Deeper Understanding, Faster Calculation
--Exam P Insights & Shortcuts

20th Edition

by Yufeng Guo

For SOA Exam P/CAS Exam 1

Fall, 2012 Edition

Table of Contents

Chapter 18 Exponential distribution

Key Points

Gain a deeper understanding of exponential distribution:

Why does exponential distribution model the time elapsed before the first or next random event occurs?

Exponential distribution lacks memory. What does this mean?

Understand and use the following integration shortcuts:

For any $\theta > 0$ and $a \geq 0$:

$$\int_a^{+\infty} \frac{1}{\theta} e^{-x/\theta} dx = e^{-a/\theta}$$

$$\int_a^{+\infty} x \left(\frac{1}{\theta} e^{-x/\theta} \right) dx = (a+\theta) e^{-a/\theta}$$

$$\int_a^{+\infty} x^2 \left(\frac{1}{\theta} e^{-x/\theta} \right) dx = \left[(a+\theta)^2 + \theta^2 \right] e^{-a/\theta}$$

You will need to understand and memorize these shortcuts to quickly solve integrations in the heat of the exam. Do not attempt to do integration by parts during the exam.

Explanations

Exponential distribution is the continuous version of geometric distribution. While geometric distribution describes the probability of having N trials before the first or next success (success is a random event), exponential distribution describes the probability of having time T elapse before the first or next success.

Let's use a simple example to derive the probability density function of exponential distribution. Claims for many insurance policies seem to occur randomly. Assume that on average, claims for a certain type of insurance policy occur once every 3 months. We want to find out the probability that T months will elapse before the next claim occurs.

To find the pdf of exponential distribution, we take advantage of geometric distribution, whose probability mass function we already know. We divide each month into n intervals, each interval being one minute. Since there are 30*24*60=43,200 minutes in a month (assuming there are 30 days in a month), we convert each month into 43,200 intervals. Because on average one claim occurs every 3 months, we assume that the chance of a claim occurring in one minute is

$$\frac{1}{3\times 43,200}$$

How many minutes must elapse before the next claim occurs? We can think of one minute as one trial. Then the probability of having *m* trials (i.e. *m* minutes) before the first success is a geometric distribution with

$$p=\frac{1}{3\times 43,200}$$

Instead of finding the probability that it takes **exactly** *m* minutes to have the first claim, we'll find the probability that it takes *m* **minutes or less** to have the first claim. The latter is the cumulative distribution function which is more useful.

P (it takes *m* minutes or less to have the first claim)
=1 – *P* (it takes more than *m* minutes to have the first claim)

The probability that it takes more than *m* trials before the first claim is $\left(1-p\right)^m$. To see why, you can reason that to have the first success only after m trials, the first *m* trials must all end with failures. The probability of having *m* failures in *m* trials is $\left(1-p\right)^m$.

Therefore, the probability that it takes *m* trials or less before the first success is $1-\left(1-p\right)^m$.

Now we are ready to find the pdf of T:

$P\left(T\le t\right)=P$ (43,200t trials or fewer before the 1st success)

$$=1-\left(1-\frac{1}{3\times 43,200}\right)^{43,200t}=1-\left[\left(1-\frac{1}{3\times 43,200}\right)^{-3\times 43,200t}\right]^{-t/3}\approx 1-e^{-t/3}$$

Of course, we do not need to limit ourselves by dividing one month into intervals of one minute. We can divide, for example, one month into *n* intervals, with each interval of 1/1,000,000 of a minute. Essentially, we want $n\to+\infty$.

$P\left(T\le t\right)=P$ (nt trials or fewer before the 1st success)

$$=1-\left(1-\frac{1}{3n}\right)^{nt}=1-\left[\left(1-\frac{1}{3n}\right)^{-3n}\right]^{-t/3}=1-e^{-t/3} \quad (\text{as } n\to+\infty)$$

If you understand the above, you should have no trouble understanding why exponential distribution is often used to model time elapsed until the next random event happens.

Here are some examples where exponential distribution can be used:

- Time until the next claim arrives in the claims office.
- Time until you have your next car accident.
- Time until the next customer arrives at a supermarket.
- Time until the next phone call arrives at the switchboard.

General formula:

Let T=time elapsed (in years, months, days, etc.) before the next random event occurs.

$$F(t)=P(T\le t)=1-e^{-t/\theta},\ f(t)=\frac{1}{\theta}e^{-t/\theta}, \text{ where } \theta=E(T)$$
$$P(T>t)=1-F(t)=e^{-t/\theta}$$

Alternatively,

$$F(t)=P(T\le t)=1-e^{-\lambda t},\ f(t)=\lambda e^{-\lambda t}, \text{ where } \lambda=\frac{1}{E(T)}$$
$$P(T>t)=1-F(t)=e^{-\lambda t}$$

Mean and variance: $E(T)=\theta=\frac{1}{\lambda}$, $Var(T)=\theta^2=\frac{1}{\lambda^2}$

Like geometric distribution, exponential distribution lacks memory:
$$P(T>a+b|T>a)=P(T>b)$$

We can easily derive this:

$$P(T>a+b|T>a)=\frac{P(T>a+b\cap T>a)}{P(T>a)}=\frac{P(T>a+b)}{P(T>a)}=\frac{e^{-(a+b)/\theta}}{e^{-a/\theta}}=e^{-b/\theta}=P(T>b)$$

In plain English, this lack of memory means that if a component's time to failure follows exponential distribution, then the component does not remember how long it has been working (i.e. does not remember wear and tear). At any moment when it is working, the component starts fresh as if it were completely new. At any moment while the component is working, if you reset the clock to zero and count the time elapsed from then until the component breaks down, the time elapsed before a breakdown is always exponentially distributed with the identical mean.

This is clearly an idealized situation, for in real life wear and tear does reduce the longevity of a component. However, in many real world situations, exponential distribution can be used to approximate the actual distribution of time until failure and still give a reasonably accurate result.

