An Integrated Component of Your Study Manual Program

Actuarial University
Your Path to Success

GOAL
Practice. Quiz. Test. PASS

Study Manual

Virtual eFlashcards

Videos & More!

a/s/m

Exam MAS-I Study Manual

5th Edition

Abraham Weishaus, Ph.D., FSA, CFA, MAAA

a/s/m
Actuarial Study Materials
Learning Made Easier

a/s/m

Exam MAS-I Study Manual

5th Edition

Abraham Weishaus, Ph.D., FSA, CFA, MAAA

TO OUR READERS:

Please check A.S.M.'s web site at **www.studymanuals.com** for errata and updates. If you have any comments or reports of errata, please e-mail us at **mail@studymanuals.com**.

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.

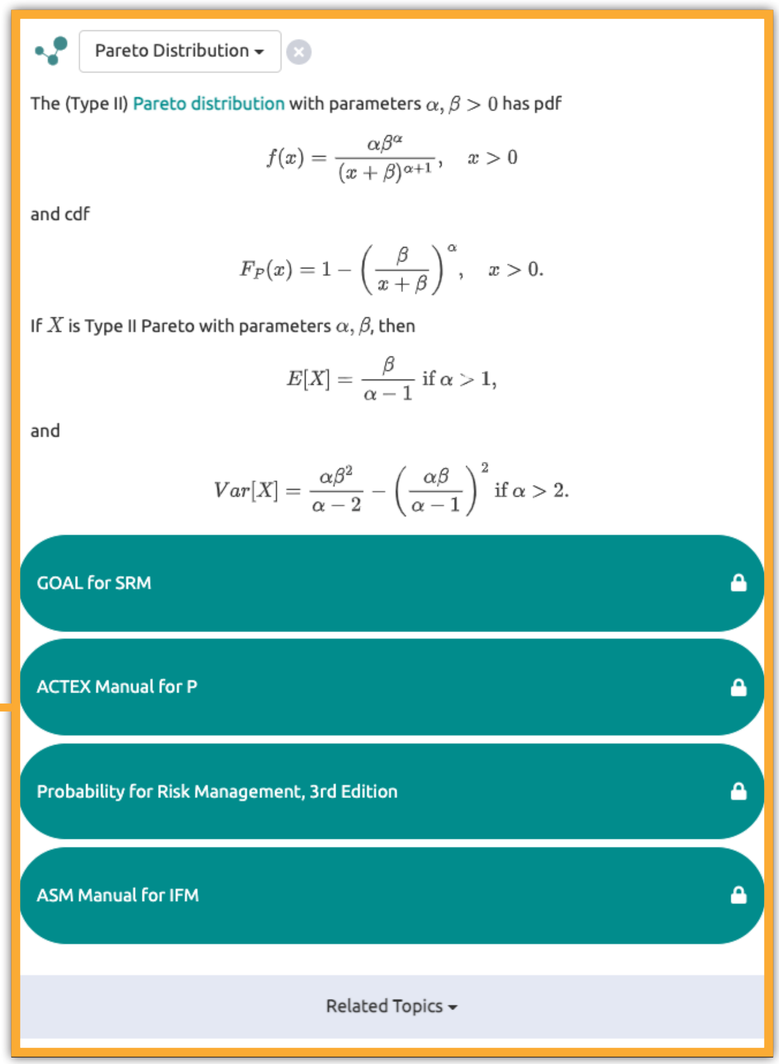
When this icon appears, it will be next to an important topic in the manual! **Click the link** in your digital manual, or **search the underlined topic** in your print manual.

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.

GOAL 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:

GOAL Practice Tool	**Instructional Videos***
Virtual Flashcards	**Formula & Review Sheet**

GOAL: Guided Online Actuarial Learning

Your Adaptable, Customizable, Online Actuarial Exam Prep Tool.

GOAL is an **eLearning test preparation tool** for students to practice skills learned in class or from independent study. The online platform offers a massive database of **SOA & CAS exam-style problems with detailed solutions**.

GOAL offers **instructor support** as well as **performance reporting and tracking** to monitor your progress. Don't forget to check out **GOAL Score** to see how ready you are to sit for your exam!

Practice. Quiz. Test. Pass!

- 10,000+ Exam-style problems with detailed solutions!
- Adaptive Quizzes
- 3 Learning Modes
- 3 Difficulty Modes

How GOAL Works

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

Practice. Predict. Pass.

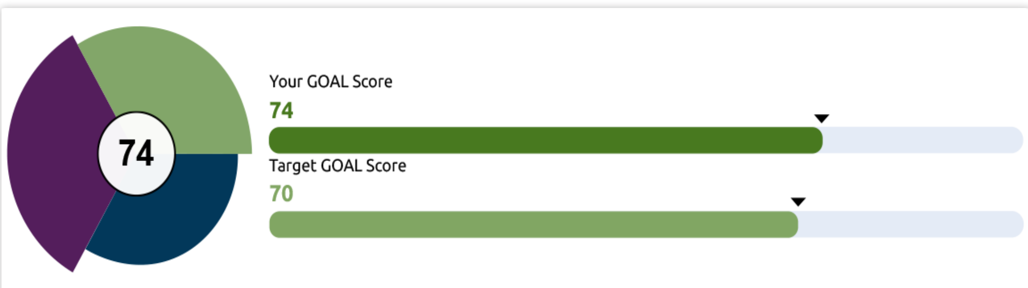
GOAL

Track your exam readiness with GOAL Score!

Available now for P, FM, FAM, ALTAM, ASTAM, SRM, MAS-I, MAS-II, CAS 6C/U

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.



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

62	74	78
General Probability	Random Variables with Univariate Probability Distributions	Random Variables with Univariate Probability Distributions
Weighting: 13.4%	Weighting: 43.3%	Weighting: 43.3%

See key areas where you can improve.

General Probability — Set functions including set notation and basic elements of probability

Difficulty	Mean	Proportion
Core	58	92 / 304
Advanced	60	169 / 304
Mastery	78	43 / 304

Detailed performance tracking.

GOAL Session Activity Summary — Filter: All Types, All Statuses

Created	Last Accessed	Completed	Mode	Categories	Questions	Status	Status
05/24/2022 21:57:43	05/24/2022 21:57:43		Quiz	Continuous P...	25	New	Resume
05/24/2022 10:34:05	05/24/2022 14:57:49	05/24/2022 14:57:49	Practice Session	Addition and ...	80	Complete	Review
05/21/2022 19:32:50	05/23/2022 20:02:00		Simulated Exam	Exam 2	30	Reviewing	Complete
05/17/2022 15:19:19	05/17/2022 15:46:03	05/23/2022 14:15:29	Simulated Exam	Exam 6	30	Complete	Review
05/14/2022 11:26:59	05/14/2020 12:02:36	05/23/2022 11:57:47	Quiz	Conditional D...	20	Complete	Review

Quickly return to previous sessions.

Contents

Introduction

Welcome to Exam MAS-I!

Syllabus

Exam MAS-I is a 4 hour exam. Every released Exam MAS-I had 45 questions, so I expect it to continue to have 45 questions.

At this writing (June 2023), the syllabus for Fall 2023, which is now called the content outline, is posted at https://www.casact.org/sites/default/files/2022-12/MASI_Content_Outline_Fall-2023.pdf. The exam has the following major topics: (Content outline weights for each topic are in parentheses.)

- Probability models: Poisson processes, Markov chains, reliability theory, life contingencies, simulation (20–30%)
- Statistics (20–30%)
- Extended linear models (45–55%)

The following table gives the number of questions on each topic on the four released exams.

	Spring 2018	Fall 2018	Spring 2019	Fall 2019
Probability	0	1	2	1
Poisson processes	6	4	5	4
Reliability theory	4	3	1	3
Markov chains	2	3	3	2
Life contingencies	2	3	2	2
Simulation	1	1	2	1
Estimator quality, kernel smoothing	1	3	3	3
Parameter estimation	2	3	3	2
Hypothesis testing	5	4	5	5
Generalized linear model	11	8	8	7
Statistical learning	6	7	6	9
Time series	5	5	5	5
Total	45	45	45	45

Time series is no longer on the syllabus.

Tables

Download the tables you will get at the exam from the CAS website. You will need them to work out exercises in this manual. These tables list characteristics such as density functions and moments for many common distributions. The tables also include the CDF of the standard normal distribution, critical values for the t, F, and chi-square distributions.

The tables for the Fall 2023 exam have not been posted at this writing. The tables used for previous sittings are found at

https://www.casact.org/sites/default/files/2021-03/masi_tables.pdf.

On the released exams, the following instructions were given regarding using the normal distribution table:

When using the normal distribution, unless told otherwise, do not interpolate. If you want $\Phi(x)$, round x to two places and use the value in the table. For example, $\Phi(1.1342)$ would be evaluated as 0.8708. If you want $\Phi^{-1}(z)$, find the 4-digit value in the table that is closest to z and use its inverse. For example, $\Phi^{-1}(0.92) = 1.41$. The exception to this rule is that the 3-digit values for special values of z shown at the bottom of the table should be used when appropriate. For example, for the 90$^{\text{th}}$ percentile of a standard normal distribution use 1.282, not 1.28.

