x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^x t^{a-1} \times (1-t)^{b-1}\, dt \text{ for } 0 \leq x \leq 1,\ a,\ b > 0 \tag{2.25}$$

References to the gamma function have been rare and the incomplete functions have not been referred to on the released exams. It is useful to remember the integral relationship $\int_0^\infty x^k e^{-cx}\, dx = \frac{\Gamma(k+1)}{c^{k+1}}$, particularly in the case in which k is a non-negative integer.

In that case, we get $\boldsymbol{\int_0^\infty x^k e^{-cx}\, dx = \frac{k!}{c^{k+1}}}$, which can occasionally simplify integral relationships. This relationship is embedded in the definition of the gamma distribution in the ASTAM Exam Table.

The pdf of the **gamma distribution** with parameters α and θ is $f(t) = \frac{t^{\alpha-1}e^{-t/\theta}}{\theta^\alpha\,\Gamma(\alpha)}$, defined on the interval $t > 0$. This means that $\int_0^\infty \frac{t^{\alpha-1}e^{-t/\theta}}{\theta^\alpha\,\Gamma(\alpha)} dt = 1$, which can be reformulated as $\int_0^\infty t^{\alpha-1}e^{-t/\theta}\, dx = \theta^\alpha \times \Gamma(\alpha)$. If we let $\theta = \frac{1}{c}$ and $k = \alpha - 1$, we get the relationship seen above

$$\int_0^\infty x^k e^{-cx}\, dx = \frac{\Gamma(k+1)}{c^{k+1}}. \tag{2.26}$$

Looking at the various continuous distributions in the ASTAM Exam Table gives some hints at calculating a number of integral forms. For instance, the pdf of the **beta distribution** with parameters $a,\ b,\ \theta = 1$ is

$$f(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \times x^{a-1}(1-x)^{b-1} \quad \text{for} \quad 0 < x < 1.$$

Therefore, $\int_0^1 \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \times x^{a-1}(1-x)^b - 1\, dx = 1$, from which we get

$$\int_0^1 x^{a-1}(1-x)^{b-1}\, dx = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}.$$

Section 2 Problem Set

Preliminary Review - Random Variables I

1. Let X be a discrete random variable with probability function

$$P[X = x] = \frac{2}{3^x} \quad \text{for} \quad x = 1, 2, 3, \ldots .$$

What is the probability that X is even?

(A) $\frac{1}{4}$ (B) $\frac{2}{7}$ (C) $\frac{1}{3}$ (D) $\frac{2}{3}$ (E) $\frac{3}{4}$

2. For a certain discrete random variable on the non-negative integers, the probability function satisfies the relationships $P(0) = P(1)$ and $P(k+1) = \frac{1}{k} \times P(k)$ for $k = 1, 2, 3, \ldots$.

Find $P(0)$.

(A) $\ln e$ (B) $e - 1$ (C) $(e+1)^{-1}$ (D) e^{-1} (E) $(e-1)^{-1}$

3. Let X be a continuous random variable with density function

$$f(x) = \begin{cases} 6x(1-x) & \text{for} \quad 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}.$$

Calculate $P\left[|X - \frac{1}{2}| > \frac{1}{4}\right]$.

(A) 0.0521 (B) 0.1563 (C) 0.3125 (D) 0.5000 (E) 0.8000

4. Let X be a random variable with distribution function

$$F(x) = \begin{cases} 0 & \text{for} \quad x < 0 \\ \frac{x}{8} & \text{for} \quad 0 \le x < 1 \\ \frac{1}{4} + \frac{x}{8} & \text{for} \quad 1 \le x < 2 \\ \frac{3}{4} + \frac{x}{12} & \text{for} \quad 2 \le x < 3 \\ 1 & \text{for} \quad x \ge 3 \end{cases}.$$

Calculate $P[1 \le X \le 2]$.

(A) $\frac{1}{8}$ (B) $\frac{3}{8}$ (C) $\frac{7}{16}$ (D) $\frac{13}{24}$ (E) $\frac{19}{24}$

5. Let X_1,X_2 and X_3 be three independent continuous random variables each with density function

$$f(x) = \begin{cases} \sqrt{2} - x & \text{for} \quad 0 < x < \sqrt{2} \\ 0 & \text{otherwise} \end{cases}.$$

What is the probability that exactly 2 of the 3 random variables exceeds 1?

(A) $\frac{3}{2} - \sqrt{2}$ (B) $3 - 2\sqrt{2}$ (C) $3(\sqrt{2}-1)(2-\sqrt{2})^2$

(D) $\left(\frac{3}{2} - \sqrt{2}\right)^2 \left(\sqrt{2} - \frac{1}{2}\right)$ (E) $3\left(\frac{3}{2} - \sqrt{2}\right)^2 \left(\sqrt{2} - \frac{1}{2}\right)$

6. Let X_1, X_2 and X_3 be three independent, identically distributed random variables each with density function

$$f(x) = \begin{cases} 3x^2 & \text{for} \quad 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}.$$

Let $Y = \max\{X_1,X_2,X_3\}$. Find $P\left[Y > \frac{1}{2}\right]$.

(A) $\dfrac{1}{64}$ (B) $\dfrac{37}{64}$ (C) $\dfrac{343}{512}$ (D) $\dfrac{7}{8}$ (E) $\dfrac{511}{512}$

7. Let the distribution function of X for $x > 0$ be $F(x) = 1 - \sum_{k=0}^{3} \frac{x^k e^{-x}}{k!}$.

What is the density function of X for $x > 0$?

(A) e^{-x} (B) $\dfrac{x^2 e^{-x}}{2}$ (C) $\dfrac{x^3 e^{-x}}{6}$ (D) $\dfrac{x^3 e^{-x}}{6} - e^{-x}$ (E) $\dfrac{x^3 e^{-x}}{6} + e^{-x}$

8. Let X have the density function $f(x) = \frac{3x^2}{\theta^3}$ for $0 < x < \theta$, and $f(x) = 0$, otherwise.

If $P[X > 1] = \frac{7}{8}$, find the value of θ.

(A) $\dfrac{1}{2}$ (B) $\left(\dfrac{7}{8}\right)^{1/3}$ (C) $\left(\dfrac{8}{7}\right)^{1/3}$ (D) $2^{1/3}$ (E) 2

9. A large wooden floor is laid with strips 2 inches wide and with negligible space between strips. A uniform circular disk of diameter 2.25 inches is dropped at random on the floor. What is the probability that the disk touches three of the wooden strips?

(A) $\dfrac{1}{\sqrt{\pi}}$ (B) $\dfrac{1}{\pi}$ (C) $\dfrac{1}{4}$ (D) $\dfrac{1}{8}$ (E) $\dfrac{1}{\pi^2}$

10. If X has a continuous uniform distribution on the interval from 0 to 10, then what is $P\left[X + \frac{10}{X} > 7\right]$?

(A) $\dfrac{3}{10}$ (B) $\dfrac{31}{70}$ (C) $\dfrac{1}{2}$ (D) $\dfrac{39}{70}$ (E) $\dfrac{7}{10}$

11. For a loss distribution where $x \ge 2$, you are given:

i) The hazard rate function:

$$h(x) = \frac{z^2}{2x}, \text{ for } x \ge 2$$

ii) A value of the distribution function: $F(5) = 0.84$

Calculate z.

(A) 2 (B) 3 (C) 4 (D) 5 (E) 6

Copyright © 2024 ACTEX Learning

Section 2 Problem Set Solutions

1. $P[X \text{ is even}] = P[X=2] + P[X=4] + P[X=6] + \cdots$

$$= \frac{2}{3} \times \left[\frac{1}{3} + \frac{1}{3^3} + \frac{1}{3^5} + \cdots\right] = \frac{2}{3^2} \times \frac{1}{1-\frac{1}{3^2}} = \frac{1}{4}$$

Answer A

2. $P(2) = P(1) = P(0)$, $P(3) = \frac{1}{2} \times P(2) = \frac{1}{2!} \times P(0)$, $\ldots$, $P(k) = \frac{1}{(k-1)!} \times P(0)$.

The probability function must satisfy the requirement $\sum\limits_{i=0}^{\infty} P(i) = 1$ so that

$$P(0) + \sum_{i=1}^{\infty} \frac{1}{(i-1)!} \times P(0) = P(0)(1+e) = 1$$

(this uses the series expansion for e^x at $x=1$). Then, $P(0) = \frac{1}{e+1}$.

Answer C

3. $P\left[\left|X - \frac{1}{2}\right| \le \frac{1}{4}\right] = P\left[-\frac{1}{4} \le X - \frac{1}{2} \le \frac{1}{4}\right] = P\left[\frac{1}{4} \le X \le \frac{3}{4}\right] = \int_{1/4}^{3/4} 6x(1-x)\,dx = .6875$

$$\implies P\left[\left|X - \frac{1}{2}\right| > \frac{1}{4}\right] = 1 - P\left[\left|X - \frac{1}{2}\right| \le \frac{1}{4}\right] = 0.3125$$

Answer C

4. $P[1 \le X \le 2] = P[X \le 2] - P[X < 1] = F(2) - \lim\limits_{x \to 1^-} F(x) = \frac{11}{12} - \frac{1}{8} = \frac{19}{24}$

Answer E

5. $P[X \le 1] = \int_0^1 (\sqrt{2} - x)\,dx = \sqrt{2} - \frac{1}{2}$, $P[X > 1] = 1 - P[X \le 1] = \frac{3}{2} - \sqrt{2}$.

With 3 independent random variables, X_1,X_2 and X_3, there are 3 ways in which exactly 2 of the X_i's exceed 1 (either X_1,X_2 or X_1,X_3 or X_2,X_3).

Each way has probability $(P[X > 1])^2 \times P[X \le 1] = \left(\frac{3}{2} - \sqrt{2}\right)^2 \left(\sqrt{2} - \frac{1}{2}\right)$

for a total probability of $3 \times \left(\frac{3}{2} - \sqrt{2}\right)^2 \left(\sqrt{2} - \frac{1}{2}\right)$.

Answer E

6. $P\left[Y > \frac{1}{2}\right] = 1 - P\left[Y \le \frac{1}{2}\right] = 1 - P\left[\left(X_1 \le \frac{1}{2}\right) \cap \left(X_2 \le \frac{1}{2}\right) \cap \left(X_3 \le \frac{1}{2}\right)\right]$

$$= 1 - \left(P\left[X \le \frac{1}{2}\right]\right)^3 = 1 - \left[\int_0^{1/2} 3x^2\,dx\right]^3 = 1 - \left(\frac{1}{8}\right)^3 = \frac{511}{512}$$

Answer E

7. $f(x) = F'(x) = -\sum\limits_{k=0}^{3} \frac{kx^{k-1}e^{-x} - x^k e^{-x}}{k!} = e^{-x} \times \sum\limits_{k=0}^{3} \left[\frac{x^k - kx^{k-1}}{k!}\right]$

$$= e^{-x} \times \left[1 + \frac{x-1}{1} + \frac{x^2 - 2x}{2} + \frac{x^3 - 3x^2}{6}\right] = \frac{e^{-x}x^3}{6}$$

Answer C

8. Since $f(x) = 0$ if $x > \theta$, and since $P[X > 1] = \dfrac{7}{8}$, we must conclude that $\theta > 1$.

Then, $P[X > 1] = \displaystyle\int_1^\theta f(x)\,dx = \int_1^\theta \frac{3x^2}{\theta^3}\,dx = 1 - \frac{1}{\theta^3} = \frac{7}{8}$, or equivalently, $\theta = 2$.

Answer E

9. Let us focus on the left-most point p on the disk. Consider two adjacent strips on the floor. Let the interval [0,2] represent the distance as we move across the left strip from left to right. If p is between 0 and 1.75, then the disk lies within the two strips.

If p is between 1.75 and 2, the disk will lie on 3 strips (the first two and the next one to the right). Since any point between 0 and 2 is equally likely as the left most point p on the disk (i.e. uniformly distributed between 0 and 2) it follows that the probability that the disk will touch three strips is $\frac{0.25}{2} = \frac{1}{8}$.

Answer D

10. Since the density function for X is $f(x) = \frac{1}{10}$ for $0 < x < 10$, we can regard X as being positive. Then

$$\begin{aligned} P\left[X + \frac{10}{X} > 7\right] &= P[X^2 - 7X + 10 > 0] = P[(X-5)(X-2) > 0] \\ &= P[X > 5] + P[X < 2] = \frac{5}{10} + \frac{2}{10} = \frac{7}{10} \end{aligned}$$

(since $(t-5)(t-2) > 0$ if either both $t-5,\ t-2 > 0$ or both $t-5,\ t-2 < 0$).

Answer E

11. The survival function $S(y)$ for a random variable can be formulated in terms of the hazard rate function: $S(y) = \exp[-\int_{-\infty}^{y} h(x)\ dx]$.

In this question, $S(5) = 1 - F(5) = 0.16 = \exp\left[-\displaystyle\int_2^5 \frac{z^2}{2x}\ dx\right] = \exp\left[-\frac{z^2}{2}\ln\left(\frac{5}{2}\right)\right]$.

Taking natural log of both sides of the equation results in $-\dfrac{z^2}{2}\ \ln\left(\dfrac{5}{2}\right) = \ln(0.16)$, and solving for z results in $z = 2$.

Answer A

Ready for more practice? Check out GOAL!

GOAL offers additional questions, quizzes, and simulated exams with helpful solutions and tips. Included with GOAL are topic-by-topic instructional videos! Head to ActuarialUniversity.com and log into your account.

Section 17

Aggregate Models - Compound Distributions I

The material in this section relates to "Loss Models", Sections 9.1-9.3.

The suggested time for this section is 2 hours.

17.1 The Compound Distribution Model for Aggregate Claims in One Period of Time

Aggregate claims or aggregate losses refers to the total of all losses that occur in a specified period of time. The components of the **compound model for aggregate claims** are as follows.

(i) N, the **number of claims** or **claim count random variable** whose distribution is called the **claim count distribution** or **frequency distribution**. N is a **discrete**, non-negative integer random variable representing the number of claims or losses or payments that occur in the period. The probability function of N is $P[N = n] = p_n$.

(ii) X, the **single** or **individual loss random variable** whose distribution is called the **severity distribution**. Each time a loss or claim occurs, its amount is assumed to follow the distribution of X. X can be continuous or discrete, or can have a mixed distribution, but is generally assumed to be non-negative. The pf or pdf will be denoted $f_X(x)$ and the cdf is $F_X(x) = P[X \leq x]$. X might also refer to the amount paid on a single loss after any modifications (such as policy deductible, or policy limit) based on insurance coverage.