A simple way to see why a component can, at least by theory, forget how long it has worked so far is to think about geometric distribution, the discrete counterpart of exponential distribution. For example, in tossing a coin, you can clearly see why a coin doesn't remember its past success history. Since getting heads or tails is a purely random event, how many times you have tossed a coin so far before getting heads really should NOT have any bearing on how many more times you need to toss the coin to get heads the second time.

The calculation shortcuts are explained in the following sample problems.

Sample Problems and Solutions

Problem 1

The lifetime of a light bulb follows exponential distribution with a mean of 100 days. Find the probability that the light bulb's life …

(1) Exceeds 100 days
(2) Exceeds 400 days
(3) Exceeds 400 days given it exceeds 100 days.

Solution

Let T = # of days before the light bulb burns out.

$F(t) = 1 - e^{-t/\theta}$, where $\theta = E(T) = 100$
$P(T > t) = 1 - F(t) = e^{-t/\theta}$

$$P(T > 100) = 1 - F(100) = e^{-100/100} = e^{-1} = 0.3679$$
$$P(T > 400) = 1 - F(400) = e^{-400/100} = e^{-4} = 0.0183$$

$$P(T > 400 \mid T > 100) = \frac{P(T > 400)}{P(T > 100)} = \frac{e^{-400/100}}{e^{-100/100}} = e^{-3} = 0.0498$$

Or use the lack of memory property of exponential distribution:

$$P\left(T>400 \middle| T>100\right)=P\left(T>400-100\right)=e^{-300/100}=e^{-3}=0.0498$$

Problem 2

The length of telephone conversations follows exponential distribution. If the average telephone conversation is 2.5 minutes, what is the probability that a telephone conversation lasts between 3 minutes and 5 minutes?

Solution

$$F\left(t\right)=1-e^{-t/2.5}$$
$$P\left(3<T<5\right)=\left(1-e^{-5/2.5}\right)-\left(1-e^{-3/2.5}\right)=e^{-3/2.5}-e^{-5/2.5}=0.1659$$

Problem 3

The random variable T has an exponential distribution with pdf $f\left(t\right)=\frac{1}{2}e^{-t/2}$.

Find $E\left(T \middle| T>3\right)$, $Var\left(T \middle| T>3\right)$, $E\left(T \middle| T\leq 3\right)$, $Var\left(T \middle| T\leq 3\right)$.

Solution

First, let's understand the conceptual thinking behind the symbol $E\left(T \middle| T>3\right)$. Here we are only interested in $T>3$. So we reduce the original sample space $T\in\left[0,+\infty\right]$ to $T\in\left[3,+\infty\right]$. The pdf in the original sample space $T\in\left[0,+\infty\right]$ is $f\left(t\right)$; the pdf in the reduced sample space $t\in\left[3,+\infty\right]$ is $\frac{f\left(t\right)}{P\left(T>3\right)}$. Here the factor $\frac{1}{P\left(T>3\right)}$ is a normalizing constant to make the total probability in the reduced sample space add up to one:

$$\int_3^{+\infty}\frac{f\left(t\right)}{P\left(T>3\right)}dt=\frac{1}{P\left(T>3\right)}\int_3^{+\infty}f\left(t\right)dt=\frac{1}{P\left(T>3\right)}\times P\left(T>3\right)=1$$

Next, we need to calculate $E\left(T \mid T>3\right)$, the expected value of T in the reduced sample space $T \in [3,+\infty]$:

$$E\left(T \mid T>3\right)=\int\limits_{\substack{\text{Reduced} \\ \text{Sample space} \\ T\in[3,+\infty]}} t\frac{f(t)}{P(T>3)}dt=\frac{1}{P(T>3)}\int_3^{+\infty} t\, f(t)dt=\frac{1}{1-F(3)}\int_3^{+\infty} t\, f(t)dt$$

$$1-F(3)=e^{-3/2}$$

$$\int_3^{+\infty} tf(t)dt=5e^{-3/2} \quad \text{(integration by parts)}$$

$$E\left(T \mid T>3\right)=\frac{1}{1-F(3)}\int_3^{+\infty} tf(t)dt=\frac{5e^{-3/2}}{e^{-3/2}}=5$$

Here is another approach. Because T does not remember wear and tear and always starts anew at any working moment, the time elapsed since T =3 until the next random event (i.e. $T-3$) has exponential distribution with an identical mean of 2. In other words, $\left(T-3 \mid T>3\right)$ is exponentially distributed with an identical mean of 2.

So $E\left(T-3 \mid T>3\right)=2$.

$$E\left(T \mid T>3\right)=E\left(T-3 \mid T>3\right)+3=2+3=5$$

Next, we will find $Var\left(T \mid T>3\right)$.

$$E\left(T^2 \mid T>3\right)=\frac{1}{\Pr(T>3)}\int_3^{+\infty} t^2 f(t)dt=\frac{1}{\Pr(T>3)}\int_3^{+\infty} t^2\frac{1}{2}e^{-t/2}dt$$

$$\int_3^{+\infty} t^2\frac{1}{2}e^{-t/2}dt=29e^{-3/2} \quad \text{(integration by parts)}$$

$$E\left(T^2 \mid T>3\right)=\frac{29e^{-3/2}}{e^{-3/2}}=29$$

$$Var\left(T \mid T>3\right)=E\left(T^2 \mid T>3\right)-E^2\left(T \mid T>3\right)=29-5^2=4=\theta^2$$

It is no coincidence that $Var(T|T>3)$ is the same as $Var(T)$. To see why, we know $Var(T|T>3)=Var(T-3|T>3)$. This is because $Var(X+c)=Var(X)$ stands for any constant c.