Currently, you get a spreadsheet with Excel functions at the exam. Using `NORMSDIST` and `NORMSINV` you can obtain normal values to high degrees of precision. So I am not sure the previous paragraph is still applicable. In this manual, the method of the previous paragraph will be used. However, on the exam, read the exam instructions carefully and follow them.

Characteristics of CAS exams

This may be the first CAS exam you are taking. CAS exams have a somewhat different style from SOA exams.

Starting with Fall 2023, there will be four types of questions:

Multiple choice Traditional 5-choice questions with only one correct choice. Previously there was no guessing penalty, and I suspect that policy will continue.

Multiple selection Multiple answer choices with more than one correct answer. These questions will probably be a list of true/false statements rather than numeric.

Point and click An image is presented after a problem. You must click the correct location on the image.

Fill in the blank You must fill in the correct answer.

CAS exam multiple choice questions with numeric answers in the past usually provided ranges rather then specific answer choices. The ranges are usually equal in size, and your answer should usually not be more than the size of a range lower than the first choice or higher than the last choice. For example, if the choices offered are

A. Less than 5
B. At least 5, but less than 7
C. At least 7, but less than 9
D. At least 9, but less than 11
E. At least 11

then your answer should not be less than 3 or more than 13. However, this rule is not hard-and-fast, so if you get an answer far out of range, it is suspect but not necessarily wrong.

Before 2020, every CAS exam was released, and answer choices were provided as well, but not worked out solutions. Exams are now administered with computer-based testing, so starting in 2020, exams are not be released.

Released CAS exams frequently had defective questions. If you look at the old Exam 3, Exam 3L, Exam ST, Exam LC, and Exam S answer lists, you will frequently find that more than one answer was accepted. There was an average of one defective question per exam. These defects were of the following types:

1. Questions with typos. These questions are usually not considered defective; the typo is annoying, but you still have to work out the question.

2. Questions with poor or incorrect wording. In these questions, you have to figure out what they meant and answer the question accordingly, rather than interpret the question literally.

3. Questions with ambiguous wording allowing more than one interpretation. These are the questions where often more than one answer choice is accepted.

4. Questions with specific answer choices (rare for the CAS) in which none of the answer choices is correct. Or more commonly, answers with range answer choices in which the answer is far out of the range. For example, a question with ranges like < 10, 10–30, 30–50, 50–70, and > 70 where the answer is 822.

5. Questions that cannot be answered with the information and tables that you are given.

With the move to computer-based testing and reused questions, perhaps questions will no longer be defective. But if you have great difficulty solving a question despite knowing the underlying material, it may be best to move on to the next question.

The exam spreadsheet

At the exam you will be given a spreadsheet. For a sample, go to

https://home.pearsonvue.com/cas

and at the bottom of the webpage, click "SAMPLE SPREADSHEET". This spreadsheet includes a TI-30XS calculator.

The CAS has stated that the computational nature of exam questions will not change for the meantime. Questions will not require use of spreadsheet functions to solve.

Features of this manual

This manual has about 1300 exercises; about 600 of them are original. The exercises taken from old exams are indicated by xxx–yy:zz where xxx is the exam number or name, yy is the date of the exam (e.g. S14=spring 2014, F18=fall 2018), and zz is the question number. Almost no exercises were taken from recent exams, but a list of relevant questions from recent exams is given at the end of each lesson, and solutions to these questions can be found in Appendix B.

In lessons that deal with topics that have been on the syllabus for a long time, there are lots of exercises, mostly old exam questions. In lessons dealing with topics new to the syllabus, or that were on the syllabus in the past only briefly, most or all exercises are original and there are only a few of them. Moreover, there can be no guarantee that exam questions will emphasize the same topics as the exercises. Whenever a lesson has only a few exercises, do all of them.

Solutions to all relevant questions from old Exams 3L, LC, ST, S, and MAS-I are in Appendix B.

This manual has an index.

This manual, whose primary purpose is exam preparation, does not cover `R`. Exams do not expect you to program in `R`; they merely expect you to understand outputs from programs. I believe the material in this manual will be sufficient for you to work out any exam question. On the other hand, applying the methods discussed in this course to real life requires running computer programs. If you are using this manual as a textbook replacement, I encourage you to go through the labs in *An Introduction to Statistical Learning* in order to learn how to use `R` to carry out the methods discussed in this course.

New to this edition

Time series was removed, since it is no longer on the Exam MAS-I syllabus.

Acknowledgements

I would like to thank the SOA and CAS for allowing me to use their old exam questions.

The creators of TeX, LaTeX, and its multitude of packages all deserve thanks for making possible the professional typesetting of this mathematical material.

I thank Geoff Tims for proofreading much of the manual. Not only did he point out many errors, he also made many helpful suggestions for improving the style.

I'd like to thank the following readers who submitted errata for this manual: Adam Karnik, Megan Benoit, Jonathan Brand, Joseph Breslin, Wenqi Chen, Deb Clough, Jorge Miguel Conceicao, Zack de la Pena, Reginald Dorsey, Janeth Fernández, Christopher Filips, Nehama Florans, Weston Hogan, Aaron Johnson, Adam Karnik, Tzong Her Lee, Matthew Lenko, Dianting Liu, Lawrence Miley, Jaime Mullen, Li Kee Ong, Dauna Papakirykos, Seth Reuter, Alyssa Rineha4rt, Juan Sancen-Bravo, Adam Tolnay, Aytan Wachspress, Woosuk Yoo.

Errata

Please report all errors you find in this manual to the author. You may send them to the publisher at mail@studymanuals.com.This is the fifth edition of the Exam MAS-I manual..

An errata list will be posted at http://errata.aceyourexams.net. Check this errata list frequently.

Part I

Probability

Lesson 2

Parametric Distributions

Reading: Tse 2.2, Hogg, McKean, Craig 2.2,2.7

A **parametric distribution** is one that is defined by a fixed number of parameters. Examples of parametric distributions are the exponential distribution (parameter θ) and the Pareto distribution (parameters α, θ). Any distribution listed in the tables you receive at the exam is parametric.

It is traditional to use parametric distributions for claim counts (frequency) and loss size (severity). Parametric distributions have many advantages. One of the advantages of parametric distributions which makes them so useful for severity is that they handle inflation easily.

2.1 Transformations

You learn how to transform a distribution when you study probability. We will discuss the easiest case here.

If $Y = g(X)$, with $g(x)$ a **one-to-one monotonically increasing** function, then

$$F_Y(y) = \Pr(Y \le y) = \Pr\big(X \le g^{-1}(y)\big) = F_X\big(g^{-1}(y)\big) \tag{2.1}$$

and differentiating,

$$f_Y(y) = f_X\big(g^{-1}(y)\big)\frac{\mathrm{d}g^{-1}(y)}{\mathrm{d}y}$$

If $g(x)$ is one-to-one monotonically decreasing, then

$$F_Y(y) = \Pr(Y \le y) = \Pr\big(X \ge g^{-1}(y)\big) = S_X\big(g^{-1}(y)\big) \tag{2.2}$$

and differentiating,

$$f_Y(y) = -f_X\big(g^{-1}(y)\big)\frac{\mathrm{d}g^{-1}(y)}{\mathrm{d}y}$$

Putting both cases (monotonically increasing and monotonically decreasing) together:

$$f_Y(y) = f_X\big(g^{-1}(y)\big)\left|\frac{\mathrm{d}g^{-1}(y)}{\mathrm{d}y}\right| \tag{2.3}$$

Example 2A X follows a two-parameter Pareto distribution with parameters α and θ. You are given

$$Y = \ln\left(\frac{X}{\theta} + 1\right)$$

Determine the distribution of Y. ■

Solution:

$$\begin{aligned}
y &= \ln\left(\frac{x}{\theta}+1\right)\\
e^y - 1 &= \frac{x}{\theta}\\
x &= \theta(e^y-1)\\
F_Y(y) &= F_X\bigl(\theta(e^y-1)\bigr)\\
&= 1-\left(\frac{\theta}{\theta+\theta(e^y-1)}\right)^\alpha\\
&= 1-\left(\frac{\theta}{\theta e^y}\right)^\alpha\\
&= 1-e^{-\alpha y}
\end{aligned}$$

So Y's distribution is exponential with parameter $\theta = 1/\alpha$. □

We see in this example that an exponential can be obtained by transforming a Pareto.

2.2 Common parametric distributions

The tables provide a lot of information about the distributions, but if you don't recognize the distribution, you won't know to use the table. Therefore, it is a good idea to be familiar with the common distributions.