(iii) If N claims occur in the period, the loss amounts will be $X_1, X_2, \ldots, X_N$, all from the distribution of X. It is assumed that $N, X_1, X_2, \ldots, X_N$ are **mutually independent** random variables.

(iv) $S = X_1 + X_2 + \cdots + X_N$ is the **aggregate loss** per period. (17.1)

S is a **random sum**, and has a **compound distribution**. The terminology "compound distribution" refers to the combination of N (the random variable representing the number of losses in one period of time) with X (the random variable representing the size of each individual loss).

The mean and variance of a compound distribution
To find the mean and variance of S, we use the **conditional expectation rules** that were reviewed earlier in Section 8.2:

If $\boldsymbol{U}$ and $\boldsymbol{W}$ are any random variables, then

$$\boldsymbol{E[U] = E[\,E[U|W]\,]} \tag{17.2}$$

and

$$\boldsymbol{Var[U] = Var[\,E[U|W]\,] + E[\,Var[U|W]\,].} \tag{17.3}$$

We apply these rules with $U = S$ and $W = N$ (the frequency). Applying Equation 17.2, we get the **mean of the compound distribution**

$$\boldsymbol{E[S] = E[\,E[S|N]\,] = E[\,N \times E[X]\,] = E[N] \times E[X].} \tag{17.4}$$

The key point in Equation (17.4) is that $E[S|N] = N \times E[X]$. The reason that this is true can be seen in the following way. Suppose we are given that $N = 3$, so that there are 3 claims, and therefore $S = X_1 + X_2 + X_3$. We see that $E[S] = E[X_1 + X_2 + X_3] = 3 \times E[X]$.

This relationship would work no matter what the actual value of N is, so if we are given the value of N, we know that there are N claims, each with mean $E[X]$, so that $E[S|N] = N \times E[X]$. But then we note that $E[X]$ is a number, so that $N \times E[X]$ is the constant $E[X]$ multiplied by N.

Therefore $E[N \times E[X]\,] = E[N] \times E[X]$ (whenever we multiply a random variable Z by a constant, say c, we know that $E[cZ] = c \times E[Z]$; in this situation, N is the random variable instead of Z, and $E[X]$ is the number c).

Similar reasoning is used in applying Equation (17.3). First, we apply the first part of (17.4). We have seen that $E[S|N] = N \times E[X]$, where N is a random variable and $E[X]$ is a number. Then we use the rule of probability $Var[cZ] = c^2 \times Var[Z]$ if Z is a random variable and c is a number. With N as Z, and $E[X]$ as c, we get

$$\boldsymbol{Var[\,E[S|N]\,] = Var[\;N \times E[X]\,] = Var[N] \times (E[X])^2.} \tag{17.5}$$

Now we apply the second part of (17.4). Similar reasoning to that above shows that $Var[S|N] = N \times Var[X]$. Again, $Var[X]$ is a number, so

$$\boldsymbol{E[\,Var[S|N]\,] = E[\,N \times Var[X]\,] = E[N] \times Var[X].} \tag{17.6}$$

Notation used in the Loss Models book denotes raw moments as $E[Z^k] = \mu'_{Zk}$, or (μ'_k) and central moments as $E[(Z - E[Z])^k] = \mu_{Zk}$ so that $Var[Z] = \mu_{Z2}$. Using this notation, we can express the mean of S as

$$E[S] = E[N] \times E[X] = \mu'_{S1} = \mu'_{N1} \times \mu'_{X1}. \tag{17.7}$$

We can express the **variance of S** as

$$Var[S] = E[N] \times Var[X] + Var[N] \times (E[X])^2. \tag{17.8}$$

This can be expressed in the Loss Models book notation as

$$Var[S] = \mu_{S2} = \mu'_{N1} \times \mu_{X2} + \mu_{N2} \times (\mu'_{X1})^2.$$

It is also possible to show that the **3rd central moment of S** can be formulated as

$$E[(S - E[S])^3] = \mu_{S3} = \mu'_{N1} \times \mu_{X3} + 3\mu_{N2} \times \mu'_{X1} \times \mu_{X2} + \mu_{N3} \times (\mu'_{X1})^3. \tag{17.9}$$

The probability generating function and moment generating function of S

An application of the double expectation rule allows us to formulate the **probability generating function (pgf)** of S in terms of the pgf's of N and X.

$$\boldsymbol{P_S(r) = E[r^S] = P_N(\,P_X(r)\,)}. \tag{17.10}$$

We have seen that for a random variable Z, the moment generating function is closely related to the pgf,

$$M_Z(t) = E[e^{tZ}] = E[(e^t)^Z] = P_Z(e^t).$$

It follows that the **mgf of S** can be formulated as

$$\boldsymbol{M_S(t) = P_S(e^t) = P_N(\,P_X(e^t)\,) = P_N(M_X(t)) = M_N(\,\ln M_X(t)\,)}. \tag{17.11}$$

The pgf and mgf of S have rarely come up on exam questions.

Example 17.1. The number of losses per week N has the following probability function: $p_0 = P[N = 0] = .25$, $p_1 = .5$ and $p_2 = .25$. The size of each loss is uniformly distributed on the interval (0,100). The number of losses and loss sizes are mutually independent. Find the mean and variance of the aggregate loss for one week. Suppose that a policy deductible of 50 is applied to each loss. Find the expected aggregate insurance payment for one week.

Solution:

For the frequency distribution we have

$$E[N] = 1 \times 0.5 + 2 \times 0.25 = 1,\ E[N^2] = 1^2 \times 0.5 + 2^2 \times 0.25 = 1.5$$

so that $Var[N] = 1.5 - 1 = .5$.

For the severity distribution we have $E[X] = 50$ and $Var[X] = \frac{100^2}{12} = \frac{2500}{3}$.

For the aggregate claim distribution, we have $E[S] = E[N] \times E[X] = 1 \times 50 = 50$,
and $Var[S] = E[N] \times Var[X] + Var[N] \times (E[X])^2 = 1 \times \frac{2500}{3} + 0.5 \times 50^2 = \frac{6250}{3}$.
If a deductible of 50 is applied to each loss, then the amount paid by the insurer for a particular loss X is $(X - 50)_+$, with mean $E[(X - 50)_+] = \int_{50}^{100}(x - 50) \times 0.01\,dx = 12.5$.

The expected aggregate payment would be $E[N] \times E[(X - 50)_+] = 12.5$. □

Example 17.2. The number of claims per day, N, has a geometric distribution with mean 2. The size of each claim has an exponential distribution with mean 1000. The number of losses and loss sizes are mutually independent. Find the mean and variance of the aggregate loss for one day.

Solution:

$E[N] = \beta = 2$ and $Var[N] = \beta(1 + \beta) = 2 \times 3 = 6$
$E[X] = \theta = 1000, Var[X] = \theta^2 = 1000^2$
$E[S] = E[N] \times E[X] = 2 \times 1000 = 2000$

$$\begin{aligned} Var[S] &= E[N] \times Var[X] + Var[N] \times (E[X])^2 \\ &= 2 \times 1000^2 + 6 \times 1000^2 = 8 \times 1000^2 \end{aligned}$$ □

17.2 The Compound Poisson Distribution

A compound distribution that appears frequently on exam questions is the **compound Poisson distribution** S. The frequency distribution random variable, N, is **Poisson**, with mean λ. Then

$$\boldsymbol{E[S] = E[N] \times E[X] = \lambda E[X]}, \tag{17.12}$$

and

$$\boldsymbol{Var[S] = E[N] \times Var[X] + Var[N] \times (E[X])^2 = \lambda \times E[X^2]}. \tag{17.13}$$

and for the compound distribution S, Equation 17.9 can be shown to simplify to

$$E[(S - E[S])^3] = \lambda \times E[X^3]. \tag{17.14}$$

Example 17.3. When an individual is admitted to the hospital, the hospital charges have the following characteristics:

(i)

Charges	Mean	Standard Deviation
Room	1000	500
Other	500	300

(ii) the covariance between an individual's Room Charges and Other Charges is 100,000.

An insurer issues a policy that reimburses 100% for Room Charges and 80% for Other Charges. The number of hospital admissions has a Poisson distribution with a parameter (mean) of 4. Determine the variance of the insurer's payout for the policy.

Solution:

The insurer's payout is $S = \sum_{i=1}^{N} X_i$, where N is Poisson with a mean of 4, and $X = R + .8A$ (room plus other charges) is the total hospital charge for one admission.

$Var[S] = E[N] \times Var[X] + Var[N] \times (E[X])^2$

Since N is Poisson, $E[N] = Var[N] = 4$.

$E[X] = E[R] + .8 \times E[A] = 1000 + 0.8 \times 500 = 1400$, and

$$\begin{aligned} Var[X] &= Var[R] + 0.64 \times Var[A] + 2 \times 0.8 \times Cov[R,A] \\ &= 500^2 + 0.64 \times 300^2 + 2 \times 0.8 \times 100{,}000 = 467{,}600. \end{aligned}$$

□

17.3 The Normal Approximation to a Compound Distribution

Many of the aggregate loss problems on the exam involve identifying frequency N and severity X, and then finding $E[S]$ and $Var[S]$, and then applying the **normal approximation** to S to compute probabilities.

Example 17.4. A claim amount distribution is normal with mean 100 and variance 9. The distribution of the number of claims, N, is:

n	$P[N=n]$
0	0.5
1	0.2
2	0.2
3	0.1

Determine the probability that aggregate claims exceed 100 using the normal approximation to the aggregate claim random variable S.

Solution:

$E[N] = .9$

$Var[N] = E[N^2] - (E[N])^2 = 1.9 - 0.9^2 = 1.09$

$E[X] = 100$ and $Var[X] = 9$

$E[S] = E[N] \times E[X] = 0.9 \times 100 = 90$

$Var[S] = E[N] \times Var[X] + Var[N] \times (E[X])^2 = 0.9 \times 9 + 1.09 \times 100^2 = 10{,}908.1$

Applying the normal approximation to S, we get

$$P[S > 100] = P\left[\frac{S-90}{\sqrt{10{,}908.1}} > \frac{100-90}{\sqrt{10{,}908.1}}\right] = P[Z > .096] = .46$$

(from the normal distribution table in the exam tables).

Note that there is no integer correction applied. The severity X is normal, which is continuous, not integer-valued. $\square$

Section 17 Problem Set

Aggregate Models - Compound Distributions I

1. An auto insurer has analyzed claim experience and has chosen as a model for aggregate daily claims a compound Poisson random variable with an average of 20 claims per day. The insurer will pay the full amount of damage when an individual claim is made and the individual claim is modeled to be exponentially distributed with a mean claim amount of \$1000. The company considers offering a policy with a deductible amount of \$250 (i.e., if damage occurs, the company will pay the amount of damage from that individual claim that is in excess of \$250). The company charges an aggregate premium (divided equally among all policy holders) that is one standard deviation larger than the expected aggregate claim. What percentage reduction (nearest 5%) in premium results if all policyholders switch from full coverage to the coverage with \$250 deductible?

   (A) 10% (B) 15% (C) 20% (D) 25% (E) 30%

2. In a group insurance scheme, the claim per member has a normal distribution with a mean of μ and a variance of σ^2. Claims by members of the group are independent. The insurer has determined that for a group of 100 members, a premium of P per member would result in a probability of .95 that aggregate claims would be less than total premium collected. The insurer has also determined that if the group were of size n, then the same premium of P per member would give a probability of .99 that aggregate claims would be less than total premium collected. What is n?

   (A) 120 (B) 140 (C) 160 (D) 180 (E) 200

3. After a study of workdays missed in a year due to illness per employee in a particular company, the following information has been determined:

	Female	Male
Mean	4	7
Variance	30	20

   The number of workdays missed in a year by a given employee has been determined to be independent of the number of workdays missed by any other employee. In a randomly chosen department of N employees, the number of female employees is (discretely) uniformly distributed between 0 and N. If X is the number of workdays missed per year in a randomly selected department with 20 employees, what is $E[X] + \sqrt{Var[X]}$?

   (A) 131 (B) 135 (C) 139 (D) 143 (E) 144

4. Aggregate claims S follow a compound distribution with a Poisson frequency distribution N with mean 1, and a geometric severity distribution X with probability function:

$$P(X = j) = \frac{\beta^j}{(1+\beta)^{1+j}}, \text{ for } j = 0, 1, 2, \ldots$$

   It is found that $E[S \mid S > 0] = \frac{1}{2-2e^{-1/3}}$.
   Determine $E(X)$.

Copyright © 2024 ACTEX Learning

5. S_1 has a compound distribution with frequency N and severity X_1 and S_2 has a compound distribution with frequency N and severity X_2.
N is from the $(a, b, 0)$ class of distributions.
X_1 has an exponential distribution with mean θ,
and the mean and variance of S_1 are 72 and 2268.
X_2 has a uniform distribution on the interval $(0, \theta)$ (same value of θ as X_1),
and the mean and variance of S_2 are 36 and 351.
Find $P(N = 0)$.

6. For an insurance portfolio:

(i) The number of claims has the probability distribution

n	0	1	2	3
p_n	0.1	0.4	0.3	0.2

(ii) Each claim amount has a Poisson distribution with mean 3; and

(iii) The number of claims and claim amounts are mutually independent.

Calculate the variance of aggregate claims.

(A) 4.8 (B) 6.4 (C) 8.0 (D) 10.2 (E) 12.4

7. For an insurance portfolio:

(i) the number of claims has the probability distribution

$n:$	0	1	2	3
$p(n):$	0.4	0.3	0.2	0.1

(ii) each claim amount has a Poisson distribution with mean 4; and

(iii) the number of claims and claim amounts are mutually independent.

Determine the variance of aggregate claims.

(A) 8 (B) 12 (C) 16 (D) 20 (E) 24

8. An insurer issues a portfolio of 100 automobile insurance policies. Of these 100 policies, one-half have a deductible of 10 and the other half have a deductible of 0. The insurance policy pays the amount of damage in excess of the deductible subject to a maximum of 125 per accident. Assume:

(i) the number of automobile accidents per year per policy has a Poisson distribution with mean 0.03 : and

(ii) given that an accident occurs, the amount of vehicle damage has the distribution:

$x:$	30	150	200
$f(x):$	1/3	1/3	1/3

Compute the total amount of claims the insurer expects to pay in a single year.