Since $(T-3|T>3)$ is exponentially distributed with an identical mean of 2, then

$$Var(T-3|T>3)=\theta^2=2^2=4.$$

Next, we need to find $E(T|T\leq 3)$.

$$E(T|T<3)=\int_0^3 t\frac{f(t)}{\Pr(T<3)}dt=\frac{1}{F(3)}\int_0^3 tf(t)dt$$

$$F(3)=1-e^{-3/2}$$

$$\int_0^3 tf(t)dt=\int_0^{+\infty} tf(t)dt-\int_3^{+\infty} tf(t)dt$$

$$\int_0^{+\infty} tf(t)dt=E(T)=2$$

$$\int_3^{+\infty} tf(t)dt=5e^{-3/2}$$ (we already calculated this)

$$E(T|T<3)=\frac{1}{F(3)}\int_0^3 tf(t)dt=\frac{2-5e^{-3/2}}{1-e^{-3/2}}$$

Here is another way to find $E(T|T<3)$.

$$E(T)=E(T|T<3)\times P(T<3)+E(T|T>3)\times P(T>3)$$

The above equation says that if we break down T into two groups, T >3 and T <3, then the overall mean of these two groups as a whole is equal to the weighted average mean of these groups.

Also note that $P(T=3)$ is not included in the right-hand side because the probability of a continuous random variable at any single point is zero. This is similar to the concept that the mass of a single point is zero.

Of course, you can also write:

$$E(T) = E(T|T \le 3) \times P(T \le 3) + E(T|T > 3) \times P(T > 3)$$

Or $E(T) = E(T|T < 3) \times P(T < 3) + E(T|T \ge 3) \times P(T \ge 3)$

You should get the same result no matter which formula you use.

$$E(T) = E(T|T < 3) \times P(T < 3) + E(T|T > 3) \times P(T > 3)$$

$$\Rightarrow \quad E(T|T < 3) = \frac{E(T) - E(T|T > 3) \times P(T > 3)}{P(T < 3)}$$

$$\Rightarrow \quad E(T|T < 3) = \frac{\theta - (\theta + 3)e^{-3/2}}{1 - e^{-3/2}} = \frac{2 - 5e^{-3/2}}{1 - e^{-3/2}}$$

Next, we will find $E(T^2|T < 3)$:

$$E(T^2|T < 3) = \frac{1}{P(T < 3)} \int_0^3 t^2 f(t) dt = \frac{1}{P(T < 3)} \int_0^3 t^2 \frac{1}{2} e^{-t/2} dt$$

$$\int_0^3 t^2 \frac{1}{2} e^{-t/2} dt = -\left[(t+2)^2 + 4\right] e^{-x/2} \Big|_0^3 = 8 - 29e^{-3/2}$$

$$E(T^2|T < 3) = \frac{1}{P(T > 3)} \int_0^3 t^2 \frac{1}{2} e^{-t/2} dt = \frac{8 - 29e^{-3/2}}{1 - e^{-3/2}}$$

Alternatively,

$$E(T^2) = E(T^2|T < 3) \times P(T < 3) + E(T^2|T > 3) \times P(T > 3)$$

$$\Rightarrow \quad E(T^2|T < 3) = \frac{E(T^2) - E(T^2|T > 3) \times P(T > 3)}{P(T < 3)}$$

$$= \frac{2\theta^2 - 29 \times P(T > 3)}{P(T < 3)} = \frac{8 - 29e^{-3/2}}{1 - e^{-3/2}}$$

$$Var\left(T\middle|T<3\right)=E\left(T^2\middle|T<3\right)-E^2\left(T\middle|T<3\right)=\frac{8-29e^{-3/2}}{1-e^{-3/2}}-\left(\frac{2-5e^{-3/2}}{1-e^{-3/2}}\right)^2$$

In general, for any exponentially distributed random variable T with mean $\theta>0$ and for any $a\geq 0$:

$T-a\middle|T>a$ is also exponentially distributed with mean θ

$$\Rightarrow \quad E\left(T-a\middle|T>a\right)=\theta, \quad Var\left(T-a\middle|T>a\right)=\theta^2$$

$$\Rightarrow \quad E\left(T\middle|T>a\right)=a+\theta, \quad Var\left(T\middle|T>a\right)=\theta^2$$

$$E\left(T-a\middle|T>a\right)=\frac{1}{P(T>a)}\int_a^{+\infty}(t-a)f(t)dt$$

$$\Rightarrow \quad \int_a^{+\infty}(t-a)f(t)dt=E\left(T-a\middle|T>a\right)\times P(T>a)=\theta e^{-a/\theta}$$

$$E\left(T\middle|T>a\right)=\frac{1}{P(T>a)}\int_a^{+\infty}tf(t)dt$$

$$\Rightarrow \quad \int_a^{+\infty}tf(t)dt=E\left(T\middle|T>a\right)\times P(T>a)=(\theta+a)e^{-a/\theta}$$

$$E\left(T\middle|T<a\right)=\frac{1}{P(T<a)}\int_0^{a}tf(t)dt$$

$$\Rightarrow \quad \int_0^{a}tf(t)dt=E\left(T\middle|T<a\right)\times P(T<a)=E\left(T\middle|T<a\right)\left(1-e^{-a/\theta}\right)$$

$$E(T)=E\left(T\middle|T<a\right)\times P(T<a)+E\left(T\middle|T>a\right)\times P(T>a)$$

$$\Rightarrow \quad \theta=E\left(T\middle|T<a\right)\times(1-e^{-a/\theta})+E\left(T\middle|T>a\right)\times e^{-a/\theta}$$

$$E\left(T^2\right)=E\left(T^2\middle|T<a\right)\times P(T<a)+E\left(T^2\middle|T>a\right)\times P(T>a)$$

You do not need to memorize the above formulas. However, make sure you understand the logic behind these formulas.

Before we move on to more sample problems, I will give you some integration-by-parts formulas for you to memorize. These formulas are critical to you when solving exponential distribution-related problems in 3 minutes. You should memorize these formulas to avoid doing integration by parts during the exam.