You should familiarize yourself with the *form* of each distribution, but not necessarily the constants. The constant is forced so that the density function will integrate to 1. If you know which distribution you are dealing with, you can figure out the constant. To emphasize this point, in the following discussion, we will use the letter c for constants rather than spelling out what the constants are. You are not trying to recognize the constant; you are trying to recognize the form.

We will mention the means and variances or second moments of the distributions. You need not memorize any of these. The tables give you the raw moments. You can calculate the variance as $\mathbf{E}[X^2]-\mathbf{E}[X]^2$. However, for frequently used distributions, you may want to memorize the mean and variance to save yourself some time when working out questions.

We will graph the distributions. You are not responsible for graphs, but they may help you understand the distributions.

The tables occasionally use the **gamma function** $\Gamma(x)$ in the formulas for the moments. You should have a basic knowledge of the gamma function; if you are not familiar with this function, see the sidebar. The tables also use the **incomplete gamma** and **beta functions**, and define them, but you can get by without knowing them.

2.2.1 Uniform

A **uniform distribution** has a constant density on $[d,u]$:

$$f(x;d,u) = \frac{1}{u-d} \qquad d\le x\le u$$

$$F(x;d,u) = \begin{cases} 0 & x\le d\\ \dfrac{x-d}{u-d} & d\le x\le u\\ 1 & x\ge u\end{cases}$$

You recognize a uniform distribution both by its finite **support**[1] and by the lack of an x in the density function.

[1]"Support" is the range of values for which the probability density function is nonzero.

The gamma function

The gamma function $\Gamma(x)$ is a generalization to real numbers of the factorial function, defined by

$$\Gamma(x) = \int_0^\infty u^{x-1} e^{-u} \mathrm{d}u$$

For positive integers n,

$$\Gamma(n) = (n-1)!$$

The most important relationship for $\Gamma(x)$ that you should know is

$$\Gamma(x+1) = x\Gamma(x)$$

for any real number x.

Example 2B Evaluate $\dfrac{\Gamma(8.5)}{\Gamma(6.5)}$.

Solution:

$$\frac{\Gamma(8.5)}{\Gamma(6.5)} = \left(\frac{\Gamma(8.5)}{\Gamma(7.5)}\right)\left(\frac{\Gamma(7.5)}{\Gamma(6.5)}\right) = (7.5)(6.5) = \boxed{\mathbf{48.75}}$$

Its moments are

$$\mathbf{E}[X] = \frac{d+u}{2}$$

$$\mathrm{Var}(X) = \frac{(u-d)^2}{12}$$

Its mean, median, and midrange are equal. The best way to calculate the second moment is to add up the variance and the square of the mean. However, some students prefer to use the following easy-to-derive formula:

$$\mathbf{E}[X^2] = \frac{1}{u-d}\int_d^u x^2 \,\mathrm{d}x = \frac{u^3 - d^3}{3(u-d)} = \frac{u^2 + ud + d^2}{3} \tag{2.4}$$

If $d = 0$, then the formula reduces to $u^2/3$.

The uniform distribution is not directly in the tables, so I recommend you memorize the formulas for mean and variance. However, if $d = 0$, then the uniform distribution is a special case of a **beta distribution** with $\theta = u$, $a = 1$, $b = 1$.

2.2.2 Beta

The probability density function of a **beta distribution** with $\theta = 1$ has the form

$$f(x; a, b) = cx^{a-1}(1-x)^{b-1} \qquad 0 \le x \le 1$$

The parameters a and b must be positive. They may equal 1, in which case the corresponding factor is missing from the density function. Thus if $a = b = 1$, the beta distribution is a uniform distribution.

You recognize a beta distribution both by its finite support—it's the only common distribution for which the density is nonzero only on a finite range of values—and by factors with x and $1-x$ raised to powers and no other use of x in the density function.

If θ is arbitrary, then the form of the probability density function is

$$f(x; a, b, \theta) = cx^{a-1}(\theta - x)^{b-1} \qquad 0 \le x \le \theta$$

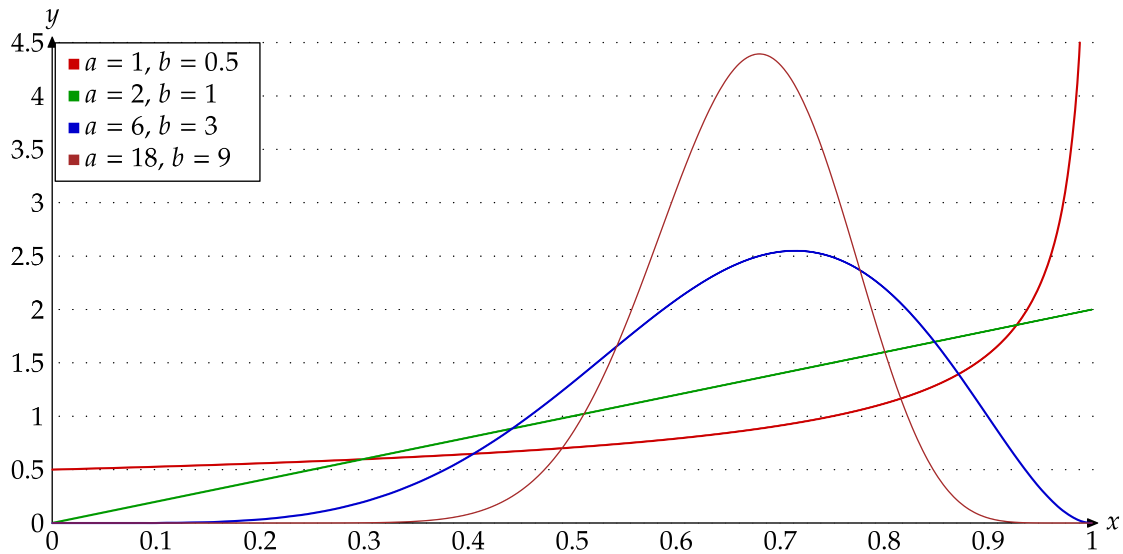


Figure 2.1: Probability density function of four beta distributions with $\theta = 1$ and mean 2/3

The distribution function can be evaluated if a or b is an integer. The moments are

$$\mathbf{E}[X] = \frac{\theta a}{a+b}$$
$$\operatorname{Var}(X) = \frac{\theta^2 ab}{(a+b)^2(a+b+1)}$$

The mode is $\theta(a-1)/(a+b-2)$ when a and b are both greater than 1, but you are not responsible for this fact.

Figure 2.1 graphs four beta distributions with $\theta = 1$ all having mean $2/3$. You can see how the distribution becomes more peaked and normal looking as a and b increase.

2.2.3 Exponential

The probability density function of an **exponential distribution** has the form

$$f(x;\theta) = ce^{-x/\theta} \qquad x \geq 0$$

θ must be positive.

You recognize an exponential distribution when the density function has e raised to a multiple of x, and no other use of x.

The distribution function is easily evaluated. The moments are:

$$\mathbf{E}[X] = \theta$$
$$\operatorname{Var}(X) = \theta^2$$

Figure 2.2 graphs three exponential distributions. The higher the parameter, the more weight placed on higher numbers.

The amount of time between events in a Poisson process is exponentially distributed. We'll discuss Poisson processes later in the course.

2.2.4 Weibull

A **Weibull distribution** is a **transformed** exponential distribution. If Y is exponential with mean μ, then $X = Y^{1/\tau}$

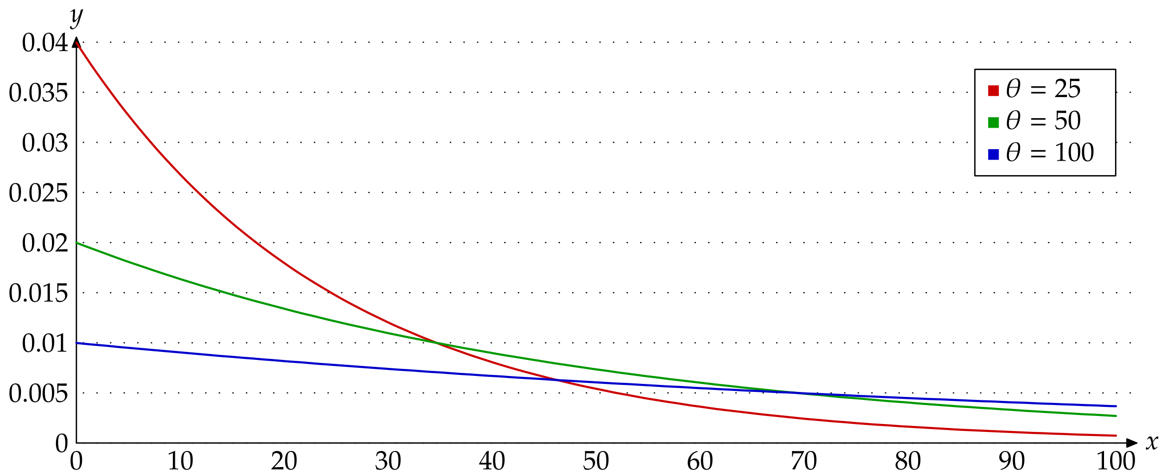


Figure 2.2: Probability density function of three exponential distributions

is Weibull with parameters $\theta = \mu^{1/\tau}$ and τ. An exponential is a special case of a Weibull with $\tau = 1$.