(A) 270 (B) 275 (C) 280 (D) 285 (E) 290

9. Let S be the aggregate claims for a collection of insurance policies. You are given: G is the premium with a relative security loading of θ (θ is the proportion by which premium is above expected claims); S has a compound Poisson distribution with Poisson parameter λ and severity X; and $R = \frac{S}{G}$ (the loss ratio). Which of the following is an expression for $Var[R]$?

(A) $\dfrac{E\left[X^2\right]}{E[X]} \times \dfrac{1}{1+\theta}$

(B) $\dfrac{E\left[X^2\right]}{(E[X])^2} \times \dfrac{1}{\lambda(1+\theta)}$

(C) $\dfrac{\left(E\left[X^2\right]\right)^2}{(E[X])^2} \times \dfrac{1}{\lambda(1+\theta)}$

(D) $\dfrac{E\left[X^2\right]}{(E[X])^2} \times \dfrac{1}{\lambda(1+\theta)^2}$

(E) $\dfrac{\left(E\left[X^2\right]\right)^2}{(E[X])^2} \times \dfrac{1}{\lambda(1+\theta)^2}$

10. For $S = X_1 + X_2 + \cdots + X_N$

(i) $X_1, X_2, \ldots$ each has an exponential distribution with mean θ;

(ii) the random variables $N, X_1, X_2, \ldots$ are mutually independent;

(iii) N has a Poisson distribution with mean 1.0; and

(iv) $M_S(1.0) = 3.0$ (moment generating function).

Determine θ.

(A) 0.50 (B) 0.52 (C) 0.54 (D) 0.56 (E) 0.58

11. You are the producer for the television show Actuarial Idol. Each year, 1000 actuarial clubs audition for the show. The probability of a club being accepted is 0.20. The number of members of an accepted club has a distribution with mean 20 and variance 20. Club acceptances and the numbers of club members are mutually independent. Your annual budget for persons appearing on the show equals 10 times the expected number of persons plus 10 times the standard deviation of the number of persons. Calculate your annual budget for persons appearing on the show.

(A) 42,600 (B) 44,200 (C) 45,800 (D) 47,400 (E) 49,000

12. For aggregate claims $S = \sum_{i=1}^{N} X_i$, you are given:

(i) the conditional distribution of N, given Λ, is Poisson with parameter Λ;

(ii) Λ has a gamma distribution with $\alpha = 3$ and $\theta = .25$;

(iii) $X_1, X_2, X_3, \ldots$ are identically distributed with $f_X(1) = f_X(3) = 0.5$;

(iv) $N, X_1, X_2, \ldots$ are mutually independent.

Determine $Var[S]$.

(A) 4.4 (B) 4.5 (C) 4.6 (D) 4.7 (E) 4.8

13. A company has a machine that occasionally breaks down. An insurer offers a warranty for this machine. The number of breakdowns and their costs are independent. The number of breakdowns each year is given by the following distribution:

# of breakdowns	Probability
0	50%
1	20%
2	20%
3	10%

The cost of each breakdown is given by the following distribution:

Cost	Probability
1,000	50%
2,000	10%
3,000	10%
5,000	30%

To reduce costs, the insurer imposes a per claim deductible of 1,000. Compute the standard deviation of the insurer's losses for this year.

(A) 1,359 (B) 2,280 (C) 2,919 (D) 3,092 (E) 3,434

14. For aggregate claims $S = \sum_{i=1}^{N} X_i$, you are given:

(i) X_i has distribution $f_X(1) = p, f_X(2) = 1 - p$.

(ii) Λ is a Poisson random variable with parameter $\frac{1}{p}$;

(iii) given $\Lambda = \lambda, N$ is Poisson with parameter λ;

(iv) the number of claims and claim amounts are mutually independent; and

(v) $Var[S] = \frac{19}{2}$.

Determine p.

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

15. Daily claim counts are modeled by the negative binomial distribution with mean 8 and variance 15. Severities have a mean 100 and variance 40,000. Severities are independent of each other and of the number of claims. Let σ be the standard deviation of a day's aggregate losses. On a certain day, 13 claims occurred, but you have no knowledge of their severities. Let σ' be the standard deviation of that day's aggregate losses, given that 13 claims occurred. Calculate $\frac{\sigma}{\sigma'} - 1$.

(A) Less than -7.5% (B) At least -7.5%, but less than 0 (C) 0

(D) More than 0, but less than 7.5% (E) At least 7.5%

16. The number of auto vandalism claims reported per month at Sunny Daze Insurance Company (SDIC) has mean 110 and variance 750. Individual losses have mean 1101 and standard deviation 70. The number of claims and the amounts of individual losses are independent. Using the normal approximation, calculate the probability that SDIC's aggregate auto vandalism losses reported for a month will be less than 100,000.

(A) 0.24 (B) 0.31 (C) 0.36 (D) 0.39 (E) 0.49

17. For an aggregate loss distribution S :

(i) The number of claims has a negative binomial distribution with $r = 16$ and $\beta = 6$.

(ii) The claim amounts are uniformly distributed on the interval (0,8).

(iii) The number of claims and claim amounts are mutually independent.

Using the normal approximation for aggregate losses, calculate the premium such that the probability that aggregate losses will exceed the premium is 5%.

(A) 500 (B) 520 (C) 540 (D) 560 (E) 580

18. You are asked to price a Workers' Compensation policy for a large employer. The employer wants to buy a policy from your company with an aggregate limit of 150% of total expected loss. You know the distribution for aggregate claims is Lognormal. You are also provided with the following:

	Mean	Standard Deviation
Number of Claims	50	12
Amount of individual loss	4,500	3,000

Calculate the probability that the aggregate loss will exceed the aggregate limit.

(A) Less than 3.5%

(B) At least 3.5%, but less than 4.5%

(C) At least 4.5%, but less than 5.5%

(D) At least 5.5%, but less than 6.5%

(E) At least 6.5%

19. For aggregate losses, S:

(i) The number of losses has a negative binomial distribution with mean 3 and variance 3.6.

(ii) The common distribution of the independent individual loss amounts is uniform from 0 to 20.

Calculate the 95-th percentile of the distribution of S using the normal approximation.

(A) 61 (B) 63 (C) 65 (D) 67 (E) 69

20. An insurance policy provides full coverage for the aggregate losses of the Widget Factory. The number of claims for the Widget Factory follows a negative binomial distribution with mean 25 and coefficient of variation 1.2. The severity distribution is given by a lognormal distribution with mean 10,000 and coefficient of variation 3. To control losses, the insurer proposes that the Widget Factory pay 20% of the cost of each loss. Calculate the reduction in the 95-th percentile of the normal approximation of the insurer's loss.

(A) Less than 5%

(B) At least 5%, but less than 15%

(C) At least 15%, but less than 25%

(D) At least 25%, but less than 35%

(E) At least 35%

21. The mean annual number of claims is 103 for a group of 10,000 insureds. The individual losses have an observed mean and standard deviation of 6,382 and 1,781, respectively. The standard deviation of the aggregate claims is 22,874. Calculate the variance of the annual number of claims.

(A) 1.47 (B) 2.17 (C) 4.82 (D) 21.73 (E) 47.23

22. Annual losses for the New Widget Factory can be modeled using a Poisson frequency model with mean of 100 and an exponential severity model with mean of \$10,000. An insurance company agrees to provide coverage for that portion of any individual loss that exceeds \$25,000. Calculate the standard deviation of the insurer's annual aggregate claim payments.

(A) Less than \$36,000

(B) At least \$36,000, but less than \$37,000

(C) At least \$37,000, but less than \$38,000

(D) At least \$38,000, but less than \$39,000

(E) \$39,000 or more

23. The number of accidents reported to a local insurance adjusting office is a Poisson process with parameter $\lambda = 3$ claims per hour. The number of claimants associated with each reported accident follows a negative binomial distribution with parameters $r = 3$ and $\beta = 0.75$. If the adjusting office opens at 8:00 a.m., calculate the variance in the distribution of the number of claimants before noon.

(A) 9 (B) 16 (C) 47 (D) 108 (E) 189

Copyright ©2024 ACTEX Learning

24. An insurance company has two independent portfolios. In Portfolio A, claims occur with a Poisson frequency of 2 per week and severities are distributed as a Pareto with mean 1,000 and standard deviation 2,000. In Portfolio B, claims occur with a Poisson frequency of 1 per week and severities are distributed as a lognormal with mean 2,000 and standard deviation 4,000. Determine the standard deviation of the combined losses for the next week.

(A) Less than 5,500

(B) At least 5,500, but less than 5,600

(C) At least 5,600, but less than 5,700

(D) At least 5,700, but less than 5,800

(E) 5,800 or more

25. The number of annual losses has a Poisson distribution with a mean of 5. The size of each loss has a two-parameter Pareto distribution with $\theta = 10$ and $\alpha = 2.5$. An insurance for the losses has an ordinary deductible of 5 per loss.
Calculate the expected value of the aggregate annual payments for this insurance.

(A) 8 (B) 13 (C) 18 (D) 23 (E) 28

26. In a CCRC, at the start of each month, residents are in one of the following three states: Independent Living (State #1), Temporarily in a Health Center (State #2) or Permanently in a Health Center (State #3). Transitions between states occur at the end of the month. If a resident receives a physical therapy, the number of sessions that the resident receives in a month has a geometric distribution with a mean which depends on the state in which the resident begins the month. The numbers of sessions are independent. The number in each state at the beginning of a given month, the probability of needing physical therapy in the month, and the mean number of sessions received for residents receiving therapy are displayed in the following table:

State #	Number in State	Probability of needing therapy	Mean number of visits
1	400	0.2	2
2	300	0.5	15
3	200	0.3	9

Using the normal approximation for the aggregate distribution, calculate the probability that more than 3000 physical therapy session will be required for the given month.

(A) 0.21 (B) 0.27 (C) 0.34 (D) 0.42 (E) 0.50

27. At the beginning of each round of a game of chance the player pays 12.5. The player then rolls one die with outcome N. The player then rolls N dice and wins an amount equal to the total of the numbers showing on the N dice. All dice have 6 sides and are fair. Using the normal approximation, calculate the probability that a player starting with 15,000 will have at least 15,000 after 1000 rounds.

(A) 0.01 (B) 0.04 (C) 0.06 (D) 0.09 (E) 0.12

28. The following information is known about a consumer electronics store:

- The number of people who make some type of purchase follows a Poisson distribution with a mean of 100 per day.
- The number of televisions bought by a purchasing customer follows a Negative Binomial distribution with parameters $r = 1.1$ and $\beta = 1.0$.

Using the normal approximation, calculate the minimum number of televisions the store must have in inventory at the beginning of each day to ensure that the probability of its inventory being depleted during the day is no more than 1.0%.

(A) Fewer than 138
(B) At least 138, but fewer than 143
(C) At least 143, but fewer than 148
(D) At least 148, but fewer than 153
(E) At least 153

29. You are given:

Number of Claims	Probability	Claim Size	Probability
0	1/5		
1	3/5	25	1/3
		150	2/3
2	1/5	50	2/3
		200	1/3

Claim sizes are independent. Determine the variance of the aggregate loss.

(A) 4,050 (B) 8,100 (C) 10,500 (D) 12,510 (E) 15,612

30. On January 1, 2005, Dreamland Insurance sold 10,000 insurance policies that pay \$100 for each day in 2005 that a policyholder is in the hospital. The following assumptions were used in pricing the policies:

- The probability that a given policyholder will be hospitalized during the year is 0.05. No policyholder will be hospitalized more than one time during the year.
- If a policyholder is hospitalized, the number of days spent in the hospital follows a lognormal distribution with $\mu = 1.039$ and $\sigma = 0.833$.

Using the normal approximation, calculate the premium per policy such that there is a 90% probability that the total premiums will exceed total losses.

(A) Less than 21.20
(B) At least 21.20, but less than 21.50
(C) At least 21.50, but less than 21.80
(D) At least 21.80, but less than 22.10
(E) At least 22.10

Copyright © 2024 ACTEX Learning

31. Computer maintenance costs for a department are modeled as follows:

(i) The distribution of the number of maintenance calls each machine will need in a year is Poisson with mean 3.

(ii) The cost for a maintenance call has mean 80 and standard deviation 200.

(iii) The number of maintenance calls and the costs of the maintenance calls are all mutually independent.

The department must buy a maintenance contract to cover repairs if there is at least a 10% probability that aggregate maintenance costs in a given year will exceed 120% of the expected costs. Using the normal approximation for the distribution of the aggregate maintenance costs, calculate the minimum number of computers needed to avoid purchasing a maintenance contract.

(A) 80 (B) 90 (C) 100 (D) 110 (E) 120

32. Two types of insurance claims are made to an insurance company. For each type, the number of claims follows a Poisson distribution and the amount of each claim is uniformly distributed as follows:

Type of Claim	Poisson Parameter λ for Number of Claims	Range of Each Claim Amount
I	12	(0,1)
II	4	(0,5)

The numbers of claims of the two types are independent and the claim amounts and claim numbers are independent. Calculate the normal approximation to the probability that the total of claim amounts exceeds 18.

(A) 0.37 (B) 0.39 (C) 0.41 (D) 0.43 (E) 0.45

33. A towing company provides all towing services to members of the City Automobile Club. You are given:

(i)

Towing Distance	Towing Cost	Frequency
0-9.99 miles	80	50 %
10-29.99 miles	100	40 %
30+ miles	160	10 %

(ii) The automobile owner must pay 10% of the cost and the remainder is paid by the City Automobile Club.

(iii) The number of towings has a Poisson distribution with mean of 1000 per year.

(iv) The number of towings and costs of individual towings are all mutually independent.

Using the normal approximation for the distribution of aggregate towing costs, calculate the probability that the City Automobile Club pays more than 90,000 in any given year.

(A) 3% (B) 10% (C) 50% (D) 90% (E) 97%

Copyright © 2024 ACTEX Learning

34. The number of claims in a period has a geometric distribution with mean 4. The amount of each claim X follows:

$$P(X = x) = 0.25, \ x = 1, 2, 3, 4$$

The number of claims and the claim amounts are independent. S is the aggregate claim amount in the period. Calculate $F_S(3)$.

(A) 0.27 (B) 0.29 (C) 0.31 (D) 0.33 (E) 0.35

35. Frequency of losses follows a binomial distribution with parameters $m = 1000$ and $q = 0.3$. Severity follows a Pareto distribution with parameters $\alpha = 3$ and $\theta = 500$. Calculate the standard deviation of the aggregate losses.

(A) Less than 7000

(B) At least 7000, but less than 7500

(C) At least 7500, but less than 8000

(D) At least 8000, but less than 8500

(E) At least 8500

36. You are given the following information for a group of policyholders:

- The frequency distribution is negative binomial with $r = 3$ and $\beta = 4$.
- The severity distribution is Pareto with $\alpha = 2$ and $\theta = 2000$.