Formulas you need to memorize:

For any $\theta > 0$ and $a \geq 0$

$$\int_a^{+\infty} \frac{1}{\theta} e^{-x/\theta} dx = e^{-a/\theta} \qquad (1)$$

$$\int_a^{+\infty} x\left(\frac{1}{\theta} e^{-x/\theta}\right) dx = (a+\theta) e^{-a/\theta} \qquad (2)$$

$$\int_a^{+\infty} x^2\left(\frac{1}{\theta} e^{-x/\theta}\right) dx = \left[(a+\theta)^2 + \theta^2\right] e^{-a/\theta} \qquad (3)$$

You can always prove the above formulas using integration by parts. However, let me give an intuitive explanation to help you memorize them.

Let X represent an exponentially random variable with a mean of θ, and $f(x)$ is the probability distribution function, then for any $a \geq 0$, Equation (1) represents $P(X > a) = 1 - F(a)$, where $F(x) = 1 - e^{-x/\theta}$ is the cumulative distribution function of X . You should have no trouble memorizing Equation (1).

For Equation (2), from Sample Problem 3, we know

$$\int_a^{+\infty} x f(x) dx = E(X \mid X > a) \times P(X > a) = (a+\theta) e^{-a/\theta}$$

To understand Equation (3), note that

$$P(X > a) = e^{-a/\theta}$$

$$\int_a^{+\infty} x^2 f(x) dx = E(X^2 \mid X > a) \times P(X > a)$$

$$E(X^2 \mid X > a) = E^2(X \mid X > a) + Var(X \mid X > a)$$

$$E^2(X \mid X > a) = (a+\theta)^2, \quad Var(X \mid X > a) = \theta^2$$

Then

$$\int_a^{+\infty} x^2 f(x) dx = \left[(a+\theta)^2 + \theta^2\right] e^{-a/\theta}$$

We can modify Equation (1),(2),(3) into the following equations:

For any $\theta > 0$ and $b \ge a \ge 0$

$$\int_a^b \frac{1}{\theta} e^{-x/\theta} dx = e^{-a/\theta} - e^{-b/\theta} \qquad (4)$$

$$\int_a^b x\left(\frac{1}{\theta} e^{-x/\theta}\right) dx = (a+\theta) e^{-a/\theta} - (b+\theta) e^{-b/\theta} \qquad (5)$$

$$\int_a^b x^2\left(\frac{1}{\theta} e^{-x/\theta}\right) dx = \left[(a+\theta)^2 + \theta^2\right] e^{-a/\theta} - \left[(b+\theta)^2 + \theta^2\right] e^{-b/\theta} \qquad (6)$$

We can easily prove the above equation. For example, for Equation (5):

$$\int_a^b x\left(\frac{1}{\theta} e^{-x/\theta}\right) dx = \int_a^{+\infty} x\left(\frac{1}{\theta} e^{-x/\theta}\right) dx - \int_b^{+\infty} x\left(\frac{1}{\theta} e^{-x/\theta}\right) dx = (a+\theta) e^{-a/\theta} - (b+\theta) e^{-b/\theta}$$

We can modify Equation (1),(2),(3) into the following equations:

For any $\theta > 0$ and $a \ge 0$

$$\int \frac{1}{\theta} e^{-x/\theta} dx = -e^{-x/\theta} + c \qquad (7)$$

$$\int x\left(\frac{1}{\theta} e^{-x/\theta}\right) dx = -(x+\theta) e^{-x/\theta} + c \qquad (8)$$

$$\int x^2\left(\frac{1}{\theta} e^{-x/\theta}\right) dx = -\left[(x+\theta)^2 + \theta^2\right] e^{-x/\theta} + c \qquad (9)$$

Set $\lambda = \frac{1}{\theta}$. For any $\lambda > 0$ and $a \ge 0$

$$\int \lambda e^{-\lambda x} dx = -e^{-\lambda x} + c \qquad (10)$$

$$\int x\left(\lambda e^{-\lambda x}\right) dx = -\left(x + \frac{1}{\lambda}\right) e^{-\lambda x} + c \qquad (11)$$

$$\int x^2\left(\lambda e^{-\lambda x}\right) dx = -\left[\left(x + \frac{1}{\lambda}\right)^2 + \frac{1}{\lambda^2}\right] e^{-\lambda x} + c \qquad (12)$$

So you have four sets of formulas. Just remember one set (any one is fine). Equations (4),(5),(6) are most useful (because you can directly apply the formulas), but the formulas are long.

If you can memorize any one set, you can avoid doing integration by parts during the exam. You definitely do not want to calculate messy integrations from scratch during the exam.

Now we are ready to tackle more problems.

Problem 4

After an old machine was installed in a factory, Worker John is on call 24-hours a day to repair the machine if it breaks down. If the machine breaks down, John will receive a service call right away, in which case he immediately arrives at the factory and starts repairing the machine.

The machine's time to failure is exponentially distributed with a mean of 3 hours. Let T represent the time elapsed between when the machine was first installed and when John starts repairing the machine.

Find $E(T)$ and $Var(T)$.

Solution

T is exponentially distributed with mean $\theta = 3$. $F(t) = 1 - e^{-t/3}$.
We simply apply the mean and variance formula:

$$E(T) = \theta = 3, \quad Var(T) = \theta^2 = 3^2 = 9$$

Problem 5

After an old machine was installed in a factory, Worker John is on call 24-hours a day to repair the machine if it breaks down. If the machine breaks down, John will receive a service call right away, in which case he immediately arrives at the factory and starts repairing the machine.

The machine was found to be working today at 10:00 a.m..

The machine's time to failure is exponentially distributed with a mean of 3 hours. Let T represent the time elapsed between 10:00 a.m. and when John starts repairing the machine.

Find $E(T)$ and $Var(T)$.

Solution

Exponential distribution lacks memory. At any moment when the machine is working, it forgets its past wear and tear and starts afresh. If we reset the clock at 10:00 and observe T, the time elapsed until a breakdown, T is exponentially distributed with a mean of 3.

$E(T)=\theta=3$, $Var(T)=\theta^2=3^2=9$

Problem 6

After an old machine was installed in a factory, Worker John is on call 24-hours a day to repair the machine if it breaks down. If the machine breaks down, John will receive a service call right away, in which case he immediately arrives at the factory and starts repairing the machine.