The form of the density function is

$$f(x;\tau,\theta) = cx^{\tau-1}e^{-(x/\theta)^{\tau}} \qquad x \ge 0$$

Both parameters must be positive. The shape parameter is τ and the scale parameter is θ.

You recognize a Weibull distribution when the density function has e raised to a multiple of a power of x, and in addition has a corresponding power of x, one lower than the power in the exponential, as a factor.

When $\theta = 1$, we say that the distribution is a *standard* Weibull.

The distribution function is easily evaluated, but the moments require evaluating the gamma function, which usually requires numerical techniques. The moments are

$$\mathbf{E}[X] = \theta\Gamma(1 + 1/\tau)$$
$$\mathbf{E}[X^2] = \theta^2\Gamma(1 + 2/\tau)$$

Figure 2.3 graphs three Weibull distributions with mean 50. The distribution has a non-zero mode when $\tau > 1$. Notice that the distribution with $\tau = 0.5$ puts a lot of weight on small numbers. To make up for this, it will also have to put higher weight than the other two distributions on very large numbers, so although it's not shown, its graph will cross the other two graphs for high x

2.2.5 Gamma

The form of the density function of a **gamma distribution** is

$$f(x;\alpha,\theta) = cx^{\alpha-1}e^{-x/\theta} \qquad x \ge 0$$

Both parameters must be positive. The shape parameter is α and the scale parameter is θ.

When α is an integer, a gamma random variable with parameters α and θ is the sum of α independent exponential random variables with parameter θ. In particular, when $\alpha = 1$, the gamma random variable is exponential. The gamma distribution is called an **Erlang distribution** when α is an integer.

You recognize a gamma distribution when the density function has e raised to a multiple of x, and in addition has x raised to a power. Contrast this with a Weibull, where e is raised to a multiple of a *power* of x.

The distribution function may be evaluated if α is an integer; otherwise numerical techniques are needed. However, the moments are easily evaluated:

$$\mathbf{E}[X] = \alpha\theta$$
$$\mathrm{Var}(X) = \alpha\theta^2$$

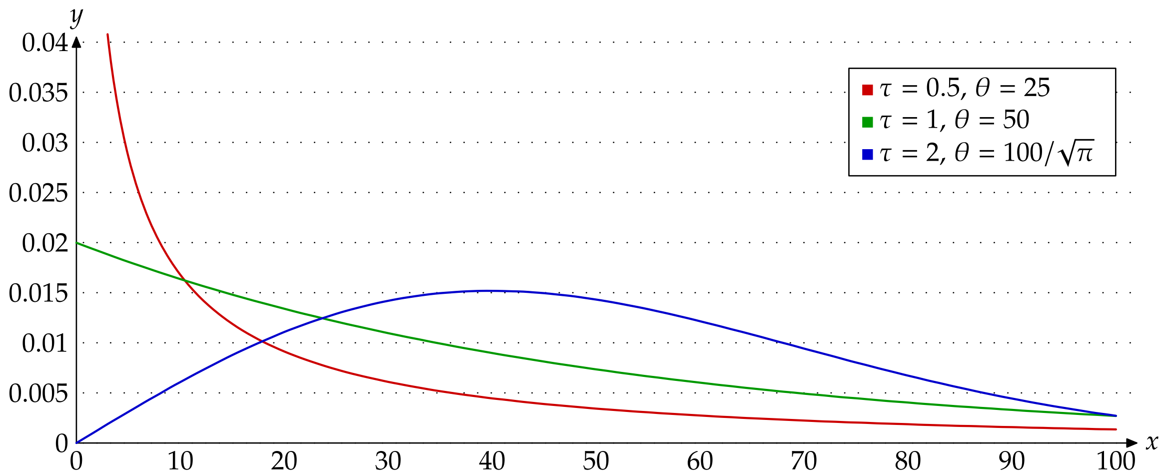


Figure 2.3: Probability density function of three Weibull distributions with mean 50

Figure 2.4 graphs three gamma distributions with mean 50. As α goes to infinity, the graph's peak narrows and the distribution converges to a normal distribution.

The gamma distribution is one of the few for which the **moment generating function** has a closed form. In particular, the moment generating function of an exponential has a closed form. The only other distributions in the tables with closed form moment generating functions are the **normal distribution** (not actually in the tables, but the formula for the lognormal moments is the MGF of a normal) and the **inverse Gaussian**.

2.2.6 Pareto

When we say "Pareto", we mean a *two-parameter* Pareto. On recent exams, they write out "two-parameter" to make it clear, but on older exams, you will often find the word "Pareto" with no qualifier. It always refers to a two-parameter Pareto, not a **single-parameter Pareto**.

The form of the density function of a two-parameter Pareto is

$$f(x) = \frac{c}{(\theta + x)^{\alpha+1}} \qquad x \geq 0$$

Both parameters must be positive.

You recognize a Pareto when the density function has a denominator with x plus a constant raised to a power.

The distribution function is easily evaluated. The moments are

$$\mathbf{E}[X] = \frac{\theta}{\alpha - 1} \qquad \alpha > 1$$
$$\mathbf{E}[X^2] = \frac{2\theta^2}{(\alpha - 1)(\alpha - 2)} \qquad \alpha > 2$$

When α does not satisfy these conditions, the corresponding moments don't exist.

A shortcut formula for the variance of a Pareto is

$$\mathrm{Var}(X) = \mathbf{E}[X]^2\left(\frac{\alpha}{\alpha - 2}\right)$$

Figure 2.5 graphs three Pareto distributions, one with $\alpha < 1$ and the other two with mean 50. Although the one with $\alpha = 0.5$ puts higher weight on small numbers than the other two, its mean is infinite; it puts higher weight on large numbers than the other two, and its graph eventually crosses the other two as $x \to \infty$.

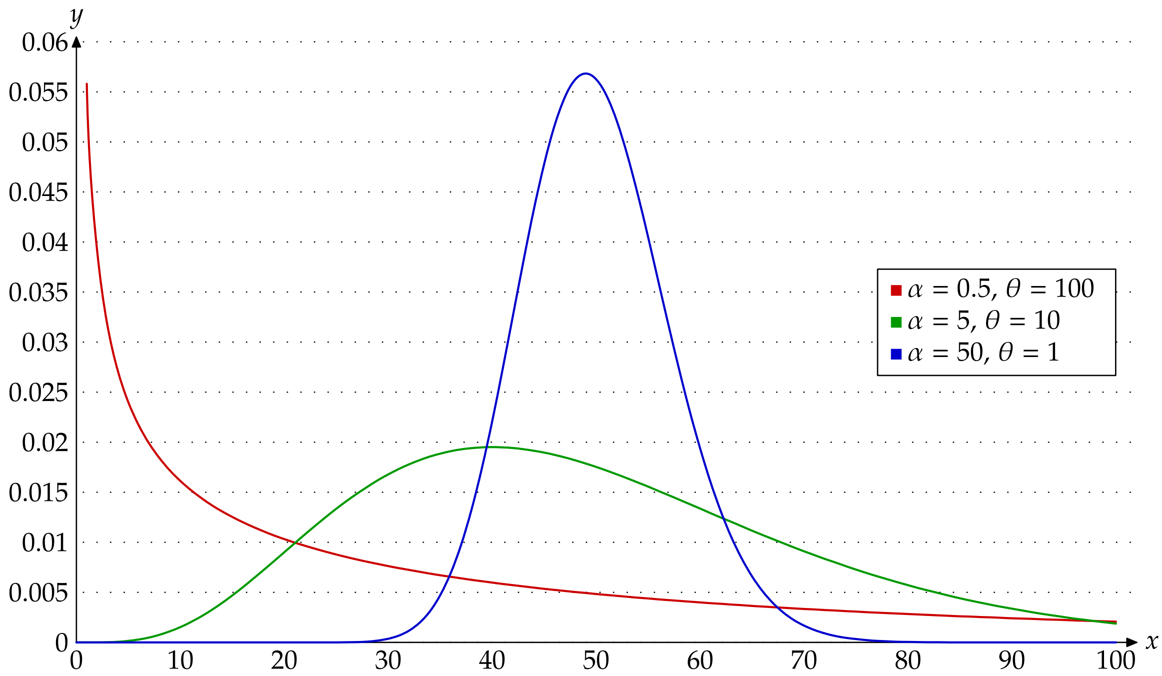


Figure 2.4: Probability density function of three gamma distributions with mean 50

