Logarithm of exponential and Cauchy random variables

Josef Bukac, Bulharska 298, Jaromer-Josefov 55102
Czech Republic
bukacjosef@seznam.cz

1. Introduction.

Exponential distribution is frequently used in reliability engineering. Lognormal distribution is also used. It is the distribution, the logaritm of which has a normal distibution. What if we take a logarithm of exponential distribution. We want to calculate the first moment and the second moment.

It can be done in two ways. One way is to integrate the transormation function, which is the logaritm in this case, multiplied by the density function. The other way is to find the density function of a new variable and use it to calculate the first moment. Both ways yield the same result.

The same idea is applied to the Cauchy distribution. The moments of the Cauchy distribution are not defined but the moments of the logarithm of its absolute value can be calculated.

2. Exponential distribution.

The density function of exponentially distributed random variable X is defined as $\mu e^{-\mu x}$ for $x \geq 0$, equal to zero for $x < 0$, provided $\mu > 0$. We will study the case when $\mu = 1$, that is, the density is just e^{-x}. We define the new variable Y as $Y = \ln(X)$.

1) The first way to define the first moment of Y is

$$E(Y) = \int_0^\infty e^{-x} \ln(x) dx = -\gamma$$

where γ is the Euler constant.

The proof may be found in Nahin (2015) p.173. It is interesting to find out that $\gamma = -0.5772\ldots$ appears also in statistics.

2) The second way is described in most textbook on probability theory. Let $F_X(x)$ denote the cdf of the random variable X and $f_X(x)$ its pdf. Assume $g(x)$ is the transformation of X giving us a random variable $Y = g(X)$. We denote by $F_Y(y)$ a cdf of Y and its pdf by $f_Y(y)$. We assume that $g(x)$ is differentiable and increasing, its inverse is denoted by $h(y) = g^{-1}(y)$.

Then we differentiate a composite function

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy} F_X(h(y)) = f_X(h(y))h'(y).$$

In our case we have

$$f_X(x) = e^{-x}, g(x) = \ln(x), h(y) = e^y$$

and the density function of Y is

$$f_Y(y) = e^{-e^y} e^y = e^{y-e^y}.$$

We just verify that

$$\int_{-\infty}^{\infty} e^{y-e^y} dy = -e^{-e^y}\Big|_{-\infty}^{\infty} = -(0-1) = 1.$$

We write the expectation of Y in both ways as

$$E(Y) = \int_{-\infty}^{\infty} y e^{y-e^y} dy = \int_0^{\infty} e^{-x} \ln(x) dx = -\gamma.$$

We have actually calculated the integral

$$\int_{-\infty}^{\infty} y e^{y-e^y} dy = -\gamma.$$

The second moment is

$$M_2(Y) = \int_0^{\infty} e^{-x} (\ln(x))^2 dx = \frac{\pi^2}{6} + \gamma^2.$$

The proof can be found at
https://www.youtube.com/watch?v=wUY6TJxfFTk

The second moment can also be written as

$$\int_{-\infty}^{\infty} y^2 e^{y-e^y} dy = \frac{\pi^2}{6} + \gamma^2.$$

The fact that the second moment is finite means, by the way, that we are free to use the central limit theorem.

3. Cauchy distribution.

The pdf of the standard Cauchy distribution X is

$$f(x) = \frac{1}{\pi(x^2+1)}$$

defined for any x. If consider the absolute value, the pdf of $|X|$ is

$$f_A(x) = \frac{2}{\pi(x^2+1)}$$

defined this way for $x \geq 0$, otherwise $f_A(x) = 0$ for $x < 0$.

We study the transformation $Y = \ln|X|$. First we find out what the pdf is

$$f_Y(y) = f_A(h(y))h'(y) = \frac{2e^y}{\pi(e^{2y}+1)}.$$

We claim that the pdf is symmetrical about zero. To show this, we substitute $z = -y$ and multiply both the numerator and denominator by e^{2y}.

$$f_Y(z) = \frac{2e^z}{\pi(e^{2z}+1)} = \frac{2e^{-y}}{\pi(e^{-2y}+1)} = \frac{2e^y}{\pi(1+e^{2y})} = f_Y(y).$$

Thus $f_Y(-y) = f_Y(y)$.

