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

by Yufeng Guo

For SOA Exam P/CAS Exam 1 on May 25, 2005

Exam P testing will be computer-based starting from Fall 2005. Using this e-manual will get you comfortable with taking an electronic test of Exam P.

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

Table of contents

Chapter 16 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 $30 \times 24 \times 60 = 43,200$ 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.

Pr (it takes m minutes or less to have the first claim)
=1 – Pr (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 $(1-p)^m$. 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 $(1-p)^m$.

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

Now we are ready to find the pdf of T:

$$\Pr(T \le t) = \Pr(43,200t \text{ trials or fewer before the first success})$$

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

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 $n \to +\infty$.

$$\Pr(T \le t) = \Pr(nt \text{ trials or fewer before the first success})$$

$$= 1 - \left(1 - \frac{1}{3n}\right)^{nt} = 1 - \left[\left(1 - \frac{1}{3n}\right)^{-3n}\right]^{-t/3}$$

$$= 1 - e^{-t/3} \quad (\text{as } n \to +\infty)$$

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.

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

$$\Pr(T > t) = 1 - F(t) = e^{-t/\theta}$$

Alternatively,

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

$$\Pr(T > t) = 1 - F(t) = e^{-\lambda t}$$

Mean and variance: $E(T) = \theta, \quad Var(T) = \theta^2$

Like geometric distribution, exponential distribution lacks memory and $\Pr(T > a + b \mid T > a) = \Pr(T > b)$. We can easily derive this:

$$\Pr(T > a + b \mid T > a) = \frac{\Pr(T > a + b)}{\Pr(T > a)} = \frac{e^{-(a+b)/\theta}}{e^{-a/\theta}} = e^{-b/\theta} = \Pr(T > b)$$

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 $\theta = E(T) = 100$

$\Pr(T > t) = 1 - F(t) = e^{-t/\theta}$

$$\Pr(T > 100) = e^{-100/100} = e^{-1} = 0.3679$$
$$\Pr(T > 400) = e^{-400/100} = e^{-4} = 0.0183$$
$$\Pr(T > 400 \mid T > 100) = \frac{\Pr(T > 400)}{\Pr(T > 100)} = \frac{e^{-4}}{e^{-1}} = e^{-3} = 0.0498$$

Or use the lack of memory property of exponential distribution:

$$\Pr(T > 400 \mid T > 100) = \Pr(T > 400 - 100 = 300) = e^{-3} = 0.0498$$

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

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

Problem 3

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

Find $E(T \mid T > 3)$, $Var(T \mid T > 3)$, $E(T \mid T \le 3)$, $Var(T \mid T \le 3)$.

Solution

$E(T \mid T > 3)$ is the same as $\frac{E(Y)}{\Pr(T > 3)}$, where $Y = 0$ for $T \le 3$ and $Y = X$ for $T > 3$.

$$E(T \mid T > 3) = \int_3^{+\infty} t \frac{f(t)}{\Pr(T > 3)} dt = \int_3^{+\infty} \frac{tf(t)}{1 - F(3)} dt = \frac{1}{1 - F(3)} \int_3^{+\infty} tf(t) dt$$

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

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

$$E(T \mid T>3) = \frac{1}{1-F(3)} \int_{3}^{+\infty} tf(t)dt = \frac{1}{1-F(3)} \int_{3}^{+\infty} tf(t)dt$$

$$= \frac{5e^{-3/2}}{e^{-3/2}} = 5$$

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, $(T-3 \mid T>3)$ is exponentially distributed with an identical mean of 2.

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

$$E(T \mid T>3) = E[(T-3) \mid T>3] + 3 = 2+3 = 5$$

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

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

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

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

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

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

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

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

Next, we need to find $E(T \mid T \le 3)$.

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

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

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

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

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

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

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

$$E(T)=E(T \mid T<3)\times\Pr(T<3)+E(T \mid T>3)\times\Pr(T>3)$$

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 $\Pr(T=3)$ is not included in the right-hand side because the probability density of a continuous random variable is zero at any single point. Put another way, the pdf of a continuous random variable X is meaningful only if you integrate it over a range of X such as $a<X<b$. At a single point, pdf is zero.

Of course, you can also write:

$$E(T)=E(T \mid T\le 3)\times\Pr(T\le 3)+E(T \mid T>3)\times\Pr(T>3)$$

Or $E(T)=E(T \mid T<3)\times\Pr(T<3)+E(T \mid T\ge 3)\times\Pr(T\ge 3)$

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

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

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

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

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

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

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

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

Alternatively,

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

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

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

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

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

$(T-a|T>a)$ is also exponentially distributed with mean θ

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

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

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

$$\Rightarrow \quad \int_{a}^{+\infty} (t-a) f(t) dt = E(T-a \mid T>a) \times \Pr(T>a) = \theta e^{-\theta a}$$

$$E(T \mid T>a) = \frac{1}{\Pr(T>a)} \int_{a}^{+\infty} t f(t) dt$$

$$\Rightarrow \quad \int_{a}^{+\infty} t f(t) dt = E(T \mid T>a) \times \Pr(T>a) = (\theta + a) e^{-\theta a}$$

$$E(T \mid T<a) = \frac{1}{\Pr(T<a)} \int_{0}^{a} t f(t) dt$$

$$\Rightarrow \quad \int_{0}^{a} t f(t) dt = E(T \mid T<a) \times \Pr(T<a) = E(T \mid T<a)(1-e^{-\theta a})$$

$$E(T) = E(T \mid T<a) \times \Pr(T<a) + E(T \mid T>a) \times \Pr(T>a)$$

$$\Rightarrow \quad \theta = E(T \mid T<a) \times (1-e^{-\theta a}) + E(T \mid T>a) \times e^{-\theta a}$$

$$E(T^2) = E(T^2 \mid T<a) \times \Pr(T<a) + E(T^2 \mid T>a) \times \Pr(T>a)$$

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 $a \ge 0$

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

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

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

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 $\Pr(X > a) = 1 - F(a)$, 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

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

To understand Equation (3), note that

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

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

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

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

Then

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

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

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

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

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

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

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 $a \geq 0$

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

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

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

So you have three 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.