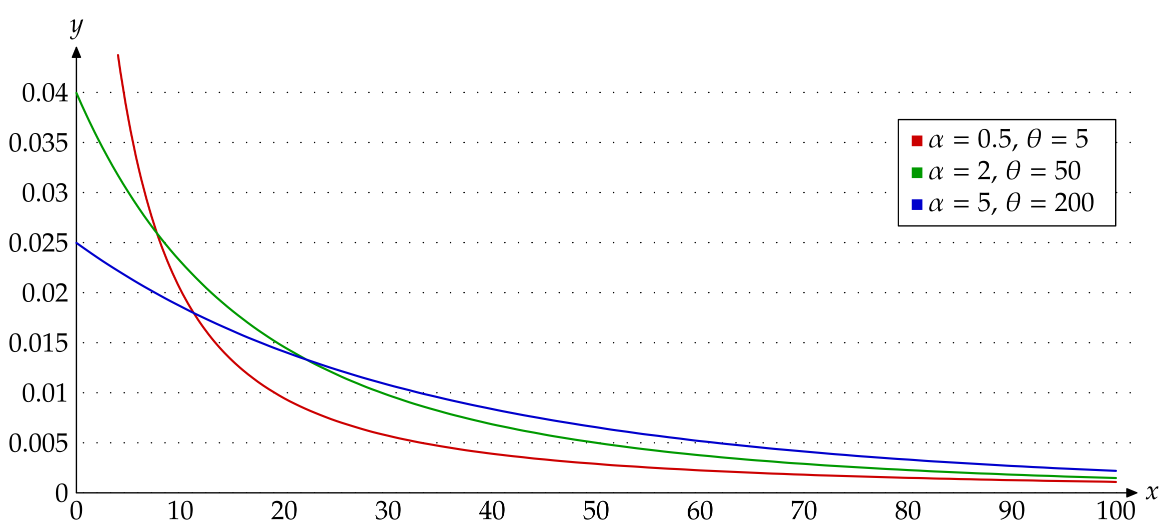


Figure 2.5: Probability density function of three Pareto distributions

2.2.7 Single-parameter Pareto

The form of the density function of a **single-parameter Pareto** is

$$f(x) = \frac{c}{x^{\alpha+1}} \qquad x \geq \theta$$

α must be positive. θ is not considered a parameter since it must be selected in advance, based on what you want the range to be.

You recognize a single-parameter Pareto by the range of nonzero values for its density function—unlike most other distributions, this range does not start at 0—and by the form of the density function, which has a denominator with x raised to a power. A beta distribution may also have x raised to a negative power, but its density function is 0 above a finite number.

A single-parameter Pareto X is a two-parameter Pareto Y shifted by θ: $X = Y + \theta$. Thus it has the same variance, and the mean is θ greater than the mean of a two-parameter Pareto with the same parameters.

$$\mathbf{E}[X] = \frac{\alpha\theta}{\alpha - 1} \qquad \alpha > 1$$
$$\mathbf{E}[X^2] = \frac{\alpha\theta^2}{\alpha - 2} \qquad \alpha > 2$$

2.2.8 Lognormal

The form of the density function of a **lognormal distribution** is

$$f(x) = \frac{ce^{-(\ln x - \mu)^2/2\sigma^2}}{x} \qquad x > 0$$

σ must be nonnegative.

You recognize a lognormal by the $\ln x$ in the exponent.

If Y is normal, then $X = e^Y$ is lognormal with the same parameters μ and σ. Thus, to calculate the distribution function, use

$$F_X(x) = F_Y(\ln x) = \Phi\left(\frac{\ln x - \mu}{\sigma}\right)$$

where $\Phi(x)$ is the **standard normal distribution function**, for which you are given tables. The moments of a lognormal are

$$\mathbf{E}[X] = e^{\mu+0.5\sigma^2}$$
$$\mathbf{E}[X^2] = e^{2\mu+2\sigma^2}$$

More generally, $\mathbf{E}[X^k] = \mathbf{E}[e^{kY}] = M_Y(k)$, where $M_Y(k)$ is the **moment generating function** of the corresponding normal distribution.

Figure 2.6 graphs three lognormals with mean 50. The mode is $\exp(\mu - \sigma^2)$, as stated in the tables. For $\mu = 2$, the mode is off the graph. As σ gets lower, the distribution flattens out.

Table 2.1 is a summary of the forms of probability density functions for common distributions.

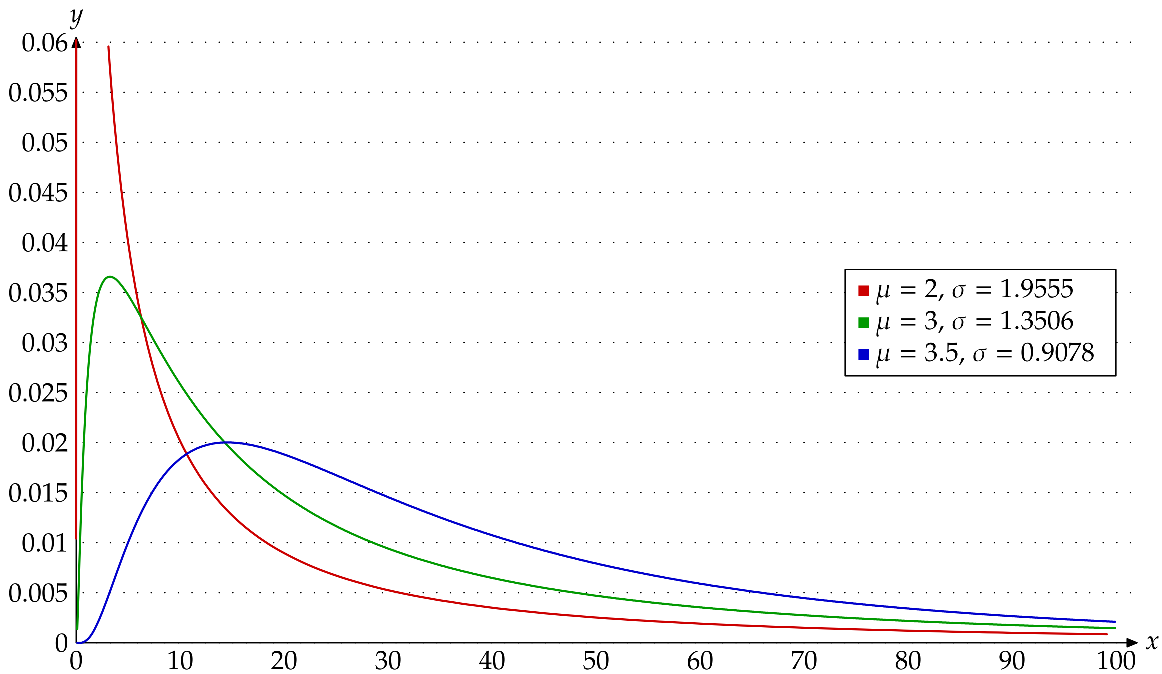


Figure 2.6: Probability density function of three lognormal distributions with mean 50

Table 2.1: Forms of probability density functions for common distributions

Distribution	Probability density function	
Uniform	c	$d \le x \le u$
Beta	$cx^{a-1}(\theta - x)^{b-1}$	$0 \le x \le \theta$
Exponential	$ce^{-x/\theta}$	$x \ge 0$
Weibull	$cx^{\tau-1}e^{-x^\tau/\theta^\tau}$	$x \ge 0$
Gamma	$cx^{\alpha-1}e^{-x/\theta}$	$x \ge 0$
Pareto	$\dfrac{c}{(x+\theta)^{\alpha+1}}$	$x \ge 0$
Single-parameter Pareto	$\dfrac{c}{x^{\alpha+1}}$	$x \ge \theta$
Lognormal	$\dfrac{ce^{-(\ln x-\mu)^2/2\sigma^2}}{x}$	$x > 0$

Exercises

2.1. X follows an exponential distribution with mean 10.

Determine the mean of X^4.

2.2. You are given

- X is exponential with mean 2.
- $Y = X^{1.5}$.

Calculate $\mathbf{E}[Y^2]$.

2.3. X follows a gamma distribution with parameters $\alpha = 2.5$ and $\theta = 10$.

$Y = 1/X$.

Calculate $\mathrm{Var}(Y)$.

2.4. **[CAS3-F05:19]** Claim size, X, follows a Pareto distribution with parameters α and θ. A transformed distribution, Y, is created such tht $Y = X^{1/\tau}$.

Determine the probability density function of Y.

A. $\dfrac{\tau\theta y^{\tau-1}}{(y+\theta)^{\tau+1}}$ B. $\dfrac{\alpha\theta^\alpha\tau y^{\tau-1}}{(y^\tau+\theta)^{\alpha+1}}$ C. $\dfrac{\theta\alpha^\theta}{(y+\alpha)^{\theta+1}}$ D. $\dfrac{\alpha\tau(y/\theta)^\tau}{y\left(1+(y/\theta)^\tau\right)^{\alpha+1}}$ E. $\dfrac{\alpha\theta^\alpha}{(y^\tau+\theta)^{\alpha+1}}$

2.5. **[CAS3-S06:27]** The following information is available regarding the random variables X and Y:

- X follows a Pareto distribution with $\alpha = 2$ and $\theta = 100$.
- $Y = \ln\big(1 + (X/\theta)\big)$

Calculate the variance of Y.