Since the pdf is symmetrical is is immediate that all the odd moments are zero. The expectation is

$$E(Y) = \int_{-\infty}^{\infty} \frac{2ye^y}{\pi(e^{2y}+1)} dy = 0$$

We get a zero if we calculate the expectation as

$$E(Y) = \int_0^{\infty} \ln(x) f_A(x) dx = \int_0^{\infty} \frac{2\ln(x)}{\pi(x^2+1)} dx = 0.$$

As a diversion we conside what Nahin (2015) p.151 presents

$$\int_1^{\infty} \frac{\ln(x)}{(x^2+1)} dx = G$$

where $G = 0.91596\ldots$ is the Catalan constant defined in Nahin (2015) p.xv. Since $0 = \int_0^{\infty} = \int_0^1 + \int_1^{\infty}$, we get

$$\int_0^{1} \frac{\ln(x)}{(x^2+1)} dx = -G$$

Let us find out what the second moment is. One way to do it is this

$$M_2(Y) = \int_0^{\infty} \frac{2(\ln(x))^2}{\pi(x^2+1)} dx = \frac{\pi^2}{4}.$$

The proof of $\int_0^{\infty} (\ln(x)^2/(x^2+1))dx = \pi^3/8$ is in Nahin (2015) p.160.

The second way is the use the pdf of Y, that is,

$$M_2(Y) = \int_{-\infty}^{\infty} \frac{2y^2e^y}{\pi(e^{2y}+1)} dy.$$

Two ways of defining $M_2(Y)$ give the same result which is $\pi^2/4$.

4. Fourth moment.

Since the pdf is symmetrical, it will suffice to calculate

$$I = \int_0^{\infty} \frac{y^4e^y}{e^{2y}+1} dy = \int_0^{\infty} \frac{y^4e^{-y}}{1+e^{-2y}} dy$$

because $M_4(Y) = 4I/\pi$. We give only an outline of the calculations. We use the geometric series $1/(1-r) = 1 + r + r^2 + \ldots$ for $|r| < 1$ in which we set $r = -q$ to yield $1/(1+q) = 1 - q + q^2 - q^3 + \ldots$ for $|q| < 1$. Thus

$$I = \int_0^{\infty} y^4e^{-y} \sum_{k=0}^{\infty} (-1)^k (e^{-2y})^k dy = \sum_{k=0}^{\infty} (-1)^k \int_0^{\infty} y^4 e^{-y}(e^{-2y})^k dy.$$

$$= \sum_{k=0}^{\infty} (-1)^k \int_0^{\infty} y^4 e^{-y} (e^{-2y})^k dy = \sum_{k=0}^{\infty} (-1)^k \int_0^{\infty} y^4 e^{-(2k+1)y} dy.$$

The antiderivative $\int x^4 e^{-mx} dx$ is easily verified to be

$$\int x^4 e^{-mx} dx = \frac{-e^{-mx}(m^4x^4 + 4m^3x^3 + 12m^2x^2 + 24mx + 24)}{m^5} + C$$

Now we calculate two limits to obtain

$$\int_0^{\infty} x^4 e^{-mx} dx = \frac{-e^{-mx}(m^4x^4 + 4m^3x^3 + 12m^2x^2 + 24mx + 24)}{m^5}\Big|_0^{\infty} = \frac{24}{m^5}.$$

When $m = 2k + 1$, we obtain

$$I = \sum_{k=0}^{\infty} (-1)^k \frac{24}{(2k+1)^5}.$$

Finally we get

$$M_4(Y) = \frac{4}{\pi} I = \frac{96}{\pi} \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)^5}.$$

The series on the right hand side converges reasonable fast and we can easily estimate the remainder of an alternating series. The sum of fifty terms of the series gives us $M_4(Y) = 30.440$. The other way to calculate $M_4(Y)$ is numerical integration.

5. Substitution.

We assume that X is a random variable with a pdf f_X positive on some interval I_X and zero otherwise. We define a random variable $Y = g(X)$ taking on values in the interval obtained by mapping the interval I_X to the interval I_Y or $I_Y = g(I_X)$. In our examples we get $I_X = [0, \infty)$ and $I_Y = (-\infty, \infty)$. We assume that $g(x)$ is differentiable and increasing, its inverse is $h(y) = g^{-1}(y)$.

When we express the n-th moment as

$$\int_{I_X} g(x)^n f_X(x) dx$$

and use the substitution $y = g(x), x = h(y), dx = h'(y)dy$, we get

$$\int_{I_Y} y^n f_X(h(y)) h'(y) dy.$$

This is exactly what we would get if we calculated the n-th moment with the help of pdf of f_Y.

6. Conclusion.

We have shown that certain integrals that are impossible to calculate with the help of antiderivatives may be applied in probability theory. The Euler constant may appear in such examples.

It is interesting to find out that the pdf of the logarithm of absolute value of a random variable with standard Cauchy distribution is symmetrical and it has finite moments.

It may be difficult to find evaluations of certain non trivial integrals. The book by Nahin (2015) was very helpful, but unfortunatly, the regular editor and its search function did not make the task easy.

6. References.

Nahin, P.J., Inside Interesting Integrals, (2015), Springer