ACTEX

LEARN TODAY. LEAD TOMORROW.

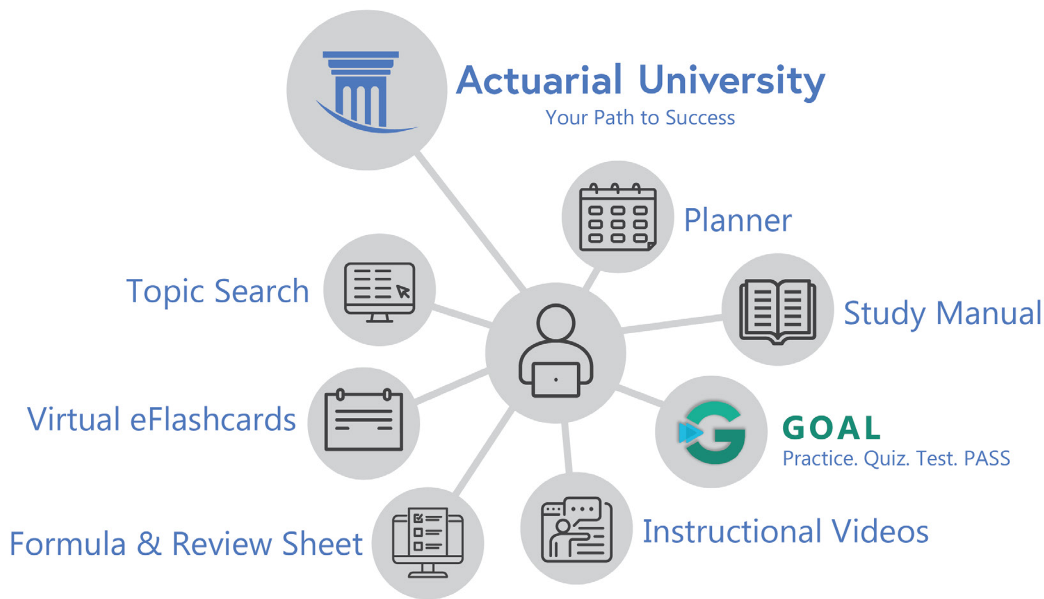
EXAM FAM-S

Study Manual for SOA Exam FAM-S

1st Edition, 2nd Printing

by
Sam A. Broverman Ph.D., ASA

Your Integrated Study Manual Program Includes:



ACTEX

Study Manual for SOA Exam FAM-S

1st Edition, 2nd Printing

by

Sam A. Broverman Ph.D., ASA

ACTEX Learning

Learn Today. Lead Tomorrow.

Actuarial & Financial Risk Resource Materials

Since 1972

Printed in the United States of America.

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.

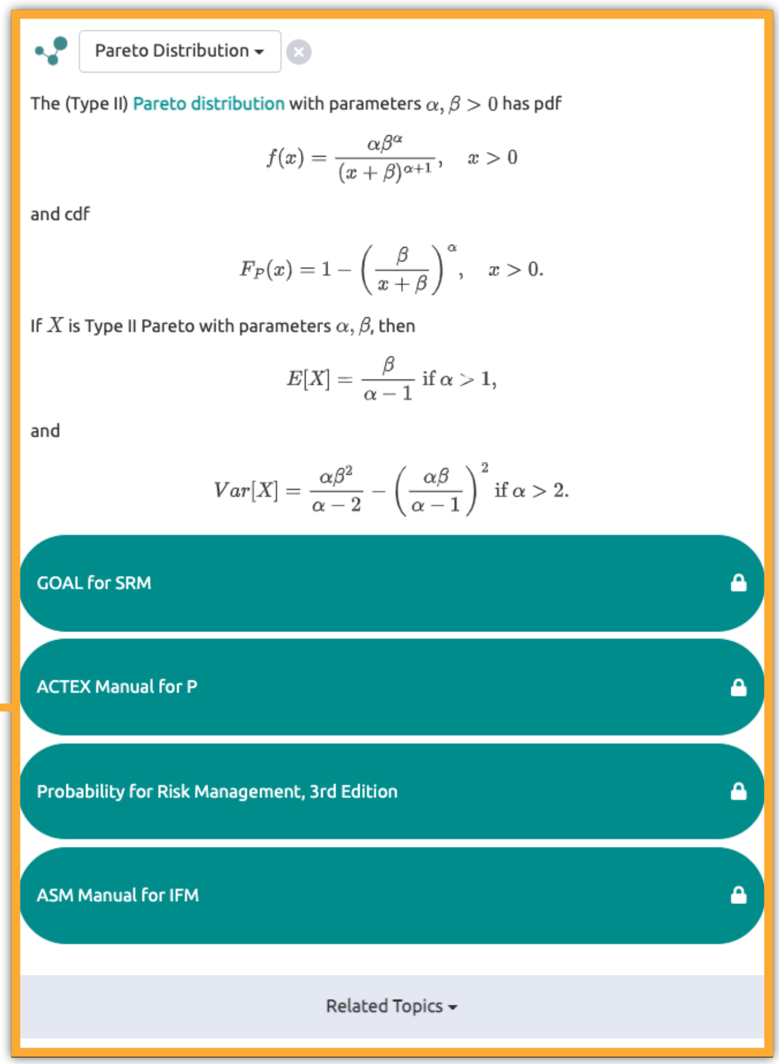
You can find integrated topics using this network icon.
When this icon appears, it will be next to an important topic in the manual! **Click the link** in your digital manual, or **search the underlined topic** in your print manual.