A. Less than 0.10
B. At least 0.10, but less than 0.20
C. At least 0.20, but less than 0.30
D. At least 0.30, but less than 0.40
E. At least 0.40

Additional released exam questions: MAS-I F18:24

Solutions

2.1. The k^{th} moment for an exponential is given in the tables:

$$\mathbf{E}[X^k] = k!\theta^k$$

for $k = 4$ and the mean $\theta = 10$, this is $4!(10^4) = \boxed{\mathbf{240{,}000}}$.

2.2. While Y is Weibull, you don't need to know that. It's easier to use $Y^2 = X^3$ and look up the third moment of an exponential.

$$\mathbf{E}[X^3] = 3!\theta^3 = 6(2^3) = \boxed{\mathbf{48}}$$

2.3. We calculate $\mathbf{E}[Y]$ and $\mathbf{E}[Y^2]$, or $\mathbf{E}[X^{-1}]$ and $\mathbf{E}[X^{-2}]$. Note that the special formula in the tables for integral moments of a gamma, $\mathbf{E}[X^k] = \theta^k(\alpha+k-1)\cdots\alpha$ only applies when k is a *positive* integer, so it cannot be used for the -1 and -2 moments. Instead, we must use the general formula for moments given in the tables,

$$\mathbf{E}[X^k] = \frac{\theta^k\Gamma(\alpha+k)}{\Gamma(\alpha)}$$

For $k=-1$, this is

$$\mathbf{E}[X^{-1}] = \frac{\theta^{-1}\Gamma(\alpha-1)}{\Gamma(\alpha)} = \frac{1}{\theta(\alpha-1)}$$

since $\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1)$. For $k=-2$,

$$\mathbf{E}[X^{-2}] = \frac{1}{\theta^2(\alpha-1)(\alpha-2)}$$

Therefore,

$$\operatorname{Var}(Y) = \frac{1}{10^2(1.5)(0.5)} - \left(\frac{1}{10(1.5)}\right)^2 = \boxed{\mathbf{0.00888889}}$$

2.4. Use formula (2.3). The transforming function is $g(x) = x^{1/\tau}$. The inverse of g is $g^{-1}(y) = y^\tau$. The derivative of $g^{-1}(y)$ is

$$\frac{\mathrm{d}g^{-1}(y)}{\mathrm{d}y} = \tau y^{\tau-1}$$

The density of a Pareto is

$$f_X(x) = \frac{\alpha\theta^\alpha}{(\theta+x)^{\alpha+1}}$$

Then using formula (2.3),

$$f_Y(y) = \frac{\alpha\theta^\alpha\tau y^{\tau-1}}{(\theta+y^\tau)^{\alpha+1}} \qquad \textbf{(B)}$$

2.5. We already worked this out in Example 2A. We established there that Y is exponential with mean $1/\alpha$ which is 0.5 here. Hence the variance, which is the square of the mean, is $\boxed{0.25}$. **(C)**

Practice Exam 1

1. Cars arrive at a toll booth in a Poisson process at the rate of 6 per minute.

Determine the probability that the third car will arrive between 30 and 40 seconds from now.

A. Less than 0.18
B. At least 0.18, but less than 0.21
C. At least 0.21, but less than 0.24
D. At least 0.24, but less than 0.27
E. At least 0.27

2. A business receives 50 pieces of mail every day in a Poisson process. One tenth of the mail contains checks. The logarithm of the amount of each check has a normal distribution with parameters $\mu = 3$, $\sigma^2 = 9$.

Determine the average number of checks for amounts greater than 10,000 that the business receives in a seven day week.

A. Less than 0.66
B. At least 0.66, but less than 0.69
C. At least 0.69, but less than 0.75
D. At least 0.75, but less than 0.75
E. At least 0.75

3. ATM withdrawals occur in a Poisson process at varying rates throughout the day, as follows:

11PM–6AM	3 per hour
6AM–8AM	Linearly increasing from 3 per hour to 30 per hour
8AM–5PM	30 per hour
5PM–11PM	Linearly decreasing from 30 per hour to 3 per hour

Withdrawal amounts are uniformly distributed on $(100, 500)$, and are independent of each other and the number of withdrawals.

Using the normal approximation, estimate the amount of money needed to be adequate for all withdrawals for a day 95% of the time.

A. Less than 137,500
B. At least 137,500, but less than 138,000
C. At least 138,000, but less than 138,500
D. At least 138,500, but less than 139,000
E. At least 139,000

Exam questions continue on the next page . . .

4. In a Poisson process, arrivals occur at the rate of 5 per hour.

Exactly one event has occurred within the last 20 minutes, but the time of the event is unknown.

Estimate the 90^{th} percentile of the time, in minutes, of the event

A. Less than 16 minutes
B. At least 16 minutes, but less than 17 minutes
C. At least 17 minutes, but less than 18 minutes
D. At least 18 minutes, but less than 19 minutes
E. At least 19 minutes

5. The amount of time between windstorms causing losses of 100 million or more is exponentially distributed with a mean of 10 years.

The amount of time between wildfires causing losses of 100 million or more is exponentially distributed with a mean of 6 years.

Determine the probability that at least 2 windstorms will occur before the third wildfire.

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

6. For a certain population, lifetime is exponentially distributed with mean 70. Every member of the population earns 100,000 per year from the 20^{th} birthday to the 65^{th} birthday.

Calculate expected lifetime earnings for a newborn.

A. Less than 2,700,000
B. At least 2,700,000, but less than 3,000,000
C. At least 3,000,000, but less than 3,300,000
D. At least 3,300,000, but less than 3,600,000
E. At least 3,600,000

7. An insurance company currently sells 20 million of inland marine insurance. The company has devised a strategy for expanding this line of business, but the strategy is risky. In any year, if the strategy is successful, sales will increase by 10 million. If the strategy is unsuccessful, sales will decrease by 10 million. If sales go down to 0, the company will exit the business.

The probability in each year that the strategy is successful is 2/3.

The company's goal is to increase sales to 60 million.

Calculate the probability that the company reaches its goal.

A. Less than 0.60
B. At least 0.60, but less than 0.65
C. At least 0.65, but less than 0.70
D. At least 0.70, but less than 0.75
E. At least 0.75

Exam questions continue on the next page . . .

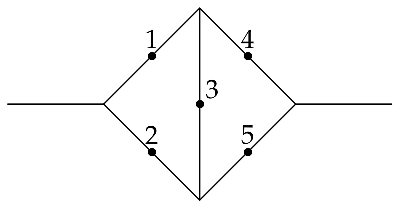
8. For a discrete irreducible Markov chain with 3 states:

- The limiting probability of state 1 is 0.6.
- The limiting probability of state 3 is 0.3.
- The probability of transition from state 2 to state 1 is 0.8.
- The probability of transition from state 3 to state 1 is 0.

Calculate the probability of staying in state 1 for one transition.

A. Less than 0.85
B. At least 0.85, but less than 0.88
C. At least 0.88, but less than 0.91
D. At least 0.91, but less than 0.94
E. At least 0.94

9. You are given the following system of 5 components:



Determine the number of minimal cut sets in this system.

A. 2 B. 3 C. 4 D. 5 E. 6

10. A graph consists of 3 nodes numbered 1, 2, 3 and arcs connecting them. The probability that an arc connects two nodes is 0.8 for nodes 1 and 2, 0.7 for nodes 1 and 3, and 0.6 for nodes 2 and 3.

Calculate the probability that the graph is connected.

A. Less than 0.65
B. At least 0.65, but less than 0.70
C. At least 0.70, but less than 0.75
D. At least 0.75, but less than 0.80
E. At least 0.80

Exam questions continue on the next page . . .

11. You are given:

- The following life table.

x	l_x	d_x
50	1000	20
51		
52		35
53		37

- ${}_2q_{52} = 0.07508$.

Determine d_{51}.

A. Less than 20
B. At least 20, but less than 22
C. At least 22, but less than 24
D. At least 24, but less than 26
E. At least 26

12. For a 30-year deferred whole life annuity on (35):

- The annuity will pay 100 per year at the beginning of each year, starting at age 65.
- If death occurs during the deferral period, the contract will pay 1000 at the end of the year of death.
- Mortality follows the Illustrative Life Table.
- $i = 0.06$.
- Y is the present value random variable for the contract.

Calculate $\mathbf{E}[Y]$.

A. Less than 204
B. At least 204, but less than 205
C. At least 205, but less than 206
D. At least 206, but less than 207
E. At least 207

Exam questions continue on the next page . . .