Calculate the variance of the number of payments if a $500 deductible is introduced.

(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

37. You are given:

- Annual frequency follows a Poisson distribution with mean 0.3.
- Severity follows a normal distribution with $F(100{,}000) = 0.6$.

Calculate that probability that there is at least one loss greater than 100,000 in a year.

(A) Less than 11%

(B) At least 11%, but less than 13%

(C) At least 13%, but less than 15%

(D) At least 15%, but less than 17%

(E) At least 17%

Copyright © 2024 ACTEX Learning

38. An insurance company pays claims at a Poisson rate of 2,000 per year. Claims are divided into three categories: "minor", "major", and "severe", with payment amounts of \$1,000, \$5,000, and \$10,000, respectively. The proportion of "minor" claims is 50%. The total expected claim payments per year is \$7,000,000. What proportion of claims are "severe"?

(A) Less than 11%

(B) At least 11%, but less than 12%

(C) At least 12%, but less than 13%

(D) At least 13%, but less than 14%

(E) 14% or more

39. You are given:

(i) Aggregate losses follow a compound model.

(ii) The claim count random variable has mean 100 and standard deviation 25.

(iii) The single-loss random variable has mean 20,000 and standard deviation 5000.

Determine the normal approximation of the probability that aggregate claims exceed 150% of expected losses.

(A) 0.023 (B) 0.056 (C) 0.079 (D) 0.092 (E) 0.159

40. An actuary has created a compound claims frequency model with the following properties:

(i) The primary distribution is the negative binomial with probability generating function

$$P(z) = [1 - 3(z - 1)]^{-2}.$$

(ii) The secondary distribution is the Poisson with probability generating function

$$P(z) = e^{\lambda(t-1)}.$$

(iii) The probability of no claims equals 0.067.

Calculate λ.

(A) 0.1 (B) 0.4 (C) 1.6 (D) 2.7 (E) 3.1

41. Lucky Tom deposits the coins he finds on the way to work according to a Poisson process with a mean of 22 deposits per month.

5% of the time, Tom deposits coins worth a total of 10.

15% of the time, Tom deposits coins worth a total of 5.

80% of the time, Tom deposits coins worth a total of 1.

The amounts deposited are independent, and are independent of the number of deposits. Calculate the variance of the monthly deposits.

(A) 180 (B) 210 (C) 240 (D) 270 (E) 30

Copyright © 2024 ACTEX Learning

42. Speedy Delivery Company makes deliveries 6 days a week. Accidents involving Speedy vehicles occur according to a Poisson process with a rate of 3 per day and are independent. In each accident, damage to the contents of Speedy's vehicle is distributed as follows:

Amount of damage	Probability
\$0	1/4
\$2,000	1/2
\$8,000	1/4

Using the normal approximation, calculate the probability that Speedy's weekly aggregate damages will not exceed \$63,000.

(A) 0.24 (B) 0.31 (C) 0.54 (D) 0.69 (E) 0.76

43. On Time Shuttle Service has one plane that travels from Appleton to Zebrashire and back and each day. Flights are delayed at a Poisson rate of two per month. Each passenger on a delayed flight is compensated \$100. The numbers of passengers on each flight are independent and distributed with mean 30 and standard deviation 50. (You may assume that all months are 30 days long and that years are 360 days long). Calculate the standard deviation of the annual compensation for the delayed flights.

(A) Less than \$25,000

(B) At least \$25,000, but less than \$50,000

(C) At least \$50,000, but less than \$75,000

(D) At least \$75,000, but less than \$100,000

(E) At least \$100,000

44. You are given:

	Mean	Standard Deviation
Number of claims	8	3
Individual Losses	10,000	3,937

Using the normal approximation, determine the probability that the aggregate loss will exceed 150% of the expected loss.

(A) $\Phi(1.25)$ (B) $\Phi(1.5)$ (C) $1-\Phi(1.25)$ (D) $1-\Phi(1.5)$ (E) $1.5\Phi(1)$

45. The normal approximation is applied to a compound Poisson distribution with Poisson parameter λ. The severity random variable is uniform on the interval $[0,\theta]$. Find the minimum value of λ for which $P(S<0)$ is at most 0.01 (to the nearest 0.1).

(A) 7.0 (B) 7.2 (C) 7.4 (D) 7.6 (E) 7.8

46. Taxicabs leave a hotel with a group of passengers at a Poisson rate $\lambda = 10$ per hour. The number of people in each group taking a cab is independent and has the following probabilities:

Number of People	Probability
1	.60
2	.30
3	.10

Using the normal approximation, calculate the probability that at least 1050 people leave the hotel in a cab during a 72-hour period.

(A) 0.60 (B) 0.65 (C) 0.70 (D) 0.75 (E) 0.80

47. The claims department of an insurance company receives envelopes with claims for insurance coverage at a Poisson rate of $\lambda = 50$ envelopes per week. For any period of time, the number of envelopes and the numbers of claims in the envelopes are independent. The numbers of claims in the envelopes have the following distribution:

Number of Claims	Probability
1	0.20
2	0.25
3	0.40
4	0.15

Using the normal approximation, calculate the 90^{th} percentile of the number of claims received in 1 weeks.

(A) Less than 1700

(B) At least 1700, but less than 1720

(C) At least 1720, but less than 1740

(D) At least 1740, but less than 1760

(E) At least 1760

48. A compound claim distribution is analyzed and it is decided to model the distribution of S by using the geometric distribution for N (claim number) and the gamma distribution for X (claim amount). With this model it is found that $E[S] = 16$, $Var[S] = 328$ and $Var[N] = 20$. An alternative model is proposed in which the distribution of N is Poisson with the same mean as the geometric distribution for N in the first model. The claim amount distribution in the new model is the same gamma distribution as in the old model, and the mean aggregate claims in the new model is the same as in the old model (16). What is the variance of the aggregate claims random variable in the new model?

(A) 18 (B) 36 (C) 54 (D) 72 (E) 90

Copyright © 2024 ACTEX Learning

49. Bob is a street musician and is paid by passersby who appreciate his music and throw money into his hat. Bob estimates that the amount of money he receives in one hour is a compound Poisson distribution with an average frequency of 4 contributions to his hat per hour. The individual amount thrown into his hat is either $1, $2 or $5, and when someone throws money into his hat there is a $\frac{1}{2}$ probability that the amount will be $1, a $\frac{1}{3}$ probability that the amount will be $2, and a $\frac{1}{6}$ probability that the amount will be $5. Bob chooses a location in front of a liquor store, but the liquor store requires that Bob pay the store up to the first $3 Bob receives in the hour. Determine the expected amount that Bob receives in an hour, after paying the store.

(A) Less than $4.00

(B) At least $4.00 but less than $4.50

(C) At least $4.50 but less than $5.00

(D) At least $5.00 but less than $5.50

(E) At least $5.50

Copyright © 2024 ACTEX Learning

Section 17 Problem Set Solutions

1. The number of claims occurring in a day will be denoted N, a Poisson random variable with mean 20. The amount of an individual claim will be denoted X, an exponential random variable with mean 1000. The cost per loss (or cost per individual claim) when a deductible of 250 is applied is $(X-250)_+$. We denote by S the aggregate daily claim when no deductible is applied to individual claims and we denote by S' the aggregate daily claim when the deductible of 250 is applied to individual claims. Both S and S' have compound Poisson distributions.

   $E[S]=E[N]\times E[X]=\lambda\times E[X]=\lambda\times p_1=20\times 1000=20{,}000$ and
   $Var[S]=\lambda\times E[X^2]=\lambda\times p_2=20\times 2{,}000{,}000=40{,}000{,}000$
   (an exponential random variable X with mean $E[X]=\mu$, has second moment $E[X^2]=\mu^2$).
   Also, $E\left[S'\right]=\lambda\times E[(X-250)_+]$ and $Var\left[S'\right]=\lambda\times E[(X-250)_+^2]$.

   We have $E\left[(X-250)_+\right]=\int_{250}^{\infty}(x-250)\times f_X(x)dx=\int_{250}^{\infty}\left[1-F_X(x)\right]dx$
   $$=\int_{250}^{\infty}e^{-0.001x}dx=1000e^{-0.25}=778.80.$$

   Alternatively, we know that if X has an exponential distribution with mean μ then the cost per payment random variable with deductible d is the conditional distribution of $X-d$ given $X>d$ and also has an exponential distribution with mean μ. We saw earlier in this study guide that $E[X-d|X>d]=\frac{E[(X-d)_+}{P[X>d]}$ so that
   $E[(X-d)_+]=E[X-d|X>d]\times P[X>d]=\mu\times e^{-d/\mu}$.
   In this example that results in $E[(X-250)_+]=1000\times e^{-250/1000}=778.80$.

   Following the same approach we saw in the previous paragraph, we have
   $$E[(X-250)_+^2]=E[(X-250)^2|X>250]\times P[X>250].$$

   Since the conditional distribution of $X-250$ given $X>250$ is exponential with mean 1000, this results in $E[(X-250)_+^2]=2\times 1000^2\times e^{-250/1000}=1{,}557{,}602$.

   Note that we can find $E[(X-250)_+]$ and $E[(X-250)_+^2]$ from
   $$E[(X-250)_+]=\int_{250}^{\infty}(x-250)\times f_X(x)dx=\int_{250}^{\infty}\left[1-F_X(x)\right]dx$$
   $$=\int_{250}^{\infty}e^{-0.001x}dx=1000e^{-0.25} \qquad \text{and}$$
   $E[(X-250)_+^2]=\int_{250}^{\infty}(x-250)^2\times e^{-x}\,dx$. This last integral requires either a change of variable or integration by parts.

   It follows that $E\left[S'\right]=20\times(778.80)=15{,}576$ and
   $Var\left[S'\right]=20\times(1{,}557{,}602)=31{,}152{,}031$.

   The (daily) premium with no deductible is $20{,}000+\sqrt{40{,}000{,}000}=26{,}325$, and the premium with the \$250 deductible is $15{,}526+\sqrt{31{,}152{,}031}=21{,}157$. The reduction is 5,168, which is $\frac{5{,}168}{26{,}325}=0.196$, or 19.6%.

   Answer C

2. For a normal random variable with mean μ and standard deviation σ, the 95-th percentile is $\mu + 1.645\sigma$, and the 99-th percentile is $\mu + 0.326\sigma$. The total claim distribution for 100 independent members would be the sum of 100 independent normal random variables each with mean μ and variance σ^2, so the total claim distribution would be normal with mean 100μ and variance $100\sigma^2$, with a similar distribution for the total claim for n members (replace 100 with n).

   Then $100P = 100\mu + 10(1.645)\sigma \to P = \mu + .1645\sigma$ and $nP = n\mu + \sqrt{n}(2.326)\sigma$.
   It follows that $P = \mu + \frac{2.326\sigma}{\sqrt{n}}$. Thus, $\frac{2.326\sigma}{\sqrt{n}} = .1645\sigma \to n = 200$.

   Answer E

3. Let N_F denote the number of females in the department. Then $E[N_F] = 10$ and
   $Var[N_F] = \frac{21^2-1}{12} = 36.6667$. The mean of X can be found from $E[X] = E[E[X \mid N_F]]$.
   If the department has N_F female employees, then there are
   $20 - N_F$ male employees and $E[X \mid N_F] = 4 \times N_F + 7 \times (20 - N_F)$.
   Thus, $E[X] = E[140 - 3 \times N_F] = 140 - 3 \times 10 = 110$.
   The variance of X can be found from $Var[X] = Var[E[X \mid N_F]] + E[Var[X \mid N_F]]$.
   But $Var[E[X \mid N_F]] = Var[140 - 3 \times N_F] = 9 \times Var[N_F] = 330$.
   With N_F female employees and $20 - N_F$ male employees, the variance of the total number of workdays missed in a year for the entire department will be $30 \times N_F + 20 \times (20 - N_F)$ (the sum of the variances for the individuals, since number of workdays missed is independent from one employee to another), thus,
   $E[Var[X \mid N_F]] = E[400 + 10 \times N_F] = 500$.
   Then, $Var[X] = 830$, and $E[X] + \sqrt{Var[X]} = 138.8$.

   Answer C

4. $E(X) = \beta$. S is a non-negative integer-valued random variable with probability function $P[S = k] = g_k$.

$$E[S \mid S > 0] = \sum_{k=1}^{\infty} k \times P[S = k \mid S > 0] = \sum_{k=1}^{\infty} k \times \frac{P[S = k]}{1 - P[S = 0]} = \frac{E[S]}{1 - P[S = 0]}$$

   The severity distribution is geometric with $P[X = 0] = \frac{1}{1+\beta}$.

   $E(S) = E(N) \times E(X) = 1 \times \beta = \beta$

$$P[S = 0] = g_0 = \sum_{k=0}^{\infty} P[S = 0 \mid N = k] \times P[N = k]$$

$$= \sum_{k=0}^{\infty} P[X_1 = 0 \cap X_2 = 0 \cap \cdots \cap X_k = 0 \mid N = k] \times P[N = k]$$

$$= \sum_{k=0}^{\infty} \frac{\left(\frac{1}{1+\beta}\right)^k \times e^{-1} \times 1^k}{k!} = e^{-1} \times e^{1/(1+\beta)} = e^{-\beta/(1+\beta)}$$

$$E[S \mid S > 0] = \frac{E(S)}{1 - P[S = 0]} = \frac{\beta}{1 - e^{-\beta/(1+\beta)}} = \frac{1}{2 - 2e^{-1/3}} = \frac{.5}{1 - e^{-.5/(1+.5)}}$$

   We see that $\beta = .5$.