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*

Contents

INTRODUCTORY COMMENTS

The FAM exam is divided almost equally into FAM-S and FAM-L topics. This study guide is designed to help in the preparation for the Society of Actuaries FAM-S Exam.

The first part of this manual consists of a summary of notes, illustrative examples and problem sets with detailed solutions. The second part consists of 4 practice exams. The SOA exam syllabus for the FAM exam indicates that the exam is 3.5 hours in length with 40 multiple choice questions. The practice exams in this manual each have 20 questions, reflecting the fact that FAM-S is 50% of the full FAM exam. The appropriate time for the 20 question FAM-S practice exams in this manual is one hour and forty-five minutes.

The level of difficulty of the practice exam questions has been designed to be similar to those on past exams covering the same topics. The practice exam questions are not from old SOA exams.

I have attempted to be thorough in the coverage of the topics upon which the exam is based, and consistent with the notation and content of the official references. I have been, perhaps, more thorough than necessary on reviewing maximum likelihood estimation.

Because of the time constraint on the exam, a crucial aspect of exam taking is the ability to work quickly. I 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 you 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. I 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 32 sections of varying lengths, with some suggested time frames for covering the material. There are almost 180 examples in the notes and over 440 exercises in the problem sets, all with detailed solutions. The 4 practice exams have 20 questions each, also with detailed solutions. Some of the examples and exercises are taken 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 almost 700 examples/problems/sample exam questions with detailed solutions. ACTEX gratefully acknowledges the SOA for allowing the use of their exam problems in this study guide.

I suggest that you work through the study guide by studying a section of notes and then attempting the exercises in the problem set that follows that section. The order of the sections of notes is the order that I recommend in covering the material, although the material on pricing and reserving in Sections 27 to 31 is independent of the other material on the exam. The order of topics in this manual is not the same as the order presented on the exam syllabus.

It has been my intention to make this study guide self-contained and comprehensive for the FAM-S Exam topics, however it is important to be familiar with 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 detail on formula derivations.

In order for the review notes in this study guide to be most effective, you should have some background at the junior or senior college level in probability and statistics. It will be assumed that you are reasonably familiar with differential and integral calculus. The prerequisite concepts to modeling and model estimation are reviewed in this study guide. The study guide begins with a detailed review of probability distribution concepts such as distribution function, hazard rate, expectation and variance. Of the various calculators that are allowed for use on the exam, I am 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 me directly at the address below. I 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 my attention. ACTEX will be maintaining a website for errata that can be accessed from www.actexmadriver.com. It is my sincere hope that you find this study guide helpful and useful in your preparation for the exam. I wish you the best of luck on the exam.

Samuel A. Broverman July 2022

Department of Statistical Sciences www.sambroverman.com

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

Section 2

Review of Random Variables I

Probability, Density and Distribution Functions

This section covers preliminary topics needed for the later FAM syllabus material and Section 2.7 relates to Section 4.2.4 in the "Loss Models" book. 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 ;

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

Differentiation

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

Product rule: $\dfrac{d}{dx}[g(x) \times h(x)] = g'(x) \times h(x) + g(x) \times h'(x)$ (2.3)

Quotient rule: $\dfrac{d}{dx}\left[\dfrac{g(x)}{h(x)}\right] = \dfrac{h(x) \times g'(x) - g(x) \times h'(x)}{[h(x)]^2}$ (2.4)

Chain rule: $\dfrac{d}{dx} f(g(x)) = f'(g(x)) \times g'(x)$, $\dfrac{d}{dx} \ln[g(x)] = \dfrac{g'(x)}{g(x)}$

$\dfrac{d}{dx}[g(x)]^n = n \times [g(x)]^{n-1} \times g'(x)$, $\dfrac{d}{dx} a^x = a^x \times \ln(a)$ (2.5)

Integration

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

Integration by parts

for definite integrals: $\int_a^b u(t)\ dv(t) = u(b) \times v(b) - u(a) \times v(a) - \int_a^b v(t)\ du(t)$ (2.7)

for indefinite integrals: $\int u\ dv = uv - \int v\ du$

(this is derived by integrating both sides of the product rule).

Note that $dv(t) = v'(t)\,dt$ and $du(t) = u'(t)\,dt$.

Some additional relationships involving integration:

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

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

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

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 FAM 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:

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

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

and for Y it is $p(k) = \dfrac{1}{2^k}$ 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 manual, 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 $k = 1, \cdots N$ 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 $P[X = k] = p(k) = \binom{m}{k} q^k (1-q)^{m-k}, k = 0{,}1{,}...{,}m$ 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
$P[X = k] = \frac{e^{-\lambda}\lambda^k}{k!}$ for $k \geq 0$, 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.

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 $\int_a^b f(x)\, dx$.