13. You are given:

- Loss sizes follow a paralogistic distribution with $\alpha = 3$, $\theta = 10$.
- The time of a loss follows a distribution with density function

$$f(t) = 2t \qquad 0 < t < 1$$

- Time of loss is independent of loss size.
- The interest rate is 0.06.
- Z, the present value of one loss, is simulated.
- Loss size is simulated using the random number 0.3 drawn from a uniform distribution on $[0, 1)$.
- Time of loss is simulated using the random number 0.6 drawn from a uniform distribution on $[0, 1)$.

Calculate the simulated value of Z.

A. Less than 4.75
B. At least 4.75, but less than 4.85
C. At least 4.85, but less than 4.95
D. At least 4.95, but less than 5.05
E. At least 5.05

14. For 2 estimators of θ, $\hat{\theta}$ and $\tilde{\theta}$, you are given:

-

	$\hat{\theta}$	$\tilde{\theta}$
Expected value	4	5
Variance	2	3

- $\theta = 5$
- $\mathrm{Cov}(\hat{\theta}, \tilde{\theta}) = -1$

Determine the mean square error of $\frac{1}{2}(\hat{\theta} + \tilde{\theta})$ as an estimator of θ.

A. Less than 1.25
B. At least 1.25, but less than 1.75
C. At least 1.75, but less than 2.25
D. At least 2.25, but less than 2.75
E. At least 2.75

15. The observations 4, 8, 18, 21, 49 are fitted to a distribution with density

$$f(x; \theta, d) = \frac{1}{\theta} e^{-(x-d)/\theta} \qquad x \geq d$$

by matching the first and second moments.

Determine the median of the fitted distribution.

A. Less than 13
B. At least 13, but less than 14
C. At least 14, but less than 15
D. At least 15, but less than 16
E. At least 16

Exam questions continue on the next page . . .

16. For an automobile liability coverage, you are given:

- The coverage is available with two policy limits: 10,000 and 20,000.
- Loss sizes under this coverage follow a two-parameter Pareto distribution with parameters $\alpha = 2$ and θ. The parameter θ does not vary by policy limit.
- Expected aggregate losses for the coverage with a 10,000 limit are 3/4 of expected aggregate losses for the coverage with a 20,000 limit.

Determine θ.

A. Less than 7,000
B. At least 7,000, but less than 9,000
C. At least 9,000, but less than 11,000
D. At least 11,000, but less than 13,000
E. At least 13,000

17. A sample of 6 observed claim sizes is

10 25 30 52 70 90

These observations are fitted to a Lognormal distribution with $\mu = 2$ using maximum likelihood.

Determine the variance of the fitted distribution.

A. Less than 21,000
B. At least 21,000, but less than 23,000
C. At least 23,000, but less than 25,000
D. At least 25,000, but less than 27,000
E. At least 27,000

18. For an insurance coverage with policy limit 100, there are five observed losses of sizes 30, 50, 60, 70, and 100. In addition, there are three losses for amounts above 100.

Loss sizes are fitted to a Pareto distribution with parameters $\theta = 50$ and α.

Calculate the maximum likelihood estimate of α.

A. 0.55 B. 0.58 C. 0.61 D. 0.66 E. 0.69

19. From a mortality study, you have five observations of time to death: 2, 3, 5, 8, 10.

Survival time is estimated using these observations with kernel-density smoothing. A rectangular kernel with bandwidth 2 is used.

Determine the 30$^{\text{th}}$ percentile of the kernel-density smoothed distribution.

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

20. For two baseball teams A and B:

- Team A wins 7 out of 10 games.
- Team B wins x out of 14 games.
- The null hypothesis is that the two teams are equally likely to win games.
- The alternative hypothesis is that the two teams are not equally likely to win games.

Determine the highest value of x for which the null hypothesis is not rejected at 5% significance.

A. 10 B. 11 C. 12 D. 13 E. 14

Exam questions continue on the next page . . .

21. In Territory 1, you have 130 policies and experience aggregate losses of 100,000, with sample standard deviation 2000.

In Territory 2, you have 80 policies and experience aggregate losses of 20,000, with sample standard deviation 1500.

You test the null hypothesis that underlying average aggregate losses per policy in both territories is equal. You assume that aggregate losses are normally distributed.

Determine the results of the test.

A. Reject the null hypothesis at 1% significance.
B. Reject the null hypothesis at 2.5% significance, but not at 1% significance.
C. Reject the null hypothesis at 5% significance, but not at 2.5% significance.
D. Reject the null hypothesis at 10% significance, but not at 5% significance.
E. Do not reject the null hypothesis at 10% significance.

22. A sample of 20 items from a normal distribution yields the following summary statistics:

$$\sum X_i = 120$$
$$\sum X_i^2 = 1100$$

Construct a 99% confidence interval of the form $(0, a)$ for the variance.

Determine a.

A. 10.0 B. 10.1 C. 10.5 D. 48.5 E. 49.8

23. X is a random variable having probability density function

$$f(x) = \alpha x^{\alpha-1} \qquad 0 < x < 1$$

You test H_0: $\alpha = 1$ against H_1: $\alpha > 1$ using 2 observations, x_1 and x_2.

Determine the form of the uniformly most powerful critical region for this test.

A. $x_1 + x_2 < k$
B. $x_1 + x_2 > k$
C. $x_1 x_2 < k$
D. $x_1 x_2 > k$
E. $\frac{1}{x_1} + \frac{1}{x_2} < k$

24. A random variable follows a two-parameter Pareto distribution with $\alpha = 1$ and $\theta = 5$.

Let Y be the minimum of a sample of 10 drawn from this random variable.

Calculate $\mathbf{E}[Y]$.

A. 5/12 B. 5/11 C. 5/10 D. 5/9 E. 5/8

25. Auto liability claim size is modeled using a generalized linear model. Based on an analysis of the data, it is believed that the coefficient of variation of claim size is constant.

Which of the following response distributions would be most appropriate to use?

A. Poisson B. Normal C. Gamma D. Inverse Gamma E. Inverse Gaussian

Exam questions continue on the next page . . .

26. A generalized linear model for automobile insurance with 40 observations has the following explanatory variables:

SEX (male or female)
AGE (4 levels)
TYPE OF VEHICLE (sedan, coupe, SUV, van)
MILES DRIVEN (continuous variable)
USE (business, pleasure, farm)

Model I includes all of these variables and an intercept. Model II is the same as Model I except that it excludes USE. You have the following statistics from these models:

	Deviance	AIC
Model I	23.12	58.81
Model II		62.61

Using the likelihood ratio test, which of the following statements is correct?

A. Accept USE at 0.5% significance.
B. Accept USE at 1.0% significance but not at 0.5% significance.
C. Accept USE at 2.5% significance but not at 1.0% significance.
D. Accept USE at 5.0% significance but not at 2.5% significance.
E. Reject USE at 5.0% significance.

27. You are given the following regression model, based on 22 observations.

$$\mathbf{y} = \beta_0 + \beta_1\mathbf{x}_1 + \beta_2\mathbf{x}_2 + \beta_3\mathbf{x}_3 + \beta_4\mathbf{x}_4 + \beta_5\mathbf{x}_5 + \varepsilon$$

The error sum of squares for this model is 156.

If the variables $\mathbf{x}_4$ and $\mathbf{x}_5$ are removed, the error sum of squares is 310.

Calculate the F ratio to determine the significance of the variables $\mathbf{x}_4$ and $\mathbf{x}_5$.

A. Less than 4.0
B. At least 4.0, but less than 5.0
C. At least 5.0, but less than 6.0
D. At least 6.0, but less than 7.0
E. At least 7.0

Exam questions continue on the next page . . .

28. Your company is trying to improve agent production. It is trying out 3 different training systems. Each training system is tried out on 20 agents, and their production is compared to a control group of 20 agents.

You test the hypothesis

H_0: There is no difference in production among the 4 groups

against

H_1: There are differences in production among the 4 groups.

You run two linear regression models:

1. A model with a categorical variable for group and no intercept.
2. A model with a constant only

The scaled deviance for the first model is 589.1 and the scaled deviance of the second model is 622.3.

Determine the F statistic to test the hypotheses.

A. Less than 1.00
B. At least 1.00, but less than 1.25
C. At least 1.25, but less than 1.50
D. At least 1.50, but less than 1.75
E. At least 1.75

29. Which of the following statements are true regarding goodness-of-fit testing for a logistic regression?

I. The chi-square distribution is a poor approximation for deviance if cell frequencies are too low.
II. The Hosmer-Lemeshow method is a method of calculating the deviance statistic when cell frequencies are low.
III. PseudoR^2 often is alarmingly low even when other measures indicate the model fits well.

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

30. A normal linear model with 2 variables and an intercept is based on 45 observations. $\hat{y}_j$ is the fitted value of y_j, and $\hat{y}_{j(i)}$ is the fitted value of y_j if observation i is removed. You are given:

- $\sum_{j=1}^{45}(y_j - y_{j(1)})^2 = 4.1.$
- The leverage of the first observation is 0.15.

Determine $|\hat{\varepsilon}_1|$, the absolute value of the first residual of the regression with no observation removed.