Today, John happens to have an appointment from 10:00 a.m. to 12:00 noon. During the appointment, he won't be able to repair the machine if it breaks down.

The machine was found working today at 10:00 a.m..

The machine's time to failure is exponentially distributed with a mean of 3 hours. Let X represent the time elapsed between 10:00 a.m. today and when John starts repairing the machine.

Find $E(T)$ and $Var(T)$.

Solution

Let T =time elapsed between 10:00 a.m. today and a breakdown. T is exponentially distributed with a mean of 3. $X=\max(2,T)$.

$$X=\begin{cases}2, & \text{if } T\le 2\\ T, & \text{if } T>2\end{cases}$$

You can also write

$$X=\begin{cases}2, & \text{if } T<2\\ T, & \text{if } T\ge 2\end{cases}$$

As said before, it doesn't matter where you include the point T =2 because the probability density function of a continuous variable at any single point is always zero.

Pdf is always $f\left(t\right)=\dfrac{1}{3}e^{-t/3}$ no matter $T\le 2$ or $T>2$.

$$E\left(X\right)=\int_0^{+\infty}x\left(t\right)f\left(t\right)dt=\int_0^2 2\left(\frac{1}{3}e^{-t/3}\right)dt+\int_2^{+\infty}t\left(\frac{1}{3}e^{-t/3}\right)dt$$

$$\int_0^2 2\left(\frac{1}{3}e^{-t/3}\right)dt=2\left(1-e^{-2/3}\right)$$

$$\int_2^{+\infty}t\left(\frac{1}{3}e^{-t/3}\right)dt=\left(2+3\right)e^{-2/3}=5e^{-2/3}$$

$$E\left(X\right)=2\left(1-e^{-2/3}\right)+5e^{-2/3}=2+3e^{-2/3}$$

$$E\left(X^2\right)=\int_0^{+\infty}x^2 f\left(t\right)dt=\int_0^2 2^2\left(\frac{1}{3}e^{-t/3}\right)dt+\int_2^{+\infty}t^2\left(\frac{1}{3}e^{-t/3}\right)dt$$

$$\int_0^2 2^2\left(\frac{1}{3}e^{-t/3}\right)dt=2^2\left(1-e^{-2/3}\right)=4\left(1-e^{-2/3}\right)$$

$$\int_2^{+\infty}t^2\left(\frac{1}{3}e^{-t/3}\right)dt=\left(5^2+3^2\right)e^{-2/3}=34e^{-2/3}$$

$$E\left(X^2\right)=4\left(1-e^{-2/3}\right)+34e^{-2/3}=4+30e^{-2/3}$$

$$Var\left(X\right)=E\left(X^2\right)-E^2\left(X\right)=4+3\;e0^{2/3}-\left(2+3e^{-2/3}\right)^2$$

We can quickly check that $E(X)=2+3e^{-2/3}$ is correct:

$$X=\begin{cases}2, & \text{if } T\le 2\\ T, & \text{if } T>2\end{cases}\quad\Rightarrow\quad X-2=\begin{cases}0, & \text{if } T\le 2\\ T-2, & \text{if } T>2\end{cases}$$

$$\Rightarrow\quad E\left(X-2\right)=0\times E\left(T\middle|T<2\right)\times P\left(T<2\right)+E\left(T-2\middle|T>2\right)\times P\left(T>2\right)$$
$$=E\left(T-2\middle|T>2\right)\times P\left(T>2\right)=3e^{-2/3}$$

$$\Rightarrow\quad E\left(X\right)=E\left(X-2\right)+2=2+3e^{-2/3}$$

You can use this approach to find $E\left(X^2\right)$ too, but this approach isn't any quicker than using the integration as we did above

Problem 7

After an old machine was installed in a factory, Worker John is on call 24-hours a day to repair the machine if it breaks down. If the machine breaks down, John will receive a service call right away, in which case he immediately arrives at the factory and starts repairing the machine.

Today is John's last day of work because he got an offer from another company, but he'll continue his current job of repairing the machine until 12:00 noon if there's a breakdown. However, if the machine does not break by noon 12:00, John will have a final check of the machine at 12:00. After 12:00 noon John will permanently leave his current job and take a new job at another company.

The machine was found working today at 10:00 a.m..

The machine's time to failure is exponentially distributed with a mean of 3 hours. Let X represent the time elapsed between 10:00 a.m. today and John's visit to the machine.

Find $E(X)$ and $Var(X)$.

Solution

Let T =time elapsed between 10:00 a.m. today and a breakdown. T is exponentially distributed with a mean of 3. $X = \min(2, T)$.

$$X = \begin{cases} t, & \text{if } T \le 2 \\ 2, & \text{if } T > 2 \end{cases}$$

Pdf is always $f(t) = \frac{1}{3} e^{-t/3}$ no matter $T \le 2$ or $T > 2$.

$$E(X) = \int_0^2 t \frac{1}{3} e^{-t/3} dt + \int_2^{+\infty} 2\left(\frac{1}{3} e^{-t/3}\right) dt$$

$$\int_0^2 t \frac{1}{3} e^{-t/3} dt = 3 - (2+3) e^{-2/3}$$

$$\int_2^{+\infty} 2\left(\frac{1}{3} e^{-t/3}\right) dt = 2e^{-2/3}$$

$$E(X) = 3 - 5e^{-2/3} + 2e^{-2/3} = 3 - 3e^{-2/3}$$

To find $Var(X)$, we need to calculate $E(X^2)$.

$$E(X^2) = \int_0^{+\infty} x^2(t)\, f(t)\, dt = \int_0^2 t^2 f(t)\, dt + \int_2^{+\infty} 2^2 f(t)\, dt$$

$$\int_0^2 t^2 f(t)\, dt = \left[(0+3)^2 + 3^2\right] e^{-0/3} - \left[(2+3)^2 + 3^2\right] e^{-2/3} = 18 - 34e^{-2/3}$$

$$\int_2^{+\infty} 2^2 f(t)\, dt = 4e^{-2/3}$$

$$E(X^2) = 18 - 34e^{-2/3} + 4e^{-2/3} = 18 - 30e^{-2/3}$$

$$Var(X) = E(X^2) - E^2(X) = \left(18 - 30e^{-2/3}\right) - \left(3 - 3e^{-2/3}\right)^2$$