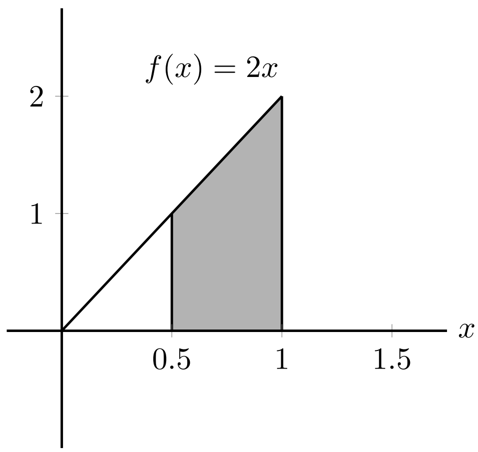
A pdf $f(x)$ must satisfy (i) $f(x) \geq 0$ for all x and (ii) $\int_{-\infty}^{\infty} f(x)\, dx = 1$ (2.12)

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 $f(x) = \begin{cases} 2x & \text{for } \ 0 < x < 1 \\ 0 & \text{elsewhere} \end{cases}$.

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,

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

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 manual, 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 mean 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

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

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

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 <u>and including</u> 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 $X \geq 0$, then the survival function satisfies

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

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.

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.

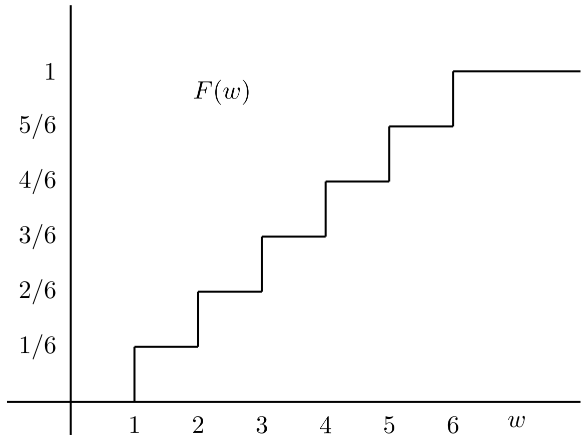
Example 2.1. Finite Discrete Random Variable (finite support)

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 $w \geq 6$).



□

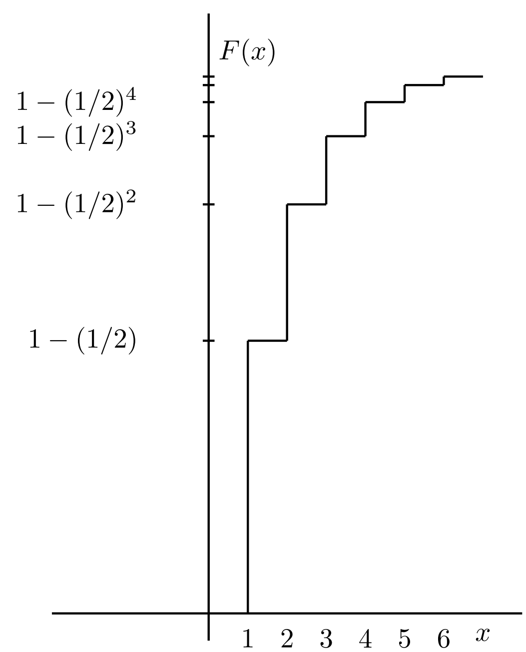
Example 2.2. Infinite Discrete Random Variable (infinite support)

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 $F_X(x) = \sum_{k=1}^{x} \dfrac{1}{2^k} = 1 - \dfrac{1}{2^x}$ 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



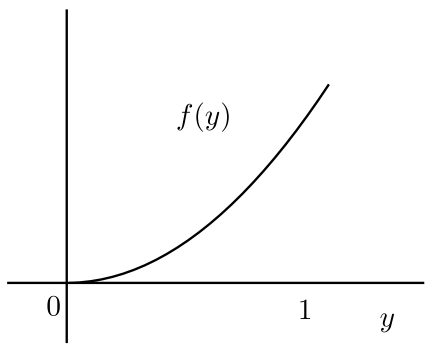
□

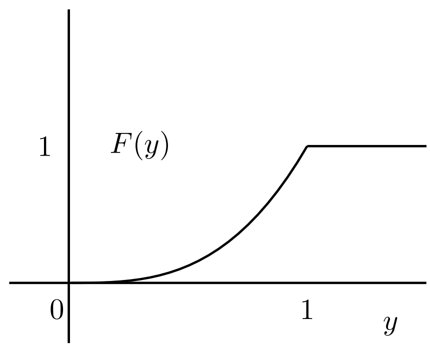
Example 2.3.