Copyright ©2024 ACTEX Learning

5. $E(S_1) = E(N) \times E(X_1) = E(N) \times \theta = 72$ (Eq. 1)
$Var(S_1) = E(N) \times Var(X_1) + Var(N) \times [E(X_1)]^2 = [E(N) + Var(N)] \times \theta^2 = 2268$ (Eq. 2)
$E(S_2) = E(N) \times E(X_2) = E(N) \times \frac{\theta}{2} = 36$ (Eq. 3)
$Var(S_2) = E(N) \times Var(X_2) + Var(N) \times [E(X_2)]^2 \, E(N) \times \frac{\theta^2}{12} + Var(N) \times \frac{\theta^2}{4} = 351$ (Eq. 4)
From Equations 2 and 4, we get $2 \times Var(N) \times \theta^2 = 1944$
so that $Var(N) \times \theta^2 = 972$, and then from equation 2 we have $E(N) \times \theta^2 = 1296$.
Now from Equation 1, we get $\theta = \frac{E(N) \times \theta^2}{E(N) \times \theta} = \frac{1296}{72} = 18$.
From this we get $E(N) = 4$ and $Var(N) = 3$.
Since N is in the $(a, b, 0)$ class, it must be either Poisson, Negative Binomial or Binomial. Binomial is the only one of these with $Var(N) < E(N)$, so N is binomial. If the parameters of N are m and q, then $mq = 4$ and $mq(1-q) = 3$, so that $1 - q = \frac{3}{4}$, and $q = \frac{1}{4}$ and $m = 16$.
Then $P(N = 0) = \binom{16}{0} q^0 (1-q)^{16} - \left(\frac{3}{4}\right)^{16} = 0.0100$.

6. $Var[S] = E[N] \times Var[X] + Var[N] \times (E[X])^2$
$E[N] = (0.4) + (2)(0.3) + (3)(0.2) = 1.6, E\left[N^2\right] = (0.4) + \left(2^2\right)(0.3) + \left(3^2\right)(0.2) = 3.4$
and $Var[N] = 3.4 - (1.6)^2 = .84$. $E[X] = Var[X] = 3$
Then, $Var[S] = (1.6)(3) + (0.84)\left(3^2\right) = 12.36$.

Answer E

7. $Var[S] = Var[N](E[X])^2 + E[N]Var[X]$
$E[N] = (0)(0.4) + (1)(0.3) + (2)(0.2) + (3)(0.1) = 1$
$Var[N] = E\left[N^2\right] - (E[N])^2 = 2 - 1^2 = 1$
$E[X] = Var[X] = 4 \rightarrow Var[S] = (1)\left(4^2\right) + (1)(4) = 20$

Answer D

8. Let W denote the annual aggregate claim from one policy with deductible 0. Then W has a compound Poisson distribution, with Poisson parameter .01, and claim-per-accident distribution X. Then $E[W] = 0.03 \times \left[\frac{1}{3}(30 + 125 + 125)\right] = 2.8$.
Let V denote the annual aggregate claim from one policy with deductible 10. Then V has a compound Poisson distribution, with Poisson parameter 0.01, and claim-per-accident distribution X. Then $E[V] = 0.03 \times \left[\frac{1}{3}(20 + 125 + 125)\right] = 2.7$. Since there are 50 policies of each type, the expected annual claim on the insurer is $50 \times [2.8 + 2.7] = 275$.

Answer B

9. For the compound Poisson distribution S with relative security loading $\theta, E[S] = \lambda E[X]$, $Var[S] = \lambda E\left[X^2\right]$, and $G = (1+\theta)E[S] = (1+\theta)\lambda E[X]$
$Var\left[\frac{S}{G}\right] = \frac{1}{G^2} \times Var[S] = \frac{1}{[(1+\theta)\lambda E[X]]^2} \times \left(\lambda E\left[X^2\right]\right) = \frac{E\left[X^2\right]}{(E[X])^2} \times \frac{1}{\lambda(1+\theta)^2}$.

Answer D

10. For the compound Poisson distribution S with Poisson parameter λ and severity random variable X, the moment generating function of S is $M_S(t) = e^{\lambda[M_X(t)-1]}$. Since X has an exponential distribution with mean $\theta, M_X(t) = \frac{1}{1-\theta t}$.
In this case, $3.0 = M_S(1.0) = \exp\left[\frac{1}{1-\theta} - 1\right] \rightarrow \theta = 0.52$.

Answer B

Copyright © 2024 ACTEX Learning

11. The number of people appearing on the show has a compound distribution with frequency N that is binomial with $m = 1000$ and $q = 0.2$, and severity X that has mean $E[X] = 20$ and variance $Var[X] = 20$. The mean and variance of N are $E[N] = (1000)(0.2) = 200$, and $Var[N] = (1000)(0.2)(0.8) = 160$. The mean of the number appearing on the show in one year is $E[S] = E[N] \times E[X] = (1000)(0.2)(20) = 4000$, and the variance is

$$Var[S] = E[N] \times Var[X] + Var[N] \times (E[X])^2 = (200)(20) + (160)(20)^2 = 68{,}000.$$

The standard deviation of S is $\sqrt{68{,}000} = 261$.
The annual budget is $10(4000 + 261) = 42{,}610$.

Answer A

12. When the conditional distribution of N given Λ is Poisson, and the distribution of Λ is gamma with parameters α and θ, the unconditional distribution of N is negative binomial with $r = \alpha$ and $\beta = \theta$. Thus, N has a negative binomial distribution with $r = 3$ and $\beta = 0.25$, so that $E[N] = r\beta = 0.75$ and $Var[N] = r\beta(1+\beta) = 0.9375$. From the information about X we have $E[X] = 2$ $Var[X] = 1$. Then, $Var[S] = E[N] \times Var[X] + Var[N] \times (E[X])^2 = 4.5$.

Answer B

13. After the deductible is imposed on a breakdown, the cost per breakdown X has distribution

Cost	Probability
0	50%
1,000	10%
2,000	10%
4,000	30%

The insurer's losses in a year has a compound distribution with frequency N as indicated in the question and severity X given above. The variance of the insurer's total losses for the year is

$$E[N] \times Var[X] + Var[N] \times (E[X])^2,$$
$$E[N] = (0)(0.5) + (1)(0.2) + (2)(0.2) + (3)(0.1) = .9,$$
$$E\left[N^2\right] = \left(0^2\right)(0.5) + \left(1^2\right)(0.2) + \left(2^2\right)(0.2) + \left(3^2\right)(0.1) = 1.9,$$
$$Var[N] = 1.9 - (.9)^2 = 1.09,$$
$$E[X] = (0)(0.5) + (1000)(0.1) + (2000)(0.1) + (4000)(0.3) = 1{,}500,$$
$$E\left[X^2\right] = \left(0^2\right)(0.5) + \left(1000^2\right)(0.1) + \left(2000^2\right)(0.1) + \left(4000^2\right)(0.3) = 5{,}300{,}000,$$
$$Var[X] = 5{,}300{,}000 - (1{,}500)^2 = 3{,}050{,}000.$$

Then the variance of the insurer's losses for the year is

$$(0.9)(3{,}050{,}000) + (1.09)(1{,}500)^2 = 5{,}197{,}500.$$

The standard deviation is $\sqrt{5{,}197{,}500} = 2{,}280$.

Answer B

14. $Var[S]=E[N]\times Var[X]+Var[N]\times(E[X])^2$
$E[X]=2-p, Var[X]=E\left[X^2\right]-(E[X])^2=4-3p-(2-p)^2=p-p^2$
$E[N]=E[E[N\mid\Lambda]]=E[\Lambda]=\frac{1}{p}$
$Var[N]=E[Var[N\mid\Lambda]]+Var[E[N\mid\Lambda]]=E[\Lambda]+Var[\Lambda]=\frac{1}{p}+\frac{1}{p}=\frac{2}{p}$
$Var[S]=\frac{1}{p}\times(p-p^2)+\frac{2}{p}\times(2-p)^2=\frac{19}{2}$
The resulting quadratic equation in p has roots $p=16,\frac{1}{2}$.
Since $0\le p\le 1$, it follows that $p=\frac{1}{2}$.

Answer E

15. The aggregate loss for the day, say S, has a compound distribution with frequency N and severity X. The variance of S is
$\sigma^2=Var[S]=E[N]\times Var[X]+Var[N]\times(E[X])^2=(8)(40{,}000)+(15)\left(100^2\right)=470{,}000.$
$(\sigma')^2=Var[S\mid N=13]=13\times Var[X]=(13)(40{,}000)=520{,}000.$
Then, $\sigma=\sqrt{470{,}000}=685.6$ and $\sigma'=\sqrt{520{,}000}=721.1$
so that $\frac{\sigma}{\sigma'}-1=-.049\,(-4.9\%)$.

Answer B

16. Aggregate losses per month S follow a compound distribution with frequency N (number of claims) with $E[N]=110$ and $Var[N]=750$, and with severity (individual claim size) Y with $E[Y]=1101$ and $\sqrt{Var[Y]}=70$. Then, $E[S]=E[N]\times E[Y]=121{,}110$, and $Var[S]=E[N]\times Var[Y]+Var[N]\times(E[Y])^2=909{,}689{,}750=30{,}161^2$.
Using the normal approximation, we have

$$P[S<100{,}000]=P\left[\frac{S-121{,}110}{30{,}161}<\frac{100{,}000-121{,}110}{30{,}161}\right]=P[Z<-0.7]=1-\Phi 0(.7)=0.242$$

(Z has a standard normal distribution).

Answer A

17. S has a compound distribution with $E[N]=r\beta=96$ and $Var[N]=r\beta(1+\beta)=672$, and $E[X]=4$ and $Var[X]=\frac{8^2}{12}=5.333$.
$E[S]=E[N]\times E[X]=384$ and $Var[S]=E[N]\times Var[X]+Var[N]\times(E[X])^2=11{,}264$.
We wish to find c so that $P(S>c)=0.05$ using the normal approximation.
$\frac{c-E[S]}{\sqrt{Var[S]}}=1.645\rightarrow c=1.645\sqrt{11{,}264}+384=559$

Answer D

18. The aggregate loss S has a compound distribution. The expected aggregate loss is $E[S]=E[N]\times E[X]=50\times 4{,}500=225{,}000$, and the variance of the aggregate loss is $Var[S]=E[N]\times Var[X]+Var[N]\times(E[X])^2=3{,}366{,}000{,}000$.
The aggregate limit on the policy is $1.5\times E[S]=1.5\times 225{,}000=337{,}500$.
Using the lognormal distribution for S, we have $E[S]=225{,}000=e^{\mu+\frac{1}{2}\sigma^2}$, and $E\left[S^2\right]=Var[S]+(E[S])^2=53{,}991{,}000{,}000=e^{2\mu+2\sigma^2}$.
Then $\mu+\frac{\sigma^2}{2}=12.3239$ and $2\mu+2\sigma^2=24.7121$, so that $\mu=12.292,\sigma=0.2518$.
$P[S>337{,}500]=1-\Phi\left(\frac{\ln(337{,}500)-12.292}{0.2518}\right)=1-\Phi(1.74)=1-.9151$

Answer B

19. S has a compound distribution with negative binomial frequency distribution N and uniform severity distribution Y. Then $E[S] = E[N] \times E[Y] = (3)\left(\frac{0+20}{2}\right) = 30$ and
$Var[S] = E[N] \times Var[Y] + Var[N] \times (E[Y])^2 = (3)\left(\frac{20^2}{12}\right) + (3.6)(10)^2 = 460$
(we have used the fact that the variance of the uniform distribution on the interval (a, b) is $\frac{(b-a)^2}{12}$). Using the normal approximation to S, the 95-th percentile of S is c, where
$P[S \le c] = P\left[\frac{S-30}{\sqrt{460}} \le \frac{c-30}{\sqrt{460}}\right] = \Phi\left(\frac{c-30}{\sqrt{460}}\right) = 0.95$. Therefore, $\frac{c-30}{\sqrt{460}} = 1.645$
so that $c = 65.3$.

Answer C

20. The frequency distribution has mean $E[N] = r\beta = 25$ and variance $Var[N] = r\beta(1+\beta)$. The square of the coefficient of variation of frequency is $\frac{Var[N]}{(E[N])^2} = \frac{r\beta(1+\beta)}{(r\beta)^2} = 1.2^2 = 1.44$. Since $r\beta = 25$, it follows that the variance of N is 900.
In a similar way, we can find the variance of the severity, from $\frac{Var[X]}{(E[X])^2} = 3^2$,
and since $E[X] = 10{,}000$ it follows that $Var[X] = 30{,}000^2$. Then for aggregate loss S
$E[S] = E[N] \times E[X] = (25)(10{,}000) = 250{,}000$, and variance

$$\begin{aligned} Var[S] &= E[N] \times Var[X] + Var[N] \times (E[X])^2 \\ &= (25)\left(30{,}000^2\right) + (900)\left(10{,}000^2\right) = (1125)\left(10{,}000^2\right). \end{aligned}$$

The 95-th percentile of insurance losses before the 20% copayment is C, where
$P[S \le C] = .95$, and using the normal approximation, we get $\frac{C-250{,}000}{\sqrt{1125(10{,}000^2)}} = 1.645$, so that $C = 801{,}750$. After the copayment, the insurance loss is 80% of what it would have been, so $S' = 0.8S$, and $E[S'] = 0.8 \times E[S] = 200{,}000$, and
$Var[S'] = .8^2 \times Var[S] = 0.8^2 \times 1125 \times (10{,}000^2)$.
Using the normal approximation, the 95-th percentile of S' is

$$C' = 1.645 \times 0.8 \times \sqrt{1125\,(10{,}000^2)} + 200{,}000 = 641{,}400.$$

This problem could have been solved much more quickly with the following observation. The 95-th percentile of S is C, which is the solution of $\frac{C-E[S]}{\sqrt{Var[S]}} = 1.645$, so that
$C = E[S] + 1.645\sqrt{Var[S]}$. When we apply the 20% copayment, the insurer loss becomes $S' = 0.8 \times S$, with $E[S'] = 0.8 \times E[S]$ and $\sqrt{Var[S']} = 0.8 \times \sqrt{Var[S]}$, so that the 95-th percentile of S' is $C' = E[S'] + 1.645 \times \sqrt{Var[S']} = 0.8 \times C$. The reduction in the percentile is 20%.

Answer C

21. This is an implied compound distribution S with frequency N and severity X.
We are given $E[N] = 103, E[X] = 6{,}382, Var[X] = 1{,}781^2$ and $Var[S] = 22{,}874^2$.
We use the relationship $Var[S] = E[N] \times Var[X] + Var[N] \times (E[X])^2$ to get
$22{,}874^2 = (103)\left(1781^2\right) + Var[N] \times \left(6382^2\right)$, from which we get $Var[N] = 4.82$

Answer C

22. The insurance applies an ordinary deductible of 25,000 to each loss Y. The severity Y is exponential with mean 10,000. The cost per loss is $(Y-25{,}000)_+$, and the 2nd moment of the cost per loss is $E\left[(Y-25{,}000)_+^2\right]$. This can be found from the integral $\int_{25{,}000}^{\infty}(y-25{,}000)^2 \frac{1}{10{,}000}e^{-y/10{,}000}dy$, which can be found by using the substitution $z=y-25{,}0000$.
The integral becomes $\int_0^{\infty} z^2 \times \frac{1}{10{,}000} \times e^{-(z+25{,}000)/10{,}000}dz$
$= e^{-2.5} \times (\text{2nd moment of an exponential with mean } 10{,}000) = e^{-2.5} \times (2 \times 10{,}000^2)$.