A. Less than 4
B. At least 4, but less than 5
C. At least 5, but less than 6
D. At least 6, but less than 7
E. At least 7

31. A gamma regression model with $\alpha = 1$ is fitted to data. The identity link is used.

This regression is equivalent to a weighted least squares regression.

Express the weights, the entries in the diagonal matrix **W**, as a function of μ_i, the mean of the response variable.

A. $1/\mu_i^2$ B. $1/\mu_i$ C. $\ln \mu_i$ D. μ_i E. μ_i^2

Exam questions continue on the next page . . .

32. You are given the following output from a GLM to estimate loss size:

- Distribution selected is Inverse Gaussian.
- The link is $g(\mu) = 1/\mu^2$.

Parameter	β
Intercept	0.00279
Vehicle Body	
Coupe	0.002
Sedan	−0.001
SUV	0.003
Vehicle Value (000)	−0.00007
Area	
B	−0.025
C	0.015
D	0.005

Calculate mean loss size for a sedan with value 25,000 from Area A.

A. Less than 100
B. At least 100, but less than 500
C. At least 500, but less than 1000
D. At least 1000, but less than 5000
E. At least 5000

33. The response variable of a generalized linear model follows a normal distribution. The link is $g(\mu) = \ln \mu$. The method of scoring is used to fit the coefficients. At each iteration, weighted least squares is performed. Which of the following is proportional to the weights?

A. $1/\mu_i^2$ B. $1/\mu_i$ C. 1 D. μ_i E. μ_i^2.

34. A generalized linear model of the form

$$\sqrt{\mu} = \beta_1 + \beta_2 \mathbf{x}_2 + \beta_3 \mathbf{x}_3 + \varepsilon$$

is estimated based on 20 observations. The resulting estimate of $\boldsymbol{\beta}$ is $b_1 = 1.80$, $b_2 = 3.28$, $b_3 = 3.21$. You are given that

$$(\mathbf{X}^T \mathbf{W} \mathbf{X})^{-1} = \begin{pmatrix} 0.85 & 0.75 & 0.23 \\ 0.75 & 1.60 & 0.85 \\ 0.23 & 0.85 & 1.32 \end{pmatrix}$$

Based on the Wald statistic, which of the β parameters are significant at the 5% level?

A. β_2 only B. β_3 only C. β_1 and β_2 only D. β_2 and β_3 only E. β_1, β_2, and β_3

Exam questions continue on the next page . . .

35. For an inverse Gaussian regression, you are given

- $y_5 = 652$.
- $\hat{y}_5 = 530$
- The inverse Gaussian has parameter $\theta = 1/2$

Calculate the deviance residual of the fifth observation, d_5.

A. Less than 0.01
B. At least 0.01, but less than 0.02
C. At least 0.02, but less than 0.03
D. At least 0.03, but less than 0.04
E. At least 0.04

36. For a generalized linear model,

- There are 72 observations.
- There are 25 parameters.
- The loglikelihood is -361.24

You are considering adding a cubic polynomial variable.

Determine the lowest loglikelihood for which this additional variable would not be rejected at 1% significance.

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

37. An insurance company is modeling the probability of a claim using logistic regression. The explanatory variable is vehicle value. Vehicle value is banded, and the value of the variable is 1, 2, 3, 4, 5, or 6, depending on the band. Band 1 is the reference level.

The fitted value of the β corresponding to band 4 is -0.695.

Let O_1 be the odds of a claim for a policy in band 1, and O_4 the odds of a claim for a policy in band 4.

Determine O_4/O_1.

A. Less than 0.35
B. At least 0.35, but less than 0.40
C. At least 0.40, but less than 0.45
D. At least 0.45, but less than 0.50
E. At least 0.50

Exam questions continue on the next page . . .

38. A generalized linear model is used to estimate the probability of death upon consuming a quantity x of a toxic chemical. The tolerance distribution is the normal distribution

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

The probability p of death is modeled using

$$g(p) = \beta_1 + \beta_2 x$$

where g is the link and x is the dose.

The fitted values of the parameters are $\beta_1 = -0.8$ and $\beta_2 = 0.4$.

Calculate the estimated median lethal dose.

A. Less than 0.4
B. At least 0.4, but less than 0.8
C. At least 0.8, but less than 1.2
D. At least 1.2, but less than 1.6
E. At least 1.6

39. A regression is performed on 4 observations. Results of the regression are:

x	y	$\hat{y}$	Leverage
1	3	2.658	0.487
2	2	3.105	0.355
4	5	4.000	0.250
9	6	6.237	0.908

Calculate the cross-validation error rate generated when using LOOCV.

A. Less than 2.7
B. At least 2.7, but less than 2.9
C. At least 2.9, but less than 3.1
D. At least 3.1, but less than 3.3
E. At least 3.3

40. You have fitted a logistic model with no predictors, just a constant. To determine the accuracy of the fitted constant, you use the bootstrap method. Based on 4 bootstrap samples, the fitted constants are 0.8, −0.2, 0.4, 0.2.

Calculate the bootstrap estimate of the accuracy of the generated binomial probability.

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

Exam questions continue on the next page . . .

41. A model with 50 observations has 7 predictors. We would like to select an optimal subset of these predictors. As part of this selection, we will have to consider models with 5 predictors.

Let

n_1 be the number of such models that need to be considered using best subset selection.

n_2 be the number of such models that need to be considered using forward stepwise selection.

n_3 be the number of such models that need to be considered using backward stepwise seletion.

Determine $n_1 + n_2 + n_3$.

42. Consider the vector $\{5, -3, 8, -2, 4\}$.

Calculate the absolute difference between the ℓ_2 norm and ℓ_1 norm of this vector.

A. Less than 12
B. At least 12, but less than 15
C. At least 15, but less than 18
D. At least 18, but less than 21
E. At least 21

43. Which of the following statements are true?

I. Partial Least Squares is a supervised method of dimension reduction.
II. Partial Least Squares directions are linear combinations of the original variables.
III. Partial Least Squares can be used for feature selection.

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

44. A least squares model with a large number of predictors is fitted to 90 observations. To reduce the number of predictors, forward stepwise selection is performed.

For a model with k predictors, RSS = c_k.

The estimated variance of the error of the fit is $\hat{\sigma}^2 = 40$.

Determine the value of $c_d - c_{d+1}$ for which you would be indifferent between the $d + 1$-predictor model and the d-predictor model based on Mallow's C_p.

A. Less than 30
B. At least 30, but less than 45
C. At least 45, but less than 60
D. At least 60, but less than 75
E. At least 75

45. A natural cubic regression spline is fitted with 6 degrees of freedom.

Determine the number of internal knots.

A. 4 B. 5 C. 6 D. 7 E. 8

Appendix A. Solutions to the Practice Exams

Answer Key for Practice Exam 1

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

Practice Exam 1

1. [**Lesson 13**] The probability that the third car will arrive in the interval $(30, 40)$ is the probability of at least 3 cars in 40 seconds minus the probability of at least 3 cars in 30 seconds. For 40 seconds, the Poisson parameter is 4 and the probability is

$$1 - e^{-4}\left(1 + 4 + \frac{4^2}{2}\right) = 1 - 0.238103$$

For 30 seconds, the Poisson parameter is 3 and the probability is

$$1 - e^{-3}\left(1 + 3 + \frac{3^2}{2}\right) = 1 - 0.423190$$

The difference is $0.423190 - 0.238103 = \boxed{0.185087}$. **(B)**

2. [**Lesson 15**] The probability of a check greater than 10,000 is

$$1 - \Phi\left(\frac{\ln 10{,}000 - 3}{3}\right) = 1 - \Phi(2.07) = 1 - 0.9808 = 0.0192$$

The Poisson distribution of just the checks over 10,000 in one week has parameter $7(50)(0.1)(0.0192) = \boxed{0.672}$. **(B)**

3. [**Lesson 18**] The Poisson parameter per day is computed by adding up the rates over the 4 periods. For 11PM–6AM, we have 7 hours times 3 per hour, or 21. For 8AM–5PM we have 9 hours times 30 per hour, or 270. For the other two periods, because of the linear increase or decrease, the average per hour is the midpoint, or $(30 + 3)/2 = 16.5$, and there are 8 hours with varying rates, for a total of $8 \times 16.5 = 132$. The total number of withdrawals per day is $21 + 270 + 132 = 423$. The mean aggregate withdrawals is $(423)(300) = 126{,}900$.

The second moment of the uniform distribution on $(100, 500)$ is the variance plus the mean squared. The variance of a uniform distribution is the range squared divided by 12, or $400^2/12$. Therefore, the second moment of the uniform distribution is $400^2/12 + 300^2 = 103{,}333\frac{1}{3}$. The variance of aggregate withdrawals, by the compound variance formula (18.2), is $\lambda\,\mathbf{E}[X^2] = (423)(103{,}333\frac{1}{3}) = 43{,}710{,}000$.