We can easily verify that $E(X) = 3 - 3e^{-2/3}$ is correct. Notice:

$$T + 2 = \min(T, 2) + \max(T, 2)$$

$$\Rightarrow \quad E(T+2) = E\left[\min(T,2)\right] + E\left[\max(T,2)\right]$$

We know that

$$E\left[\min(T,2)\right] = 3 - 3e^{-2/3} \quad \text{(from this problem)}$$

$$E\left[\max(T,2)\right] = 2 + 3e^{-2/3} \quad \text{(from the previous problem)}$$

$$E(T+2) = E(T) + 2 = 3 + 2$$

So the equation $E(T+2) = E\left[\min(T,2)\right] + E\left[\max(T,2)\right]$ holds.

We can also check that $E(X^2) = 18 - 30e^{-2/3}$ is correct.

$$T + 2 = \min(T, 2) + \max(T, 2)$$

$$\Rightarrow \qquad (T+2)^2 = \left[\min(T,2) + \max(T,2)\right]^2$$

$$= \left[\min(T,2)\right]^2 + \left[\max(T,2)\right]^2 + 2\min(T,2)\max(T,2)$$

$$\min(T,2)=\begin{cases} t & \text{if } t\le 2 \\ 2 & \text{if } t>2 \end{cases}, \qquad \max(T,2)=\begin{cases} 2 & \text{if } t\le 2 \\ t & \text{if } t>2 \end{cases}$$

$$\Rightarrow \qquad \min(T,2)\max(T,2)=2t$$

$$\Rightarrow \qquad (T+2)^2=\left[\min(T,2)\right]^2+\left[\max(T,2)\right]^2+2(2t)$$

$$\Rightarrow \qquad E(T+2)^2=E\left[\min(T,2)+\max(T,2)\right]^2$$

$$=E\left[\min(T,2)\right]^2+E\left[\max(T,2)\right]^2+E\left[2(2t)\right]$$

$$E(T+2)^2=E\left(T^2+4t+4\right)=E\left(T^2\right)+4E(t)+4=2\left(3^2\right)+4(3)+4=34$$

$E\left[\min(T,2)\right]^2=18-30e^{-2/3}$ (from this problem)

$E\left[\max(T,2)\right]^2=4+30e^{-2/3}$ (from previous problem)

$E\left[2(2t)\right]=4E(t)=4(3)=12$

$$E\left[\min(T,2)\right]^2+E\left[\max(T,2)\right]^2+E\left[2(2t)\right]$$
$$=18-30e^{-2/3}+4+30e^{-2/3}+12=34$$

So the equation $E(T+2)^2=E\left[\min(T,2)+\max(T,2)\right]^2$ holds.

Problem 8

An insurance company sells an auto insurance policy that covers losses incurred by a policyholder, subject to a deductible of \$100 and a maximum payment of \$300. Losses incurred by the policyholder are exponentially distributed with a mean of \$200. Find the expected payment made by the insurance company to the policyholder.

Solution

Let X =losses incurred by the policyholder. X is exponentially distributed with a mean of 200, $f(x)=\dfrac{1}{200}e^{-x/200}$.

Let Y =claim payment by the insurance company.

$$Y=\begin{cases} 0, & \text{if } X\le 100 \\ X-100, & \text{if } 100\le X\le 400 \\ 300, & \text{if } X\ge 400 \end{cases}$$

$$E(Y)=\int_0^{+\infty} y(x)\, f(x)\,dx=\int_0^{100} 0\, f(x)\,dx+\int_{100}^{400}(x-100)\, f(x)\,dx+\int_{400}^{+\infty} 300\, f(x)\,dx$$

$$\int_0^{100} 0\, f(x)\,dx=0$$

$$\int_{100}^{400}(x-100)\, f(x)\,dx=\int_{100}^{400} xf(x)\,dx-\int_{100}^{400} 100\, f(x)\,dx$$

$$\int_{100}^{400} xf(x)\,dx=(100+200)e^{-100/200}-(400+200)e^{-400/200}=300e^{-1/2}-600e^{-2}$$

$$\int_{100}^{400} 100\, f(x)\,dx=100\left(e^{-100/200}-e^{-400/200}\right)=100\left(^{-1/2}-e^{-2}\right)$$

$$\int_{400}^{+\infty} 300\, f(x)\,dx=300e^{-400/200}=300e^{-2}$$

Then we have

$$E(X)=300e^{-1/2}-600e^{-2}-100\left(e^{-1/2}-e^{-2}\right)+300e^{-2}=200\left(e^{-1/2}-e^{-2}\right)$$

Alternatively, we can use the shortcut developed in Chapter 22:

$$E(X)=\int_d^{d+L}\Pr(X>x)\,dx=\int_{100}^{100+300} e^{-x/200}dx=200\left[e^{-x/200}\right]_{400}^{100}=200\left(e^{-1/2}-e^{-2}\right)$$

Problem 9

An insurance policy has a deductible of 3. Losses are exponentially distributed with mean 10. Find the expected non-zero payment by the insurer.

Solution

Let X represent the losses and Y the payment by the insurer. Then $Y=0$ if $X\le 3$; $Y=X-3$ if $X>3$. We are asked to find $E(Y|Y>0)$.

$$E(Y|Y>0)=E(X-3|X>3)$$

$X-3|X>3$ is an exponential random variable with the identical mean of 10. So $E(X-3|X>3)=E(X)=10$.