We can also find $E\left[(Y-25{,}000)_+^2\right]$ by using the fact that
$E\left[(Y-25{,}000)_+^2 \mid Y > 25{,}000\right] = \frac{E\left[(Y-25{,}000)_+^2\right]}{1-F_Y(25{,}000)}$, so that
$$E\left[(Y-25{,}000)_+^2\right] = E\left[(Y-25{,}000)_+^2 \mid Y > 25{,}000\right] \times [1-F_Y(25{,}000)]$$
$$= E\left[(Y-25{,}000)_+^2 \mid Y > 25{,}000\right] \times e^{-2.5}.$$
Since the frequency is Poisson, the variance of the annual aggregate claim payments is $E[N] \times (\text{2nd moment of payment per claim}) = 100 \times 2 \times 10{,}000^2 \times e^{-2.5}$, and the standard deviation is $\sqrt{100 \times 2 \times 10{,}000^2 \times e^{-2.5}} = 40{,}518$.

Answer E

23. The number of accidents reported by noon has a Poisson distribution with a mean of 12 (3 accidents per hour, on average, for 4 hours from 8 AM to noon). This is the frequency N. The severity X for an accident is the number of claimants associated with the accident. X has a negative binomial distribution with $r=3$ and $\beta=.75$. The total number of claimants before noon is $S = X_1 + X_2 + \cdots + X_N$, a compound distribution (actually, a compound Poisson distribution). The mean and the variance of N are $E[N] = Var[N] = 12$, and the mean and variance of X are $E[X] = r\beta = 2.25$ and $Var[X] = r\beta(1+\beta) = 3.9375$.
Then $Var[S] = E[N] \times Var[X] + Var[N] \times (E[X])^2 = 108$.

Answer D

24. Since the portfolios have compound Poisson distributions for aggregate losses, the variance of aggregate losses for portfolio A is
$Var[S_A] = E[N_A] \times E[X_A^2] = E[N_A] \times \left(Var[X_A] + (E[X_A])^2\right) = 2\left(2000^2 + 1000^2\right)$.
Similarly, $Var[S_B] = E[N_B] \times E\left[X_B^2\right]$
$$= E[N_B] \times \left(Var[X_B] + (E[X_B])^2\right) = 1\left(4000^2 + 2000^2\right).$$
Since the portfolios are independent, we have
$Var[S_A + S_B] = Var[S_A] + Var[S_B] = 10{,}000{,}000 + 20{,}000{,}000 = 30{,}000{,}000$.
The standard deviation is $\sqrt{30{,}000{,}000} = 5{,}477$.

Answer A

25. The expected cost per loss with a deductible of 5 per loss is
$$E[(X-5)_+] = E[X] - E[X \wedge 5] = \frac{\theta}{\alpha-1} - \left(\frac{\theta}{\alpha-1}\right)\left[1 - \left(\frac{\theta}{5+\theta}\right)^{\alpha-1}\right]$$
$$= \frac{10}{2.5-1} \times \left(\frac{10}{5+10}\right)^{1.5} = 3.63.$$
The expected aggregate payment is $E[N] \times E[(X-5)_+] = 5 \times 3.63 = 18.1$.

Answer C

26. For a resident in state 1, let N_1 be the number of therapy visits needed in a month. N_1 is a mixture of the constant 0 with probability. 8 (no therapy needed) and geometric X_1 with mean 2 with probability 0.2. The first and second moments of a geometric distribution with mean β are β and $\beta + 2\beta^2$.
The first and second moments of X_1 are 2 and 10. The first and second moments of N_1 are $0.2 \times 2 = 0.4$ and $0.2 \times 10 = 2$ so the variance of N_1 is $2 - 0.4^2 = 1.84$.
An alternative calculation for the variance of Y is based on the following rule: if Y is a mixture of 0 with probability $1-q$ and W with probability q, the mean of Y is $qE[W]$, and the variance of Y is $qE\left[W^2\right] - (qE[W])^2 = qVar[W] + q(1-q)(E[W])^2$.
Using this rule with $Y = N_1$ and W geometric with mean 2,
$Var\left[N_1\right] = (0.2)(2)(1+2) + (0.2)(0.8)(2)^2 = 1.84$.
In a similar way, if N_2 is the number of therapy visits needed in a month for a resident in state 2, then the mean of N_2 is $0.5 \times 15 = 7.5$, and the second moment of N_2 is $0.5 \times 465 = 232.5$, and the variance of N_2 is 176.25.
Again, in a similar way, the mean and second moment of N_3 are 2.7 and $0.3 \times 171 = 51.3$, and the variance of N_3 is 44.01. The mean number of visits for all residents in all states is $(400)(0.4) + (300)(7.5) + (200)(2.7) = 2950$, and because of independence of residents, the variance of the number of visits needed for all residents in all states is
$(400)(1.84) + (300)(176.25) + (200)(44.01) = 62{,}413$.
Applying the normal approximation to N (an integer), the total number of visits by all residents, we get

$$P[N > 3000] = P\left[\frac{N-2950.5}{\sqrt{62{,}413}} > \frac{3000-2950.5}{\sqrt{62{,}413}}\right] = 1 - \Phi\left(\frac{3000-2950.5}{\sqrt{62{,}413}}\right) = 1 - \Phi(0.2) = .42.$$

Answer D

27. Suppose S is the amount won in a single play of the game. S has a compound distribution with "frequency" N (number of dice tossed) and "severity" Y (number turning up on an individual die). N is an integer from 1 to 6, with $P[N=n] = \frac{1}{6}$ for $n = 1,2,\ldots,6$.
$E[N] = \frac{7}{2}$, $Var[N] = \frac{6^2-1}{12} = \frac{35}{12}$.
Y has the same distribution as N, since Y is the outcome of a single die toss.
Then $E[S] = E[N] \times E[Y] = \left(\frac{7}{2}\right)\left(\frac{7}{2}\right) = \frac{49}{4}$, and
$Var[S] = E[N] \times Var[Y] + Var[N] \times (E[Y])^2 = \left(\frac{7}{2}\right)\left(\frac{35}{12}\right) + \left(\frac{35}{12}\right)\left(\frac{7}{2}\right)^2 = 45.9375$.
$X = \sum_{i=1}^{1000} S_i$ is the total amount won in 1000 rounds of the game.
Then $E[X] = 1000 \times E[S] = 12{,}250$ and since the rounds are independent of one another,
$Var[S] = 1000 \times Var[X] = 45{,}937.5$.
Since the cost is 12.5 for each play of the game, the cost for 1000 plays is 12,500. In order for the player to have at least as much as they started with after 1000 plays of the game, he must win a total of at least 12,500. We use the normal approximation with continuity correction:

$$\begin{aligned} P[X \geq 12{,}499.5] &= P\left[\frac{X-12{,}250}{\sqrt{45{,}937.5}} \geq \frac{12{,}499.5-12{,}250}{\sqrt{45{,}937.5}}\right] = P[Z \geq 1.17] \\ &= 1 - \Phi(1.17) = 1 - 0.879 = 0.12 \end{aligned}$$

Answer E

28. The number of televisions purchased in a day has a compound distribution with Poisson frequency N with a mean of 100, and Negative Binomial severity X with parameters $r = 1.1$ and $\beta = 1.0$. The total number of television purchased in one day, say S, has mean
$E[S] = E[N] \times E[X] = 100(1.1)(1.0) = 110$ and variance
$Var[S] = E[N] \times Var[X] + Var[N] \times (E[X])^2 = 100(1.1)(1.0)(2.0) + 100 \times 1.1^2 = 341.$
We wish to find the number of televisions k so that, using the normal approximation,
$P[S > k] = .01$. This probability can be written as $P\left[\frac{S-110}{\sqrt{341}} > \frac{k-110}{\sqrt{341}}\right] = .01$, so that $\frac{k-110}{\sqrt{341}}$ is the 99-th percentile of the standard normal distribution. Therefore, $\frac{k-110}{\sqrt{341}} = 2.326$ and $k = 153$.

Answer E

29. We formulate the variance of S by conditioning on N.
$Var[S] = Var[E[S \mid N]] + E[Var[S \mid N]]$.
We work from the inside out. First, if $N = 0$ then $E[S \mid N = 0] = 0$,
if $N = 1$ then $E[S \mid N = 1] = (25)\left(\frac{1}{3}\right) + (150)\left(\frac{2}{3}\right) = \frac{325}{3}$, and
if $N = 2$ then $E[S \mid N = 2] = 2\left[(50)\left(\frac{2}{3}\right) + (200)\left(\frac{1}{3}\right)\right] = 200$.

Therefore, $E[S \mid N] = \begin{cases} 0 & \text{prob. } 1/5 \\ \frac{325}{3} & \text{prob. } 3/5, \\ 200 & \text{prob. } 1/5 \end{cases}$ and

$$Var[E[S \mid N]] = \left[(0)^2\left(\tfrac{1}{5}\right) + \left(\tfrac{325}{3}\right)^2\left(\tfrac{3}{5}\right) + (200)^2\left(\tfrac{1}{5}\right)\right] - \left[(0)\left(\tfrac{1}{5}\right) + \left(\tfrac{325}{3}\right)\left(\tfrac{3}{5}\right) + (200)\left(\tfrac{1}{5}\right)\right]^2 = 4017.$$

If $N = 0$ then $Var[S \mid N = 0] = 0$ (no claims),
if $N = 1$ then $Var[S \mid N = 1] = (150 - 25)^2\left(\frac{1}{3}\right)\left(\frac{2}{3}\right) = \frac{31{,}250}{9}$, and
if $N = 2$ then $Var[S \mid N = 2] = 2\left[(200 - 50)^2\left(\frac{2}{3}\right)\left(\frac{1}{3}\right)\right] = 10{,}000$

(2 independent claims results in the variance of each claim being counted).

Therefore $Var[S \mid N] = \begin{cases} 0 & \text{prob. } 1/5 \\ \frac{31{,}250}{9} & \text{prob. } 3/5, \text{ and} \\ 10{,}000 & \text{prob. } 1/5 \end{cases}$

$E[Var[S \mid N]] = (0)\left(\frac{1}{5}\right) + \left(\frac{31{,}250}{9}\right)\left(\frac{3}{5}\right) + (10{,}000)\left(\frac{1}{5}\right) = 4083.$

Then, $Var[S] = 4{,}017 + 4{,}083 = 8{,}100$.

Answer B

30. For a single policy, the number of hospitalizations in the year, N, is a Bernoulli random variable with $P[N=0]=0.95$ and $P[N=1]=0.05$. When a hospitalization takes place, the cost is $Y=100X$, where X has a lognormal distribution with $\mu=1.039$ and $\sigma=0.833$. The total cost, S, in 2005 for a single policy is a compound distribution based on frequency N and severity Y. The mean and variance of N are $E[N]=0.05, Var[N]=0.05\times 0.95=0.0475$.
The mean and second moment of Y are
$E[Y]=100E[X]=100e^{\mu+\frac{1}{2}\sigma^2}=100e^{1.039+\frac{1}{2}(0.833^2)}=400$,
$E\left[Y^2\right]=100^2E\left[X^2\right]=10{,}0000e^{2\mu+2\sigma^2}=10{,}000e^{2.078+2(.833^2)}=320{,}013$.
The variance of Y is $E\left[Y^2\right]-(E[Y])^2=160{,}013$.
The mean and variance of S are $E[S]=E[N]\times E[Y]=20$ and
$Var[S]=E[N]\times Var[Y]+Var[N]\times(E[Y])^2=15{,}601$.
The aggregate cost for 10,000 policies is $W=S_1+S_2+\cdots+S_{10{,}000}$.

The mean and variance of W are $E[W]=10{,}000E[S]=10{,}000\times 20$ and
$Var[W]=10{,}000\times Var[S]=10{,}000\times 15{,}601$ (we assume that the policies are independent, so the variance of the sum does not involve any covariances).

We apply the normal approximation to find the fund size F which satisfies the probability $P[W\le F]=0.9$. The probability can be written as $P\left[\frac{W-E[W]}{\sqrt{Var[W]}}\le\frac{F-10{,}000\times 20}{\sqrt{10{,}000\times 15{,}601}}\right]=0.9$.

Then $\frac{F-10{,}000\times 20}{\sqrt{10{,}000\times 15{,}601}}$ is the 90-th percentile of the standard normal distribution, which, from the table, is 1.282. Solving for F results in $F=216{,}013$.
This is the premium for 10,000 policies. The premium per policy is $\frac{216{,}013}{10{,}000}=21.60$.
This solution interprets the question as meaning that the policy will pay \$100 for each day for a hospitalization that occurs in 2005 (so that if the hospitalization occurs on Dec. 31,2005, all days are still covered, even though the stay may carry over to 2006).

Answer C

31. The annual maintenance cost for one computer, say W, has a compound Poisson distribution with a Poisson frequency distribution N with a mean of 3 and severity distribution X with mean 80 and variance 40,000. Then $E[W]=E[N]\times E[X]=(3)(80)=240$ and
$Var[W]=E[N]\times Var[X]+Var[N]\times(E[N])^2=(3)(40{,}000)+(3)(80)^2=139{,}200$.

With n computers, the annual maintenance cost will be $T_n=W_1+\cdots+W_n$, with mean $E\left[T_n\right]=240n$ and variance $Var\left[T_n\right]=139{,}200n$.

The probability that annual maintenance costs exceed 120% of expected costs is
$P\left[T_n>1.2E\left[T_n\right]\right]=P\left[T_n>288n\right]$.

Applying the normal approximation, this probability is
$P\left[\frac{T_n-240n}{\sqrt{139{,}200n}}>\frac{288n-240n}{\sqrt{139{,}200n}}\right]=1-\Phi(0.1287\sqrt{n})$.

A maintenance contract will be avoided if this probability is less than 10%.
The probability is less than 10% if $\Phi(0.1287\sqrt{n})>.9$.

From the normal table, the 90-th percentile of the standard normal distribution is 1.282, so that the maintenance contract will be avoided if $.1287\sqrt{n}>1.282$, or equivalently, if $n>99.3$.