Continuous Random Variable on a Finite Interval

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

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





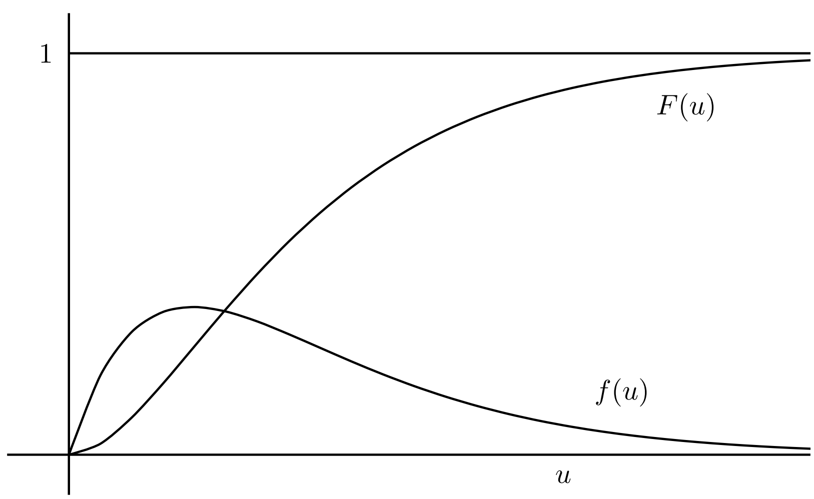
□

Example 2.4.

Continuous Random Variable on an Infinite Interval

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

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

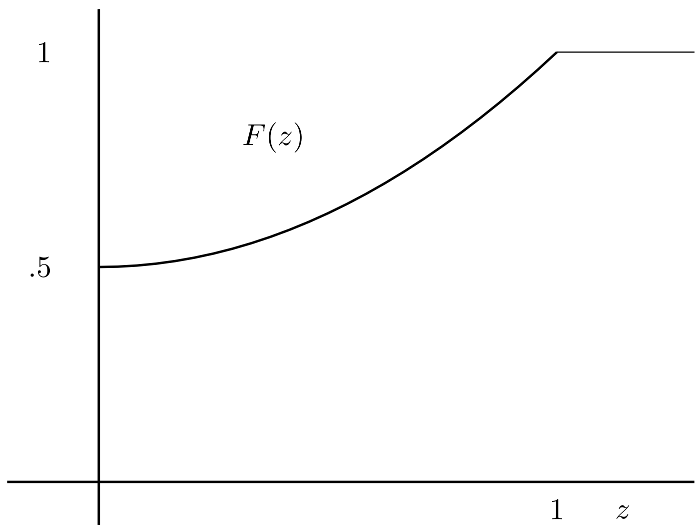


□

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 $y_1 < y_2 < \ldots < y_k$, 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,...,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:

$$F_n(t) = \frac{\text{number of } x_i\text{'s} \le t}{n}. \tag{2.19}$$

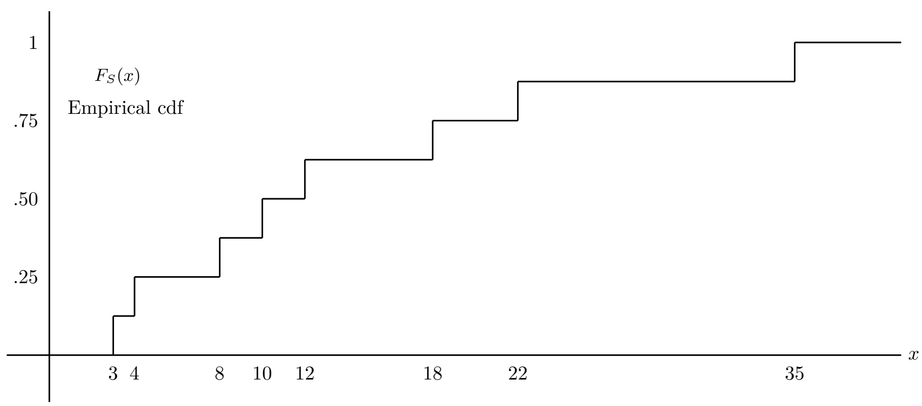
In the example above, $F_8(4) = \frac{5}{8}$. The subscript "8" in F_8 just indicates the total number of data points in the sample.

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. $F_8(t) = \frac{\text{number of } x_i\text{'s} \le t}{8}$.
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 FAM Exam Tables make reference to the **gamma function** and the **incomplete gamma function**. The definitions of these functions are

- **gamma function**: $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t}\, dt \quad \text{for} \quad \alpha > 0$

- **incomplete gamma function**: $\Gamma(\alpha; x) = \dfrac{1}{\Gamma(\alpha)} \times \int_0^x t^{\alpha-1} e^{-t}\, dt \quad \text{for} \quad \alpha > 0,\, x > 0$ (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 $\Gamma(\alpha+k) = (\alpha+k-1) \times (\alpha+k-2) \cdots \times \alpha \times \Gamma(\alpha)$ (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**:

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

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 $\int_0^\infty x^k e^{-cx}\, dx = \dfrac{\Gamma(k+1)}{c^{k+1}}$, particularly in the case in which k is a non-negative integer. In that case, we get $\boldsymbol{\int_0^\infty x^k e^{-cx}\, dx = \dfrac{k!}{c^{k+1}}}$, which can occasionally simplify integral relationships. This relationship is embedded in the definition of the gamma distribution in the FAM Exam Table.