Generally, if X is an exponential loss random variable with mean θ, then for any positive deductible d

$$E(X-d \mid X>d)=E(X)=\theta, \quad E(X \mid X>d)=E(X-d \mid X>d)+d=\theta+d$$

Problem 10

Claims are exponentially distributed with a mean of $8,000. Any claim exceeding $30,000 is classified as a big claim. Any claim exceeding $60,000 is classified as a super claim.

Find the expected size of big claims and the expected size of super claims.

Solution

This problem tests your understanding that the exponential distribution lacks memory.
Let X represents claims. X is exponentially distributed with a mean of θ=8,000.
Let Y =big claims, Z =super claims.

$$\begin{aligned} E(Y) &= E(X \mid X>30,000)=E(X-30,000 \mid X>30,000)+30,000 \\ &= E(X)+30,000=\theta+30,000=38,000 \end{aligned}$$

$$\begin{aligned} E(Z) &= E(X \mid X>60,000)=E(X-60,000 \mid X>60,000)+60,000 \\ &= E(X)+60,000=\theta+60,000=68,000 \end{aligned}$$

Problem 11

Evaluate $\int_2^{+\infty}(x^2+x)e^{-x/5}dx$.

Solution

$$\int_2^{+\infty}(x^2+x)e^{-x/5}dx=5\int_2^{+\infty}(x^2+x)\left(\frac{1}{5}e^{-x/5}\right)dx=5\int_2^{+\infty}x^2\left(\frac{1}{5}e^{-x/5}\right)dx+5\int_2^{+\infty}x\left(\frac{1}{5}e^{-x/5}\right)dx$$

$$\int_2^{+\infty}x^2\left(\frac{1}{5}e^{-x/5}\right)dx=\left[5^2+(5+2)^2\right]e^{-2/5}, \quad \int_2^{+\infty}x\left(\frac{1}{5}e^{-x/5}\right)dx=(5+2)e^{-2/5}$$

$$\Rightarrow \quad \int_2^{+\infty}(x^2+x)e^{-x/5}dx=5\left[5^2+((5+2)^2+(5+2)\right]e^{-2/5}=405e^{-2/5}$$

Problem 12 (two exponential distributions competing)

You have a car and a van. The time-to-failure of the car and the time-to-failure the van are two independent exponential random variables with mean of 8 years and 4 years respectively.

Calculate the probability that the car dies before the van.

Solution

Let X and Y represent the time-to-failure of the car and the time-to-failure of the van respectively. We are asked to find $P(X < Y)$.

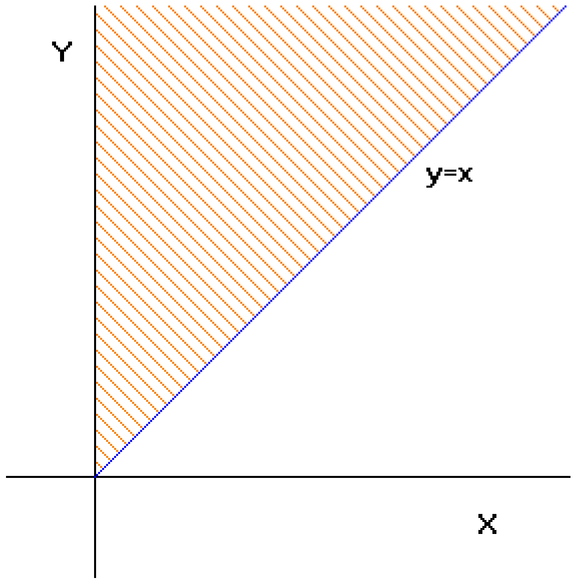
X and Y are independent exponential random variables with mean of 8 and 4 respectively. Their pdf is:

$$f_X(x) = \frac{1}{8}e^{-x/8}, \quad F_X(x) = 1 - e^{-x/8}, \text{ where } x \geq 0$$

$$f_Y(y) = \frac{1}{4}e^{-y/4}, \quad F_Y(y) = 1 - e^{-y/4}, \text{ where } y \geq 0$$

Method #1 X and Y have the following joint pdf:

$$f_{X,Y}(x,y) = f_X(x)\, f_Y(y) = \frac{1}{8}e^{-x/8}\left(\frac{1}{4}e^{-y/4}\right)$$



The shaded area is $x \ge 0$, $y \ge 0$, and $x < y$.

$$P(X < Y) = \iint_{shaded\ Area} f(x,y)\,dxdy = \int_0^{+\infty}\int_x^{+\infty} f(x,y)\,dydx = \int_0^{+\infty}\int_x^{+\infty} \frac{1}{8}e^{-x/8}\left(\frac{1}{4}e^{-y/4}\right)dydx$$

$$= \int_0^{+\infty} \frac{1}{8}e^{-x/8}\left(e^{-x/4}\right)dx = \int_0^{+\infty} \frac{1}{8}e^{-3x/8}dx = \frac{1}{3}$$

Method 2

$$P(X < Y) = \int_0^{+\infty} P(x < X \le x + dx)P(Y > x + dx)$$

The above equation says that to that find $P(X < Y)$, we first fix X at a tiny interval $(x, x+dx]$. Next, we set $Y > x + dx$. This way, we are guaranteed that $X < Y$ when X falls in the interval $(x, x+dx]$. To find $P(X < Y)$ when X falls $[0, +\infty]$, we simply integrate $P(x < X \le x + dx)P(Y > x + dx)$ over the interval $[0, +\infty]$.

$$P(x < X \le x + dx) = f(x)\,dx = \frac{1}{8}e^{-x/8}dx$$

$$P(Y > x + dx) = P(Y > x) \qquad \text{because } dx \text{ is tiny}$$

$$= 1 - F_Y(x) = 1 - \left(1 - e^{-x/4}\right) = e^{-x/4}$$

$$P(X<Y)=\int_0^{+\infty} f_X(x)P(Y>x)=\int_0^{+\infty}\left(\frac{1}{8}e^{-x/8}dx\right)e^{-x/4}=\frac{1}{8}\int_0^{+\infty} e^{-\left(\frac{1}{8}+\frac{1}{4}\right)x}dx=\frac{\frac{1}{8}}{\frac{1}{8}+\frac{1}{4}}=\frac{1}{3}$$

This is the intuitive meaning behind the formula $\frac{\frac{1}{8}}{\frac{1}{8}+\frac{1}{4}}$. In this problem, we have a car and a van. The time-to-failure of the car and the time-to-failure the van are two independent exponential random variables with mean of 8 years and 4 years respectively. So on average a car failure arrives at the speed of 1/8 per year; van failure arrives at the speed of ¼ per year; and total failure (for cars and vans) arrives a speed of $\left(\frac{1}{8}+\frac{1}{4}\right)$ per year. Of the total failure, car failure accounts for $\frac{\frac{1}{8}}{\frac{1}{8}+\frac{1}{4}}=\frac{1}{3}$ of the total failure.