Answer C

32. The aggregate claim distribution is the sum of two compound Poisson distributions.
S_I has mean and variance $E[S_I] = 12 \times 0.5 = 6$, and $Var[S_I] = 12 \times \frac{1}{3} = 4$.
We use the rule for a compound Poisson distribution that $Var[S] = E[N] \times E[X^2]$, and for the uniform distribution on (a,b) the second moment is $\frac{b^3-a^3}{3(b-a)}$.

S_{II} has mean $E[S_{II}] = 4 \times 2.5 = 10$, and variance $Var[S_I] = 4 \times \frac{125}{3\times 5} = 33.33$.
Since the two distributions are independent, the variance of W is
$Var[W] = Var[S_I] + Var[S_{II}] = 4 + 33.33 = 37.33$.

The mean of W is 16. Using the normal approximation to W, we have
$P[W > 18] = P\left[\frac{W-16}{\sqrt{37.33}} > \frac{18-16}{\sqrt{37.33}}\right] = P[Z > 0.33] = 1 - \Phi(0.33) = 0.37$.

Answer A

33. The annual cost to the Club S has a compound Poisson distribution with Poisson mean $\lambda = 1000$, and 3-point severity distribution $X = \begin{cases} 72 & \text{prob. } 0.5 \\ 90 & \text{prob. } 0.4 \\ 144 & \text{prob. } 0.1 \end{cases}$.
This is 90% of the towing cost.

Then for compound Poisson S we have
$E[S] = \lambda \times E[X] = 1000 \times [(72)(0.5) + (90)(0.4) + (144)(0.1)] = 86{,}400$,
$Var[S] = \lambda E[X^2]$ (valid only for compound Poisson) $= 7{,}905{,}600$.
Applying the normal approximation with continuity correction to S (since X is an integer), we have

$$P[S > 90{,}000.5] = P\left[\frac{S-86{,}400}{\sqrt{7{,}905{,}600}} > \frac{90{,}000.5-86{,}400}{\sqrt{7{,}905{,}600}}\right] = 1 - \Phi(1.28) = .10.$$

Answer B

34. S has a compound distribution with frequency N and severity X.
There are a couple of approaches that can be taken to solve this problem.
The first approach uses the relationship $F_S(y) = \sum_{n=0}^{\infty} P[S \le y \mid N = n] \times P[N = n]$,
where N is the frequency distribution which is geometric with mean $\beta = 4$ in this case.
The probability function of N is $P[N = n] = \frac{4^n}{5^{n+1}}, n = 0,1,2,\ldots$.
Since X must be at least 1, $P[S \le 3 \mid N = n] = 0$ for $n \ge 4$ (if there are 4 or more claims, S must be at least 4). Therefore,
$F_S(3) = \sum_{n=0}^{\infty} P[S \le 3 \mid N = n] \times P[N = n] = \sum_{n=0}^{3} P[S \le 3 \mid N = n] \times P[N = n]$.
$P[S \le 3 \mid N = 0] = 1$ since if there are no claims, S must be 0.
$P[S \le 3 \mid N = 1] = \frac{3}{4}$ since 3 of the 4 claim amounts are ≤ 3.
$P[S \le 3 \mid N = 2] = \frac{3}{16}$ since each pair X_1, X_2 of claims has $\frac{1}{4} \times \frac{1}{4} = \frac{1}{16}$ probability of occurring, and 3 of the 16 pairs results in $S \le 3$ (the 1,1 pair, the 1,2 pair and the 2,1 pair).
$P[S \le 3 \mid N = 3] = \frac{1}{64}$ since if there are 3 claims, the only way the total is ≤ 3 is if all three are for amount 1 each; this probability is $\frac{1}{4} \times \frac{1}{4} \times \frac{1}{4}$. Then

$$\begin{aligned} F_S(3) &= \sum_{n=0}^{3} P[S \le 3 \mid N = n] \times P[N = n] \\ &= (1)\left(\frac{1}{5}\right) + \left(\frac{3}{4}\right)\left(\frac{4}{25}\right) + \left(\frac{3}{16}\right)\left(\frac{16}{125}\right) + \left(\frac{1}{64}\right)\left(\frac{64}{625}\right) = 0.3456. \end{aligned}$$

An alternative combinatorial approach is to find
$F_S(3) = P[S=0] + P[S=1] + P[S=2] + P[S=3]$.
$P[S=0] = P[N=0] = \frac{1}{5}$, since $S=0$ only if $N=0$
$P[S=1] = P[N=1] \times P[X=1] = \frac{4}{25} \times \frac{1}{4} = \frac{1}{25}$
(one claim of amount 1 is the only combination which results in $S=1$).
$P[S=2] = P[N=1] \times P[X=2] + P[N=2] \times (P[X=1])^2 = \frac{4}{25} \times \frac{1}{4} + \frac{16}{125} \times \left(\frac{1}{4}\right)^2 = \frac{6}{125}$
(one claim of amount 2, or two claims each of amount 1).

$P[S=3]$
$= P[N=1] \times P[X=3] + P[N=2] \times 2P[X=1] \times P[X=2] + P[N=3] \times (P[X=1])^3$
$= \left(\frac{4}{25}\right)\left(\frac{1}{4}\right) + \left(\frac{16}{125}\right)(2)\left(\frac{1}{4}\right)\left(\frac{1}{4}\right) + \left(\frac{64}{625}\right)\left(\frac{1}{4}\right)^3 = \frac{36}{625}$

Then $F_S(3) = \frac{1}{5} + \frac{1}{25} + \frac{6}{125} + \frac{36}{625} = \frac{216}{625} = 0.3456$.

Answer E

35. Aggregate loss S has a compound distribution S with $E[N] = 1000 \times 0.3 = 300$, $Var[N] = 1000 \times 0.3 \times 0.7 = 210$ for frequency, and $E[X] = \frac{500}{3-1} = 250$, $E\left[X^2\right] = \frac{2 \times 500^2}{(3-1)(3-2)} = 250{,}000$ and $Var[X] = 250{,}000 - 250^2 = 187{,}500$ for severity. $Var[X] = E[N] \times Var[X] + Var[N] \times (E[X])^2 = 69{,}375{,}000$ and $\sqrt{Var[X]} = 8329$.

Answer D

36. The probability that a payment is above the deductible of 500 is $1 - F(500) = \left(\frac{2000}{500+2000}\right)^2 = 0.64$. The negative binomial distribution satisfies the "infinite divisibility" property, which implies that the number of losses that are above the deductible, say M, is also negative binomial with the same r, and with "new β"="old β" $\times 0.64 = 2.56$. Then $Var[M] = 3 \times 2.56 \times 3.56 = 27.34$.

Answer A

37. The Poisson distribution also satisfies the infinite divisibility property. The number of losses greater than 100,000, say M, has a Poisson distribution with "new λ"= "old λ"$\times P[X > 100{,}000] = 0.3 \times 0.4 = 0.12$. Probability of at least one loss greater than 100,000 in a year is $P[M \geq 1] = 1 - P[M=0] = 1 - e^{-0.12} = 0.113$.

Answer B

38. Let us suppose that the proportion of major claims is p. Then the proportion of severe claims must be $1 - 0.5 - p = 0.5 - p$ (since the three proportions must add to 1).
We expect $2000 \times .5 = 1000$ minor claims per year, $2000p$ major claims per year, and $2000(0.5-p) = 1000 - 2000p$ severe claims per year.
The total expected claim payments per year will be

$$1{,}000(1{,}000) + 2{,}000p(5{,}000) + (1000 - 2000p)(10{,}000) = 11{,}000{,}000 - 10{,}000{,}000p.$$

We are told that this total is 7,000,000. Therefore $11{,}000{,}000 - 10{,}000{,}000p = 7{,}000{,}000$ from which it follows that $p = 0.4$. The proportion of severe claims is $0.5 - p = 0.1$.

Answer A

39. Aggregate losses will be represented by the random variable S.
We wish to find $P[\,S > 1.5E[S]\,]$, using the normal approximation.
This will be $P\left[\frac{S-E[S]}{\sqrt{Var[S]}} > \frac{1.5E[S]-E[S]}{\sqrt{Var[S]}}\right] = 1-\Phi\left(\frac{.5E[S]}{\sqrt{Var[S]}}\right)$.
For the compound distribution, we have $E[S] = E[N] \times E[X] = 100 \times 20{,}000 = 2{,}000{,}000$. ($N$ is the claim count, and X is the single-loss). We also have
$Var[S] = E[N] \times Var[X] + Var[N] \times (E[X])^2 = 100 \times 5000^2 + \times 25^2 \times 20{,}000^2 = 2.525 \times 10^{11}$.
The probability is $1-\Phi\left(\frac{.5E[S]}{\sqrt{Var[S]}}\right) = 1-\Phi\left(\frac{.5(2{,}000{,}000)}{\sqrt{2.525\times 10^{11}}}\right) = 1-\Phi(1.99)$
$= 1 - 0.9767 = 0.0233$.
Answer A

40. If K is a discrete non-negative integer random variable with probability generating function $P_K(z)$, then $P[K=0] = P_K(0)$. Suppose that N is the primary distribution (negative binomial) and M is the secondary (Poisson), and let K denote the compound claims frequency model. Then the probability generating function of K is
$$P_K(z) = P_N\left(P_M(z)\right) = \left[1-3\left(e^{\lambda(z-1)}-1\right)\right]^{-2}.$$
Then, $0.067 = P_K(0) = \left[1-3\left(e^{\lambda(0-1)}-1\right)\right]^{-2}$. Solving this equation for λ results in $\lambda = 3.1$.
Answer E

41. The number of deposits per month (frequency), N, is Poisson with mean 22. The amount of each deposit (severity), X, has a 3-point distribution,
$$P[X=1] = 0.8, \quad P[X=5] = 0.15, \quad P[X=10] = 0.05.$$
Aggregate monthly deposits, S, has a compound Poisson distribution with variance
$$Var[S] = \lambda \times E\left[X^2\right] = (22)[(1)(0.8) + (25)(0.15) + (100)(0.05)] = 210.$$
Answer B

42. In 6 days the expected number of accidents is 18 (6 days and an average of 3 accidents per day). The aggregate damages for one week, say S, has a compound Poisson distribution with $\lambda = 18$ (expected number of claims per week). The severity distribution X has mean
$E[X] = (0)\left(\frac{1}{4}\right) + (2{,}000)\left(\frac{1}{2}\right) + (8{,}000)\left(\frac{1}{4}\right) = 3{,}000$, and the second moment is
$E\left[X^2\right] = (0^2)\left(\frac{1}{4}\right) + (2{,}000^2)\left(\frac{1}{2}\right) + (8{,}000^2)\left(\frac{1}{4}\right) = 18{,}000{,}000$.
The mean and variance of S are $E[S] = \lambda E[X] = 54{,}000$ and
$Var[S] = \lambda E\left[X^2\right] = 324{,}000{,}000$. X is an integer, so we apply the normal approximation with continuity correction. We get
$$P[S \le 63{,}000.5] = P\left[\frac{S-54{,}000}{\sqrt{324{,}000{,}000}} \le \frac{63{,}000.5-54{,}000}{\sqrt{324{,}000{,}000}}\right] = \Phi\left(\frac{63{,}000.5-54{,}000}{\sqrt{324{,}000{,}000}}\right) = \Phi(.50) = .6915.$$
Answer D

43. The number of passengers compensated in one year S has a compound distribution.
The frequency is N, the number of delayed flights in one year, which is Poisson with mean 24 (2 per month for 12 months).
The severity Y is the number of passengers on a delayed flight.
We have $E[N] = Var[N] = 24, E[Y] = 30$ and $Var[Y] = 50^2 = 2500$.
Since the frequency has a Poisson distribution, it follows that $Var[S] = E[N] \times E\left[Y^2\right]$.
From the given information we have $E\left[Y^2\right] = Var[Y] + (E[Y])^2 = 2500 + 30^2 = 3400$.
Then, $Var[S] = (24)(3400) = 81{,}600$.

Alternatively, $Var[S] = E[N] \times Var[Y] + Var[N] \times (E[Y])^2 = 81{,}600$.
Each passenger on a delayed flight receives \$100 in compensation. The total compensation paid in one year is $100S$, and the variance is $Var[100S] = 100^2 \times 81{,}600$.
The standard deviation of annual compensation is $100\sqrt{81{,}600} = 28{,}566$.

Answer B

44. It must be assumed that number of claims N and individual loss amounts X are mutually independent. The aggregate loss S has mean and variance
$E[S] = E[N] \times E[X] = 8 \times 10{,}000 = 80{,}000$ and
$Var[S] = Var[N] \times (E[X])^2 + E[N] \times Var[X] = 1{,}023{,}999{,}752$.

Using the normal approximation,
$P[\,S > 1.5E[S]\,] = P\left[\frac{S-E[S]}{\sqrt{Var[S]}} > \frac{1.5E[S]-E[S]}{\sqrt{Var[S]}}\right] = P[Z > 1.25] = 1 - \Phi(1.25)$.

Answer C

45. $E[S] = \frac{\lambda\theta}{2}$ and $Var[S] = \frac{\lambda\theta^2}{3}$. According to the normal approximation, $Z = \frac{S-\frac{\lambda\theta}{2}}{\sqrt{\frac{\lambda\theta^2}{3}}}$ has a standard normal distribution.
If $P[S < 0] = P\left[Z < -\frac{\lambda\theta/2}{\sqrt{\lambda\theta^2/3}}\right] = 0.01$ then $\frac{\lambda\theta/2}{\sqrt{\lambda\theta^2/3}} = \frac{\sqrt{3\lambda}}{2} = 2.326$, so that $\lambda = 7.2$.
If λ is any smaller than 7.2, the probability is greater than .01.

Answer B

46. $S =$ number leaving by in a one hour period has compound Poisson distribution, with $\lambda = 10$ (average number of cabs (claims) per hour). The number per cab ("claim amount") X is 1,2 or 3, with probabilities $0.6, 0.3$ and 0.1, respectively. The expected number of people leaving per hour (expected aggregate claims per period) is

$$E[S] = \lambda E[X] = 10[(1)(.6) + (2)(.3) + (3)(.1)] = 15,$$

and the variance of the number leaving per hour is

$$Var[S] = \lambda E\left[X^2\right] = 10\left[(1)^2(.6) + (2)^2(.3) + (3)^2(.1)\right] = 27.$$

In 72 (independent) hours, W, the number of people leaving by cab has a mean of $72 \times 15 = 1080$ and variance $72 \times 27 = 1944$. If W is assumed to be approximately normal, then since W is an integer
$P[W \geq 1050] = P[W \geq 1049.5] = P\left[\frac{W-1080}{\sqrt{1944}} \geq \frac{1049.5-1080}{\sqrt{1944}}\right]$
$= P[Z \geq -0.69] = P[Z \leq 0.69] = .75$ (Z has a standard normal distribution).