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

If we let $\theta = \dfrac{1}{c}$ 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 FAM 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, $\int_0^1 \dfrac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \times x^{a-1}(1-x)^b\, dx = 1$, 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

$P[X = x] = \dfrac{2}{3^x}$ for $x = 1,2,3,\ldots$. 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 $P(k+1) = \dfrac{1}{k} \times P(k)$ 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

$f(x) = \begin{cases} 6x(1-x) & \text{for } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$. Calculate $P[|X - \frac{1}{2}| > \frac{1}{4}]$.

(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

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

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

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

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

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) $\frac{1}{2}$ (B) $\left(\frac{7}{8}\right)^{1/3}$ (C) $\left(\frac{8}{7}\right)^{1/3}$ (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) $\frac{1}{\sqrt{\pi}}$ (B) $\frac{1}{\pi}$ (C) $\frac{1}{4}$ (D) $\frac{1}{8}$ (E) $\frac{1}{\pi^2}$

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

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

i) The hazard rate function: $h(x) = \frac{z^2}{2x}$, 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

Section 2 Problem Set Solutions

1. $P[X \text{ is even}] = P[X=2] + P[X=4] + P[X=6] + \cdots = \frac{2}{3} \times [\frac{1}{3} + \frac{1}{3^3} + \frac{1}{3^5} + \cdots] = \frac{2}{3^2} \times \frac{1}{1-\frac{1}{3^2}} = \frac{1}{4}$

Answer A

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

The probability function must satisfy the requirement $\sum_{i=0}^{\infty} P(i) = 1$ 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, $P(0) = \dfrac{1}{e+1}$.

Answer C

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

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

Answer C

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

Answer E

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

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 $(P[X > 1])^2 \times P[X \le 1] = \Big(\dfrac{3}{2} - \sqrt{2}\Big)^2 \Big(\sqrt{2} - \dfrac{1}{2}\Big)$

for a total probability of $3 \times \Big(\dfrac{3}{2} - \sqrt{2}\Big)^2 \Big(\sqrt{2} - \dfrac{1}{2}\Big)$.

Answer E

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

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

Answer E

7. $f(x) = F'(x) = -\sum_{k=0}^{3} \dfrac{kx^{k-1}e^{-x} - x^k e^{-x}}{k!} = e^{-x} \times \sum_{k=0}^{3} \Big[\dfrac{x^k - kx^{k-1}}{k!}\Big].$

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

Answer C

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

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

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

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

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

Answer A

Practice Exam 1

1. Each of decision-makers X, Y, and Z has the opportunity to participate in a game with payoff uniformly distributed on (0, 10000). Assume that X, Y and Z value assets of amount $w \geq 0$ according to the following utility functions:

Decision Maker	Utility Function
X	$u_X(w) = \sqrt{w}$
Y	$u_Y(w) = \frac{w}{100}$
Z	$u_Z(w) = \left(\frac{w}{100}\right)^2$

Which decision maker would not be willing to pay more than 5,000 to participate in the game?

(A) X and Y only

(B) X and Z only

(C) Y and Z only

(D) X, Y and Z

(E) The correct answer is not given by (A), (B), (C) or (D)

2. The XYZ Insurance Company sells property insurance policies with a deductible of \$5,000, policy limit of \$500,000, and a coinsurance factor of 80%. Let X_i be the individual loss amount of the ith claim and Y_i be the claim payment of the ith claim. Which of the following represents the relationship between X_i and Y_i?

(A) $Y_i = \begin{cases} 0 & X_i \leq 5{,}000 \\ 0.80\,(X_i - 5{,}000) & 5{,}000 < X_i \leq 625{,}000 \\ 500{,}000 & X_i > 625{,}000 \end{cases}$

(B) $Y_i = \begin{cases} 0 & X_i \leq 5{,}000 \\ 0.80\,(X_i - 4{,}000) & 4{,}000 < X_i \leq 500{,}000 \\ 500{,}000 & X_i > 500{,}000 \end{cases}$

(C) $Y_i = \begin{cases} 0 & X_i \leq 5{,}000 \\ 0.80\,(X_i - 5{,}000) & 5{,}000 < X_i \leq 630{,}000 \\ 500{,}000 & X_i > 630{,}000 \end{cases}$

(D) $Y_i = \begin{cases} 0 & X_i \leq 6{,}250 \\ 0.80\,(X_i - 6{,}250) & 6{,}250 < X_i \leq 631{,}500 \\ 500{,}000 & X_i > 631{,}500 \end{cases}$

(E) $Y_i = \begin{cases} 0 & X_i \leq 5{,}000 \\ 0.80\,(X_i - 5{,}000) & 5{,}000 < X_i \leq 505{,}000 \\ 500{,}000 & X_i > 505{,}000 \end{cases}$

3. A casino has a game that makes payouts at a Poisson rate of 5 per hour and the payout amounts can be any non-negative integer, 1, 2, 3, ... without limit. The probability that any given payout is equal to $i > 0$ is $\frac{1}{2^i}$. Payouts are independent. Calculate the probability that there are no payouts of 1,2, or 3 in a given 20 minute period.

(A) 0.08 (B) 0.13 (C) 0.18 (D) 0.23 (E) 0.28

4. Zoom Buy Tire Store, a nationwide chain of retail tire stores, sells 2,000,000 tires per year of various sizes and models. Zoom Buy offers the following road hazard warranty: "If a tire sold by us is irreparably damaged in the first year after purchase, we'll replace it free, regardless of the cause."

