CHAPTER 10

Some continuous distributions

10.1. Examples of continuous random variables

We look at some other continuous random variables besides normals.

Uniform distribution

A continuous random variable has *uniform distribution* if its density is $f(x) = 1/(b-a)$ if $a \leqslant x \leqslant b$ and 0 otherwise.

For a random variable X with uniform distribution its expectation is

$$\mathbb{E}X = \frac{1}{b-a}\int_a^b x\,dx = \frac{a+b}{2}.$$

Exponential distribution

A continuous random variable has *exponential distribution* with parameter $\lambda > 0$ if its density is $f(x) = \lambda e^{-\lambda x}$ if $x \geqslant 0$ and 0 otherwise.

Suppose X is a random variable with an exponential distribution with parameter λ. Then we have

$$\begin{aligned} \mathbb{P}(X > a) &= \int_a^\infty \lambda e^{-\lambda x} dx = e^{-\lambda a}, \\ F_X(a) &= 1 - \mathbb{P}(X > a) = 1 - e^{-\lambda a}, \end{aligned} \tag{10.1.1}$$

and we can use integration by parts to see that $\mathbb{E}X = 1/\lambda$, $\operatorname{Var} X = 1/\lambda^2$. Examples where an exponential random variable is a good model is the length of a telephone call, the length of time before someone arrives at a bank, the length of time before a light bulb burns out.

Exponentials are memoryless, that is,

$$\mathbb{P}(X > s+t \mid X > t) = \mathbb{P}(X > s),$$

or given that the light bulb has burned 5 hours, the probability it will burn 2 more hours is the same as the probability a new light bulb will burn 2 hours. Here is how we can prove this

$$\begin{aligned}\mathbb{P}(X > s+t \mid X > t) &= \frac{\mathbb{P}(X > s+t)}{\mathbb{P}(X > t)} \\ &= \frac{e^{-\lambda(s+t)}}{e^{-\lambda t}} = e^{-\lambda s} \\ &= \mathbb{P}(X > s),\end{aligned}$$

where we used Equation (10.1.1) for $a = t$ and $a = s + t$.

Gamma distribution

A continuous random variable has a *Gamma distribution* with parameters α and θ if its density is

$$f(x) = \frac{\alpha e^{-\alpha x}(\alpha x)^{\theta-1}}{\Gamma(\theta)}$$

if $x \geqslant 0$ and 0 otherwise. Here $\Gamma(\theta) = \int_0^\infty e^{-y}y^{\theta-1}dy$ is the gamma function. We denote such a distribution by $\Gamma(\alpha, \theta)$.

Note that $\Gamma(1) = \int_0^\infty e^{-y}dy = 1$, and using induction on n and integration by parts one can see that

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

so we say that gamma function interpolates the factorial.

While an exponential random variable is used to model the time for something to occur, a gamma random variable is the time for θ events to occur. A gamma random variable, $\Gamma\left(\frac{1}{2}, \frac{n}{2}\right)$, with parameters $\frac{1}{2}$ and $\frac{n}{2}$ is known as a χ_n^2, a chi-squared random variable with n degrees of freedom. Recall that in Exercise 8.9 we had a different description of a χ^2 random variable, namely, Z^2 with $Z \sim \mathcal{N}(0,1)$. Gamma and χ^2 random variables come up frequently in statistics.

Beta distribution

A continuous random variable has a *Beta distribution* if its density is

$$f(x) = \frac{1}{B(a,b)}x^{a-1}(1-x)^{b-1}, \qquad 0 < x < 1,$$

where $B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}$.

This is also a distribution that appears often in statistics.

Cauchy distribution

A continuous random variable has a *Cauchy distribution* if its density is

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

What is interesting about the Cauchy random variable is that it does not have a finite mean, that is, $\mathbb{E}|X| = \infty$.

Densities of functions of continuous random variables

Often it is important to be able to compute the density of $Y = g(X)$, where X is a continuous random variable. This is explained later in Theorem 11.1.

Let us give a couple of examples.