Answer D

47. The number of claims received in one week has a compound Poisson distribution with $E[N] = 50$, and $Y = 1, 2, 3, 4$ with the given probabilities.

The expected amount of claims received in one week is
$E[N] \times E[Y] = 50 \times [(1)(0.2) + (2)(0.25) + (3)(0.4) + (4)(0.15)] = 125.$

The variance of the amount of claims received in one week is
$E[N] \times E\left[Y^2\right] = 50 \times \left[\left(1^2\right)(0.2) + \left(2^2\right)(0.25) + \left(3^2\right)(0.4) + \left(4^2\right)(0.15)\right] = 360.$

The expected number of W claims in 13 weeks is $(13)(125) = 1{,}625$, and the variance of the number of claims in 13 weeks is $13 \times 360 = 4{,}680$.
The 90-th percentile is c, where $P[W \le c] = P\left[\frac{W-1625}{\sqrt{4680}} \le \frac{c-1625}{\sqrt{4680}}\right] = \Phi\left(\frac{c-1625}{\sqrt{4680}}\right) = .90.$
Using the normal approximation $\frac{c-1625}{\sqrt{4680}} = 1.282 \to c = 1713.$

Answer B

48. With a geometric distribution for $N, Var[N] = \beta(1+\beta) = 20 \to \beta = 4$ (ignore the negative root -5). Then $E[N] = \beta = 4$.

Since $E[S] = E[N] \times E[X]$ it follows from the parameters in the old model that $E[X] = 4$.

Since $328 = Var[S] = E[N] \times Var[X] + Var[N] \times (E[X])^2 = 4 \times Var[X] + 20 \times 4^2$,
it follows that $Var[S] = 2$. Then, $E[X^2] = Var[X] + (E[X])^2 = 18$, and
$Var[S] = \lambda \times E[X^2] = 4 \times 18 = 72.$

Answer D

49. The amount Bob keeps is the stop loss amount $E\left[(S-3)_+\right]$, where S is the compound random variable representing the total amount thrown into Bob's hat in one hour
$E\left[(S-3)_+\right] = E[S] - E[S \wedge 3]$. We see that the average amount thrown into Bob's hat by any one individual is $1 \times \frac{1}{2} + 2 \times \frac{1}{3} + 5 \times \frac{1}{6} = 2$, so the average total thrown into Bob's hat in one hour is $E[S] = E[N] \times E[X] = 4 \times 2 = 8$ (N is the number of people who contribute to Bob's hat in an hour and X is the amount of an individual contribution).

$$S \wedge 3 = \begin{cases} 0 & P[S=0] \\ 1 & P[S=1] \\ 2 & P[S=2] \\ 3 & 1 - P[S=0{,}1{,}2] \end{cases}$$

where $P[S=0] = P[N=0] = e^{-4}$,

$$P[S=1] = P[N=1] \times P[X=1] = 4e^{-4} \times \frac{1}{2} = 2e^{-4},$$

$$P[S=2] = P[N=1] \times P[X=2] + P[N=2] \times P[X=1]^2$$
$$= 4e^{-4} \times \tfrac{1}{3} + \tfrac{4^2e^{-4}}{2} \times \left(\tfrac{1}{2}\right)^2 = \tfrac{10e^{-4}}{3}.$$

Then $E[S \wedge 3] = 1 \times 2e^{-4} + 2 \times \frac{10e^{-4}}{3} + 3 \times \left(1 - \frac{19e^{-4}}{3}\right) = 2.81$,
and $E\left[(S-3)_+\right] = 8 - 2.81 = 5.19.$

Answer D

Practice Exam 1

1. A portfolio of risks models the annual loss of an individual risk as having an exponential distribution with a mean of Λ. For a randomly selected risk from the portfolio, the value of Λ has an inverse gamma distribution with a mean of 40 and a standard deviation of 20.

   (a) (4 Points) Determine the values of the parameters α and θ for the distribution of Λ.

   (b) (4 Points) Determine the probability $P[X > 20]$.

   (c) (4 Points) Show that the unconditional distribution of X is Pareto with the same values of α and θ as the distribution of Λ.

2. [Workbook question] S has a compound distribution.

   The severity random variable is X with the following distribution:

   $$P(X = 1) = 0.4, \quad P(X = 2) = 0.3, \quad P(X = 3) = 0.2, \quad P(X = 4) = 0.1$$

   (a) (4 Points) Suppose that the frequency random variable N has a Poisson distribution with a mean of 0.5. Create an Excel workbook that applies the recursive method for $(a,b,0)$ frequency distributions to calculate $P(S = x)$ for successive integer values of x starting at $x = 0$.

   Calculate the cumulative distribution function $F_S(x)$ for successive values of x until $|F_S(x) - 1| < 10^{-8}$. Find the approximate value of $E[S]$ by taking the sum of $x \times P(S = x)$ up to the last value of x for which $F_S(x)$ was just calculated. Verify that this value is consistent with the theoretical value of $E[S]$.

   (b) (4 Points) Repeat part (a) for a geometric distribution with a mean of 0.5.

3. You are given the following:

   - Losses follow a distribution (prior to the application of any deductible) with mean 2000.
   - The loss elimination ratio (LER) at a deductible of 1000 is 0.3.
   - 60 percent of the losses (in number) are less than the deductible of 1000.

   (a) (2 Points) Find $E[X \wedge 1000]$.

   (b) (6 Points) Find $E[X|X < 1000]$.

4. The following table was obtained by fitting both a Poisson distribution and a binomial distribution to a data set of 100,000 integer-valued observations.

		$\widehat{\lambda} = .11527$		$\widehat{m} = 4, \widehat{q} = .0288175$	
		Fitted Poisson		Fitted Binomial	
k	n_k	expected	χ^2	expected	χ^2
0	89,000	89,112.6	.1422	88,961.8	.0164
1	10,487	10,272.0	4.500	10,558.9	.4897
2	500	592.0	14.3050	470.0	1.9195
≥ 3	13	23.4	4.622	9.3	1.3787
Totals	100,000	100,000	23.57	100,000	3.80
Degrees of freedom		$4 - 1 - 1 = 2$			$4 - 2 - 1 = 1$
p -value		$p < .001$			$.05 < p < .1$

You are also given that the negative loglikelihood of the fitted Binomial model is 36,787.

(a) (3 Points) Calculate the negative loglikelihood of the fitted Poisson model.

(b) (3 Points) Apply the Chi-square Goodness-of-Fit test to both fitted models and determine the level of signficance for each model for which the null hypothesis is rejected: H_0 - the data is consistent with the estimated model.

(c) (3 Points) A likelihood ratio test is performed to determine whether or not the fitted Binomial Model is preferable to the fitted Poisson model, with the null hypothesis being that the Binomial is not preferable to the Poisson model. Calculate the value of the test statistic and determine the level of significance (based on the χ^2 table) at which the fitted Binomial model is preferable to the fitted Poisson model.

(d) (3 Points) Apply the Schwarz Bayesian Criterion to choose between the Poisson and Binomial models.

Copyright © 2024 ACTEX Learning

5. A risk class is made up of three equally sized groups of individuals. Groups are classified as Type A, Type B and Type C. Any individual of any type has probability of .5 of having no claim in the coming year and has a probability of .5 of having exactly 1 claim in the coming year. Each claim is for amount 1 or 2 when a claim occurs. Suppose that the claim distributions given that a claim occurs, for the three types of individuals are

$$P(\text{claim of amount } x|\text{Type A and a claim occurs}) = \begin{cases} 2/3 & x=1 \\ 1/3 & x=2 \end{cases},$$

$$P(\text{claim of amount } x|\text{Type B and a claim occurs}) = \begin{cases} 1/2 & x=1 \\ 1/2 & x=2 \end{cases},$$

$$P(\text{claim of amount } x|\text{Type C and a claim occurs}) = \begin{cases} 5/6 & x=1 \\ 1/6 & x=2 \end{cases}.$$

(a) (5 Points) Find the hypothetical mean and find the expected hypothetical mean and the variance of the hypothetical mean.

(b) (3 Points) Find the process variance and find the expected process variance.

(c) (2 Points) If an individual is chosen at random from the risk population and the individual is observed to have claim amount X in the coming year, find the Buhlmann credibility premium for the following year for this individual.

6. Suppose that $X_1, X_2, \ldots, X_n, \ldots$ are independent and identically distributed random variables, which follow the exponential distribution with mean 5. Let $M_n = \max(X_1, X_2, \ldots, X_n)$.

(a) (3 Points) Show that the p-level VaR of X_1 is given by $VaR_p(X) = -5\ln(1-p)$ for any $p \in (0,1)$, and that the p-level expected shortfall $ES_p(X_1) = VaR_p(X_1) + 5$.

(b) (3 Points) Use the Fisher-Tippett-Gnedenko theorem to find the normalizing constants a_n and b_n such that the distribution of $\frac{M_n - a_n}{b_n}$ converges to the Gumbel distribution.

(c) (2 Points) Evaluate the probability $P(M_{10} \le VaR_{0.8}(X_1))$ using the approximated Gumbel distribution, and calculate the difference between it and the precise probability (to the nearest 0.0001).

(d) (2 Points) Evaluate the probability $P(M_{10} \le ES_{0.8}(X_1))$ using the approximated Gumbel distribution, and calculate the difference between it and the precise probability (to the nearest 0.0001).

Practice Exam 1 Solutions

1. (a) We are given that the mean and standard deviation of Λ are 40 and 20. Therefore, the variance of Λ is 20^2 and the 2nd moment of Λ is

$$E\left[\Lambda^2\right] = Var[\Lambda] + (E[\Lambda])^2 = 20^2 + 40^2 = 2000.$$

The mean of an inverse gamma is $\frac{\theta}{\alpha-1}$. and the 2nd moment is $\frac{\theta^2}{(\alpha-2)(\alpha-1)}$.
From the two equations $\frac{\theta}{\alpha-1} = 40$ and $\frac{\theta^2}{(\alpha-2)(\alpha-1)} = 2000$,
we get $\frac{2000}{40^2} = \frac{\theta^2}{(\alpha-2)(\alpha-1)} \Big/ \left(\frac{\theta}{\alpha-1}\right)^2 = \frac{\alpha-1}{\alpha-2}$. Then solving for α results in $\alpha = 6$.
Substituting back into $\frac{\theta}{\alpha-1} = 40$, we get that $\theta = 200$.

(b) Given $\Lambda = \lambda$, the annual loss X has an exponential distribution with mean λ and Λ has an inverse gamma distribution. X is a continuous mixture distribution of an "exponential over an inverse gamma". Suppose that the inverse gamma distribution of Λ has parameters α and θ. When we have a continuous mixture distribution for X over Λ, the pdf, expected values and probabilities for the marginal distribution of X can be found by conditioning over Λ.

The conditional pdf of X given $\Lambda = \lambda$ is $f(x|\Lambda = \lambda) = \frac{1}{\lambda}e^{-x/\lambda}$
and the pdf of the inverse gamma distribution of Λ is $f_\Lambda(\lambda) = \frac{\theta^\alpha e^{-\theta/\lambda}}{\lambda^{\alpha+1}\cdot\Gamma(\alpha)}$.
From the calculated parameter values in part (a), we have $f_\Lambda(\lambda) = \frac{200^6 e^{-200/\lambda}}{\lambda^7\cdot\Gamma(6)}$.
We can find the probability $P(X > 20)$ by conditioning over λ:

$$P(X > 20) = \int_0^\infty P(X > 20|\lambda)\cdot f_\Lambda(\lambda)d\lambda = \int_0^\infty e^{-20/\lambda}\cdot\frac{200^6 e^{-200/\lambda}}{\lambda^7\cdot\Gamma(6)}d\lambda$$
$$= \frac{200^6}{\Gamma(6)}\cdot\int_0^\infty \frac{e^{-220/\lambda}}{\lambda^7}d\lambda$$

We use the identity $\int_0^\infty \frac{e^{-c/\lambda}}{\lambda^k}d\lambda = \frac{\Gamma(k-1)}{c^{k-1}}$. It follows that $\int_0^\infty \frac{e^{-\theta/\lambda}}{\lambda^{\alpha+1}}d\lambda = \frac{\Gamma(\alpha)}{\theta^\alpha}$.

Therefore, $\int_0^\infty \frac{e^{-220/\lambda}}{\lambda^7}d\lambda = \int_0^\infty \frac{e^{-220/\lambda}}{\lambda^{6+1}}d\lambda = \frac{\Gamma(6)}{220^6}$, and $P(X > 20) = \frac{200^6}{\Gamma(6)}\cdot\frac{\Gamma(6)}{220^6} = .564$.

(c) In general suppose that X has an exponential distribution with mean Λ and that Λ has an inverse gamma distribution with parameters α and θ. The unconditional pdf of X can be found by conditioning in a way similar to method of part (b).

$$f_X(x) = \int_0^\infty f_X(x|\Lambda = \lambda)\times f_\Lambda(\lambda)\ d\lambda = \int_0^\infty \frac{1}{\lambda}e^{-x/\lambda}\times\frac{\theta^\alpha e^{-\theta/\lambda}}{\lambda^{\alpha+1}\cdot\Gamma(\alpha)}\ d\lambda = \int_0^\infty \frac{\theta^\alpha e^{-(\theta+x)/\lambda}}{\lambda^{\alpha+2}\cdot\Gamma(\alpha)}\ d\lambda$$

This becomes $\frac{\theta^\alpha}{\Gamma(\alpha)}\times\int_0^\infty \frac{e^{-(\theta+x)/\lambda}}{\lambda^{\alpha+2}}\ d\lambda$.
Using the integral identity from part (b), we have $\int_0^\infty \frac{e^{-(\theta+x)/\lambda}}{\lambda^{\alpha+2}}\ d\lambda = \frac{\Gamma(\alpha+1)}{(\theta+x)^{\alpha+1}}$.

The unconditional pdf of X becomes $f_X(x) = \frac{\theta^\alpha}{\Gamma(\alpha)}\times\frac{\Gamma(\alpha+1)}{(\theta+x)^{\alpha+1}}$.

Using the relationship $\Gamma(\alpha+1) = \alpha\times\Gamma(\alpha)$, the pdf of X becomes $f_X(x) = \frac{\alpha\theta^\alpha}{(\theta+x)^{\alpha+1}}$.

This is the pdf of a Pareto random variable with parameters α and θ.

Note that we could also have found the unconditional cdf of X by conditioning in a similar way.