The average annual cost of honoring this warranty is \$10,000,000, with a standard deviation of \$40,000. Individual claim counts follow a binomial distribution, and the average cost to replace a tire is \$100. All tires are equally likely to fail in the first year, and tire failures are independent. Calculate the standard deviation of the replacement cost per tire.

(A) Less than \$60

(B) At least \$60, but less than \$65

(C) At least \$65, but less than \$70

(D) At least \$70, but less than \$75

(E) At least \$75

5. A compound Poisson claim distribution has $\lambda = 3$ and individual claims amounts distributed as follows:

x	$f_X(x)$
5	0.6
10	0.4

Determine the expected cost of an aggregate stop-loss insurance with a deductible of 6.

(A) Less than 15.0

(B) At least 15.0 but less than 15.3

(C) At least 15.3 but less than 15.6

(D) At least 15.6 but less than 15.9

(E) At least 15.9

Use the following information for questions 6 and 7. You are the producer of a television quiz show that gives cash prizes. The number of prizes, N, and prize amounts, X are independent of one another and have the following distributions:

N: $P[N=1]=0.8$, $P[N=2]=0.2$

X: $P[X=0]=0.2$, $P[X=100]=0.7$, $P[X=1000]=0.1$

6. Your budget for prizes equals the expected prizes plus $1.25\times$ standard deviation of prizes. Calculate your budget.

(A) 384 (B) 394 (C) 494 (D) 588 (E) 596

7. You buy stop-loss insurance for prizes with a deductible of 200. The cost of insurance includes a 175% relative security load (the relative security load is the percentage of expected payment that is added). Calculate the cost of the insurance.

(A) 204 (B) 227 (C) 245 (D) 273 (E) 357

8. An actuary determines that claim counts follow a negative binomial distribution with unknown β and r. It is also determined that individual claim amounts are independent and identically distributed with mean 700 and variance 1,300. Aggregate losses have a mean of 48,000 and variance 80 million. Calculate the values for β and r.

(A) $\beta=1.20, r=57.19$

(B) $\beta=1.38, r=49.75$

(C) $\beta=2.38, r=28.83$

(D) $\beta=1{,}663.81, r=0.04$

(E) $\beta=1{,}664.81, r=0.04$

9. Let X_1, X_2, X_3 be independent Poisson random variables with means $\theta, 2\theta$, and 3θ respectively. What is the maximum likelihood estimator of θ based on sample values x_1, x_2, and x_3 from the distributions of X_1, X_2 and X_3, respectively,

(A) $\dfrac{\bar{x}}{2}$

(B) $\bar{x}$

(C) $\dfrac{x_1+2x_2+3x_3}{6}$

(D) $\dfrac{3x_1+2x_2+x_3}{6}$

(E) $\dfrac{6x_1+3x_2+2x_3}{11}$

10. Loss random variable X has a uniform distribution on $(0,\theta)$. A sample is taken of n insurance payments from policies with a limit of 100. Eight of the sample values are limit payments of 100. The maximum likelihood estimate of θ is $\widehat{\theta}$. Another sample is taken, also of n insurance payments, but from policies with a limit of 150. Three of the sample values are limit payments of 150. The maximum likelihood estimate of θ is $\frac{4}{3}\widehat{\theta}$.

Determine n.

(A) 40 (B) 42 (C) 44 (D) 46 (E) 48

11. For a group of policies, you are given:

(i) Losses follow a uniform distribution on the interval $(0, \theta)$, where $\theta > 25$.

(ii) A sample of 20 losses resulted in the following:

Interval	Number of Losses
$x \leq 10$	n_1
$10 < x \leq 25$	n_2
$x > 25$	n_3

The maximum likelihood estimate of θ can be written in the form $25 + y$. Determine y.

(A) $\dfrac{25n_1}{n_2 + n_3}$ (B) $\dfrac{25n_2}{n_1 + n_3}$

(C) $\dfrac{25n_3}{n_1 + n_2}$ (D) $\dfrac{25n_1}{n_1 + n_2 + n_3}$

(E) $\dfrac{25n_2}{n_1 + n_2 + n_3}$

12. The number of claims follows a negative binomial distribution with parameters β and r, where β is unknown and r is known. You wish to estimate β based on n observations, where $\bar{x}$ is the mean of these observations. Determine the maximum likelihood estimate of β.

(A) $\dfrac{\bar{x}}{r^2}$ (B) $\dfrac{\bar{x}}{r}$ (C) $\bar{x}$ (D) $r\bar{x}$ (E) $r^2\bar{x}$

13. The following 6 observations are assumed to come from the continuous distribution with pdf $f(x;\theta) = \frac{1}{2}x^2\theta^3 e^{-\theta x}$: 1, 3, 4, 4, 5, 7.

Find the mle of θ.