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

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 $E(X)$ 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. $X = \max(2, T)$.

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

You can also write

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

You can use this approach to find $E(X^2)$ 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 $T > 2$.

$$E(X) = \int_0^2 t \frac{1}{3} e^{-t/3} dt + \int_2^{+\infty} 2 (\frac{1}{3} e^{-t/3}) dt$$
$$\int_0^2 t \frac{1}{3} e^{-t/3} dt = 3 - (2+3) e^{-2/3}$$
$$\int_2^{+\infty} 2 (\frac{1}{3} e^{-t/3}) dt = 2e^{-2/3}$$
$$E(X) = 3 - 5e^{-2/3} + 2e^{-2/3} = 3 - 3e^{-2/3}$$

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

$$E(X^2) = \int_0^{+\infty} x^2(t) f(t) dt = \int_0^2 t^2 f(t) dt + \int_2^{+\infty} 2^2 f(t) dt$$
$$\int_0^2 t^2 f(t) dt = [(0+3)^2 + 3^2] e^{-0/3} - [(2+3)^2 + 3^2] e^{-2/3} = 18 - 34e^{-2/3}$$
$$\int_2^{+\infty} 2^2 f(t) dt = 4e^{-2/3}$$
$$E(X^2) = 18 - 34e^{-2/3} + 4e^{-2/3} = 18 - 30e^{-2/3}$$
$$Var(X) = E(X^2) - E^2(X) = (18 - 30e^{-2/3}) - (3 - 3e^{-2/3})^2$$

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

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

We know that

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

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.

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

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

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

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

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

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

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

$E\left[\min(T,2)\right]^2=18-30e^{-2/3}$ (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, $f(x) = \frac{1}{200} e^{-x/200}$.

Let Y =claim payment by the insurance company.

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

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

$$\int_0^{100} 0 f(x) dx = 0$$
$$\int_{100}^{400} (x-100) f(x) dx = \int_{100}^{400} x f(x) dx - \int_{100}^{400} 100 f(x) dx$$

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

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

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

Then we have

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

Alternatively, we can use the shortcut developed in Chapter 20:

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

Problem 9

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.

$$E(Y) = E(X | X > 30{,}000) = E(X - 30{,}000 | X > 30{,}000) + 30{,}000$$
$$= E(X) + 30{,}000 = \theta + 30{,}000 = 38{,}000$$

$$E(Z) = E(X | X > 60{,}000) = E(X - 60{,}000 | X > 60{,}000) + 60{,}000$$
$$= E(X) + 60{,}000 = \theta + 60{,}000 = 68{,}000$$

Problem 10

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

Solution

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

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

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

Homework for you: #3 May 200; #9,#14,#34 Nov 2000; #20 May 2001; #35 Nov 2001; #4 May 2003.

About the author

Yufeng Guo is a life actuary at Bloomington, IL. He 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, and 6 in rapid succession after discovering a successful study strategy. He completed this study manual while preparing for Course 5 for Fall 2004.

Mr. Guo's exam records include:

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

If you have any comments or questions, you can contact Mr. Guo at yufeng_guo@fastmail.fm

FAQ

Q How is this study manual different from other study manuals?

A This study manual is written with a simple philosophy: to provide hard-to-find conceptual insights and computational shortcuts that really work in the heat of the exam. This manual has no academic solutions that sound good on the paper but fall apart in the exam. Every solution offered in this manual is proven to work under the intense pressure of actual exam conditions.

Q Can I buy a hard copy of your manual?

A No. This book is an e-book only and no hard copy is available. E-manual has many advantages:

<u>Cost saving</u>. The author eliminates the printing cost and passes the saving to you. As a result, you can buy this e-manual at low cost of $45 per copy.
<u>Buying the manual instantly on-line</u>. If you buy a traditional paper copy of a study manual, you have to wait for several days before the book gets to your hands. When you buy this e-manual, you can purchase it instantly from the web.

<u>Getting ready for the computer based testing</u>. Starting from fall, 2005, Exam P will be computer based. Traditional paper-and-pencil exams will go away. By using this e-manual, you'll get used to the computer based problem testing.

Q Where can I buy this manual?

A This e-manual is sold exclusively at Actuarial Bookstore www.actuarialbookstore.com for $45 a copy. Because this is an e-book, you don't pay any shipping or handling charges.

Q I have a Mac computer. Does your e-manual work in my computer?

A Yes. If you buy this book, you'll get a PDF (portable document format) file. This PDF file works for both PC and Mac, as long as you have Acrobat Reader 6.0 installed in your computer. You can download a free Acrobat Reader 6.0 from the internet if you don't have one already.

Q Does your e-manual have bookmarks so I can quickly find a chapter?

A Yes. This PDF has bookmarks listing the table of the content. You can quickly find any chapter by clicking a bookmark.

Q I already purchased other study manuals. Is this manual compatible with my existing manuals?

A Yes. You can use this manual as the only manual or use it in combination with any other study manuals you currently have.