With this intuitive explanation, you should easily memorize the following shortcut:

> In general, if X and Y are two independent exponential random variables with parameters of λ_1 and λ_2 respectively:
>
> $$f_X(x)=\lambda_1 e^{-\lambda_1 x} \text{ and } f_Y(y)=\lambda_2 e^{-\lambda_2 y}$$
>
> Then $P(X<Y)=\int_0^{+\infty} f_X(x)P(Y>x)dx=\int_0^{+\infty}\lambda_1 e^{-\lambda_1 x}e^{-\lambda_2 x}dx=\lambda_1\int_0^{+\infty} e^{-(\lambda_1+\lambda_2)x}dx=\frac{\lambda_1}{\lambda_1+\lambda_2}$
>
> Similarly, $P(Y<X)=\frac{\lambda_2}{\lambda_1+\lambda_2}$.

Now you see that $P(X<Y)+P(X>Y)=\frac{\lambda_1}{\lambda_1+\lambda_2}+\frac{\lambda_2}{\lambda_1+\lambda_2}=1$. This means that $P(X=Y)=0$. To see why $P(X=Y)=0$, please note that $X=Y$ is a line in the 2-D plane. A line doesn't have any area (i.e. the area is zero). If you integrate the joint pdf over a line, the result is zero.

If you have trouble understanding why $P(X=Y)=0$, you can think of probability in a 2-D plane as a volume. You can think of the joint pdf in a 2-D plane as the height function. In order to have a volume, you must integrate the height function over an area. A line doesn't have any area. Consequently, it doesn't have any volume.

Problem 13 (Sample P #90, also May 2000 #10)

An insurance company sells two types of auto insurance policies: Basic and Deluxe. The time until the next Basic Policy claim is an exponential random variable with mean two days. The time until the next Deluxe Policy claim is an independent exponential random variable with mean three days.

What is the probability that the next claim will be a Deluxe Policy claim?
(A) 0.172
(B) 0.223
(C) 0.400
(D) 0.487
(E) 0.500

Solution

Let T^B = time until the next Basic policy is sold. T^B is exponential random variable with $\lambda^B = \frac{1}{\theta^B} = \frac{1}{2}$.

Let T^D = time until the next Deluxe policy is sold. T^D is exponential random variable with $\lambda^D = \frac{1}{\theta^D} = \frac{1}{3}$.

"The next claim is a Deluxe policy" means that $T^D < T^B$.

$$P\left(T^D < T^B\right) = \frac{T^D}{T^D + T^B} = \frac{\frac{1}{3}}{\frac{1}{3} + \frac{1}{2}} = \frac{2}{5} = 0.4$$

Homework for you: #3 May 2000; #9, #14, #34 Nov 2000; #20 May 2001; #35 Nov 2001; #4 May 2003.

About the author

Yufeng Guo was born in central China. After receiving his Bachelor's degree in physics at Zhengzhou University, he attended Beijing Law School and received his Masters of law. He was an attorney and law school lecturer in China before immigrating to the United States. He received his Masters of accounting at Indiana University. He has pursued a life actuarial career and passed exams 1, 2, 3, 4, 5, 6, and 7 in rapid succession after discovering a successful study strategy.

Mr. Guo's exam records are as follows:

Fall 2002	Passed Course 1
Spring 2003	Passed Courses 2, 3
Fall 2003	Passed Course 4
Spring 2004	Passed Course 6
Fall 2004	Passed Course 5
Spring 2005	Passed Course 7

Mr. Guo currently teaches an online prep course for Exam P, FM, MFE, and MLC. For more information, visit http://actuary88.com.

If you have any comments or suggestions, you can contact Mr. Guo at **yufeng.guo.actuary@gmail.com**.

Value of this PDF study manual

1. Don't pay the shipping fee (can cost $5 to $10 for U.S. shipping and over $30 for international shipping). Big saving for Canadian candidates and other international exam takers.

2. Don't wait a week for the manual to arrive. You download the study manual instantly from the web and begin studying right away.

3. Load the PDF in your laptop. Study as you go. Or if you prefer a printed copy, you can print the manual yourself.

4. Use the study manual as flash cards. Click on bookmarks to choose a chapter and quiz yourself.

5. Search any topic by keywords. From the Adobe Acrobat reader toolbar, click Edit ->Search or Edit ->Find. Then type in a key word.

User review of Mr. Guo's P Manual

Mr. Guo's P manual has been used extensively by many Exam P candidates. For user reviews of Mr. Guo's P manual at http://www.actuarialoutpost.com, click here Review of the manual by Guo.

Testimonies:

"**Second time I used the Guo manual and was able to do some of the similar questions in less than 25% of the time because of knowing the shortcut**."
Testimony #1 of the manual by Guo

"I just took the exam for the second time and feel confident that I passed. I used Guo the second time around. It was very helpful and gives a lot of shortcuts that I found very valuable. I thought the manual was kind of expensive for an e-file, but if it helped me pass it was well worth the cost."
Testimony # 2 of the manual by Guo

"I took the last exam in Feb 2006, and I ran out of time and I ended up with a five. I needed to do the questions quicker and more efficiently. The Guo's study guide really did the job."
Testimony #3 of the manual by Guo