(A) 0.25 (B) 0.50 (C) 0.75 (D) 1.00 (E) 1.25

14. An analysis of credibility premiums is being done for a particular compound Poisson claims distribution, where the criterion is that the total cost of claims is within 5% of the expected cost of claims with a probability of 90%. It is found that with $n = 60$ exposures (periods) and $\bar{X} = 180.0$, the credibility premium is 189.47. After 20 more exposures (for a total of 80) and revised $\bar{X} = 185$, the credibility premium is 190.88. After 20 more exposures (for a total of 100) the revised $\bar{X}$ is 187.5. Assuming that the manual premium remains unchanged in all cases, and assuming that full credibility has not been reached in any of the cases, find the credibility premium for the 100 exposure case.

(A) 191.5 (B) 192.5 (C) 193.5 (D) 194.5 (E) 196.5

15. For an insurance portfolio, you are given:

(i) For each individual insured, the number of claims follows a Poisson distribution.

(ii) The mean claim count varies by insured, and the distribution of mean claim counts follows gamma distribution.

(iii) For a random sample of 1000 insureds, the observed claim counts are as follows:

Number of Claims, n	0	1	2	3	4	5
Number of Insureds, f_n	512	307	123	41	11	6

$$\sum n f_n = 750, \qquad \sum n^2 f_n = 1494$$

(iv) Claim sizes follow a Pareto distribution with mean 1500 and variance 6,750,000.

(v) Claim sizes and claim counts are independent.

(vi) The full credibility standard is to be within 5% of the expected aggregate loss 95% of the time.

Determine the minimum number of insureds needed for the aggregate loss to be fully credible.

(A) Less than 8300

(B) At least 8300, but less than 8400

(C) At least 8400, but less than 8500

(D) At least 8500, but less than 8600

(E) At least 8600

Information on Questions 16 and 17 is as follows. You are given the following information on cumulative incurred losses through development years shown.

	Cumulative Incurred Losses				Paid-to-Date
Accident	Development Year				at Dec 31, AY4
Year	0	1	2	3	
AY1	2325	3749	4577	4701	4701
AY2	2657	4438	5529		4500
AY3	2913	4995			3500
AY4	3163				2500

16. Using an average factor model, calculate the estimated total loss reserve as of Dec. 31, AY4.

(A) Less than 7500

(B) At least 7500 but less than 7800

(C) At least 7800 but less than 8100

(D) At least 8100 but less than 8400

(E) At least 8400

17. As of Dec. 31, AY4, calculate

Estimated reserve for AY2 based on an average factor model)
− (Estimated reserve for AY2 based on a mean factor model)

- (A) Less than −300
- (B) At least −300 but less than −100
- (C) At least −100 but less than 100
- (D) At least 100 but less than 300
- (E) At least 300

18. For a one-period binomial model for the price of a stock with price 100 at time 0, you are given:

(i) The stock pays no dividends.

(ii) The stock price is either 110 or 95 at the end of the year.

(iii) The risk free force of interest is 5%.

Calculate the price at time 0 of a one-year call option with strike price 100.

- (A) Less than 6.00
- (B) At least 6.00 but less than 6.25
- (C) At least 6.25 but less than 6.50
- (D) At least 6.50 but less than 6.75
- (E) At least 6.75

19. Using the following information, determine the incurred losses for the 2017 accident year as reported at Dec. 31, 2018.

Occurrence #1: Occurrence date Feb. 1/16, Report date Apr. 1/16

Loss History:

Date	Total Paid to Date	Unpaid Loss Reserve	Total Incurred
Apr. 1/16	1000	1000	2000
Dec. 31/16	1500	1000	2500
Dec. 31/17	1500	1000	2500
Mar. 1/18	3000	0	3000

Occurrence #2: Occurrence date May 1/17, Report date July 1/17

Loss History:

Date	Total Paid to Date	Unpaid Loss Reserve	Total Incurred
July 1/17	1000	2000	3000
Dec. 31/17	3000	1000	4000
Dec. 31/18	5000	0	5000

Occurrence #3: Occurrence date Nov. 1/17, Report date Feb. 1/18

Loss History

Date	Total Paid to Date	Unpaid Loss Reserve	Total Incurred
Mar. 1/18	0	8000	8000
Dec. 31/18	5000	5000	10,000

(A) 0
(B) 3,000
(C) 5,000
(D) 8,000
(E) 15,000

20. You are given the following calendar year earned premium.

Year	CY2	CY3	CY4
Earned Premium	4200	4700	5000

You are also given the following rate changes

Date	April 1, CY1	September 1, CY2	July 1, CY3
Average Rate Change	+12 %	+6 %	+10 %

Determine the approximate earned premium at current (end of CY4) rates for CY3.

(A) Less than 5000
(B) At least 5000 but less than 5100
(C) At least 5100 but less than 5200
(D) At least 5200 but less than 5300
(E) At least 5300

Auto-Grade My Exam (Click Here)

1. A decision maker with utility function $u(w)$ will pay amount C to play a game (random variable) W based on the following relationship: $u(C) = E[u(W)]$.

   For individual X: $E[u_X(W)] = \int_0^{10{,}000} \frac{\sqrt{w}}{10{,}000}\, dw = \frac{200}{3} = u_X(C_X) = \sqrt{C_X}$.

   Solving for C_X results in $C_X = \left(\frac{200}{3}\right)^2 = 4{,}444.44$

   For individual Y: $E[u_Y(W)] = \int_0^{10{,}000} \frac{w}{10{,}000} \times \frac{1}{10{,}000}\, dw = 50 = u_Y(C_Y) = \frac{C_Y}{100}$.

   Solving for C_Y results in $C_Y = 5{,}000$

   For individual Z: $E[u_Z(W)] = \int_0^{10{,}000} \left(\frac{w}{10{,}000}\right)^2 \times \frac{1}{10{,}000}\, dw = \frac{100}{3} = u_Z(C_Z) = \left(\frac{C_Z}{100}\right)^2$.

   Solving for C_Z results in $C_Z = 1000 \times \sqrt{\frac{100}{3}} = 5{,}773.50$

   Only Z will pay more than 5,000 for the gamble. In general the concavity of the utility function determines the risk profile of the individual. X has a concave utility function $(u_X''(w) < 0)$ and will be risk-averse and will not be willing to pay more that the expected value (fair cost) of the gamble of 5,000. Y has a linear utility function and will be risk-neutral. Y will be willing to pay at most the expected payoff of the gamble of 5.000 but not more. Z has a convex utility function $(u_X''(w) > 0)$ and will be risk-preferring and will pay more the 5,000 for the gamble.

   Answer A

2. With coinsurance factor α, deductible d, policy limit $\alpha(u-d)$, the amount paid per loss is

   (we are assuming in inflation rate of $r = 0$)$Y = \begin{cases} 0 & X \le d \\ \alpha(X-d) & d < X \le u. \\ \alpha(u-d) & X > u \end{cases}$

   In this problem, the coinsurance factor is $\alpha = .8$, the deductible is $d = 5{,}000$, and the policy limit is $.8(u - 5{,}000) = 500{,}000$, so that the maximum covered loss is $u = 630{,}000$.

   The amount paid per loss becomes $Y = \begin{cases} 0 & X \le 5{,}000 \\ 0.80(X - 5{,}000) & 5{,}000 < X \le 630{,}000 \\ 500{,}000 & X > 630{,}000 \end{cases}$

   Answer C

3. When a payout occurs, it is 1, 2 or 3 with probability $\frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} = \frac{7}{8}$. The number of payouts that are 1,2 or 3 follows a Poisson process with an hourly rate of $5 \times \frac{7}{8} = \frac{35}{8}$.
   The expected number of payouts that are 1, 2 or 3 in 20 minutes, say N, has a Poisson distribution with mean $\frac{35}{8} \times \frac{20}{60} = \frac{35}{24}$. The probability that there are no payouts of 1, 2, or 3 in a given 20 minute period is the probability that $N = 0$, which is $e^{-35/24} = .233$.

   Answer D

ACTEX Learning

Advanced Topics in Predictive Analytics - Video Course

Johnny Li, PhD, FSA

Designed to help you more easily climb the steep learning curve of the new SOA ATPA Exam. In this course, leading Professor of Predictive Analytics Johnny Li will provide you with detailed explanations of the three models outlined in Topic 3 of the ATPA syllabus. The strong technical knowledge developed through this course will enable students to write the ATPA exam more confidently, and increase the likelihood of passing the exam.

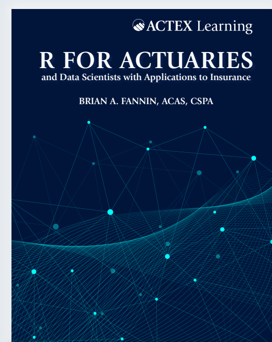
- **Lecture Slides** so you can follow along with the instructor and take notes.
- **Access to Instructor** during the duration of the course.
- **Access to Discussion Forum** to ask questions and see what other students are saying.
- **15+ Videos (6 Hours!)**, meticulously organized into three modules, helping you achieve the learning objectives of this course with ease.
- **3 End-of-Module Assessments** to evaluate your understanding of the material.
- **Certificate of Completion**

R for Actuaries and Data Scientists with Applications to Insurance

Brian Fannin, ACAS, CSPA

Written in a light, conversational style, this book will show you how to install and get up to speed with R in no time. It will also give you an overview of the key modeling techniques in modern data science including generalized linear models, decision trees, and random forests, and illustrates the use of these techniques with real datasets from insurance.

Engaging and at times funny, this book will be valuable for both newcomers to R and experienced practitioners who would like a better understanding of how R can be applied in insurance.



ACTEX Study Manual for SOA Exam PA + Videos | Spring 2022

Ambrose Lo, PhD, FSA, CERA

The ACTEX Study manual for Exam PA takes an integrated approach to learning Predictive Analytics with a three-component structure: A crash course in R covering the elements of R programming that are particularly germane to this exam, the best practices of all the statistical modeling techniques, and the written communication skills necessary for the exam.

Included with the manual and videos:

- **8 Practice Problems**
- **2 Full-Length Sample Projects**
- **91 Practice Exercises**
- **1800+ Pages of Full Syllabus Coverage**
- **All datasets, R scripts, & R markdown files** used are available virtually
- **Over 50 instructional videos** to go along with the study manual. In these videos, Professor Lo will walk you through the fundamental concepts in Predictive Analytics, with a strong emphasis on key test items in Exam PA. Videos are not sold separately.