Example 10.1 (Log of a uniform random variable). If X is uniform on the interval $[0,1]$ and $Y = -\log X$, so $Y > 0$. If $x > 0$, then

$$\begin{aligned} F_Y(x) &= \mathbb{P}(Y \leqslant x) = \mathbb{P}(-\log X \leqslant x) \\ &= \mathbb{P}(\log X \geqslant -x) = \mathbb{P}(X \geqslant e^{-x}) = 1 - \mathbb{P}(X \leqslant e^{-x}) \\ &= 1 - F_X(e^{-x}). \end{aligned}$$

Taking the derivative we see that

$$f_Y(x) = \frac{d}{dx} F_Y(x) = -f_X(e^{-x})(-e^{-x}),$$

using the chain rule. Since $f_X(x) = 1$, $x \in [0,1]$, this gives $f_Y(x) = e^{-x}$, or Y is exponential with parameter 1.

Example 10.2 (χ^2, revisited). As in Exercise 8.9 we consider $Y = Z^2$, where $Z \sim \mathcal{N}(0,1)$. Then

$$\begin{aligned} F_Y(x) &= \mathbb{P}(Y \leqslant x) = \mathbb{P}(Z^2 \leqslant x) = \mathbb{P}(-\sqrt{x} \leqslant Z \leqslant \sqrt{x}) \\ &= \mathbb{P}(Z \leqslant \sqrt{x}) - \mathbb{P}(Z \leqslant -\sqrt{x}) = F_Z(\sqrt{x}) - F_Z(-\sqrt{x}). \end{aligned}$$

Taking the derivative and using the chain rule we see

$$f_Y(x) = \frac{d}{dx} F_Y(x) = f_Z(\sqrt{x}) \left(\frac{1}{2\sqrt{x}} \right) - f_Z(-\sqrt{x}) \left(-\frac{1}{2\sqrt{x}} \right).$$

Recall that $f_Z(y) = \frac{1}{\sqrt{2\pi}} e^{-y^2/2}$ and doing some algebra, we end up with

$$f_Y(x) = \frac{1}{\sqrt{2\pi}} x^{-1/2} e^{-x/2},$$

which is $\Gamma\left(\frac{1}{2}, \frac{1}{2}\right)$. As we pointed out before, this is also a χ^2 distributed random variable with one degree of freedom.

Example 10.3 (Tangent of a uniform random variable). Suppose X is a uniform random variable on $[-\pi/2, \pi/2]$ and $Y = \tan X$. Then

$$F_Y(x) = \mathbb{P}(X \leqslant \tan^{-1} x) = F_X(\tan^{-1} x),$$

and taking the derivative yields

$$f_Y(x) = f_X(\tan^{-1} x)\frac{1}{1+x^2} = \frac{1}{\pi}\frac{1}{1+x^2},$$

which is a Cauchy distribution.

10.2. Further examples and applications

Example 10.4. Suppose that the length of a phone call in minutes is an exponential random variable with average length 10 minutes.

(1) What is the probability of your phone call being more than 10 minutes?
Solution: Here $\lambda = \frac{1}{10}$, thus

$$\mathbb{P}(X > 10) = e^{-\left(\frac{1}{10}\right)10} = e^{-1} \approx 0.368.$$

(2) Between 10 and 20 minutes?
Solution: We have that

$$\mathbb{P}(10 < X < 20) = F(20) - F(10) = e^{-1} - e^{-2} \approx 0.233.$$

Example 10.5. Suppose the life of an Uphone has exponential distribution with mean life of 4 years. Let X denote the life of an Uphone (or time until it dies). Given that the Uphone has lasted 3 years, what is the probability that it will 5 more years.

Solution: in this case $\lambda = \frac{1}{4}$.

$$\mathbb{P}\left(X > 5 + 3 \mid X > 3\right) = \frac{\mathbb{P}\left(X > 8\right)}{\mathbb{P}\left(X > 3\right)}$$

$$= \frac{e^{-\frac{1}{4}\cdot 8}}{e^{-\frac{1}{4}\cdot 3}} = e^{-\frac{1}{4}\cdot 5} = \mathbb{P}\left(X > 5\right).$$

Recall that an exponential random variable is memoryless, so our answer is consistent with this property of X.

10.3. Exercises

Exercise 10.1. Suppose that the time required to replace a car's windshield can be represented by an exponentially distributed random variable with parameter $\lambda = \frac{1}{2}$.

(a) What is the probability that it will take at least 3 hours to replace a windshield?
(b) What is the probability that it will take at least 5 hours to replace a windshield given that it hasn't been finished after 2 hours?

Exercise 10.2. The number of years a Uphone functions is exponentially distributed with parameter $\lambda = \frac{1}{8}$. If Pat buys a used Uphone, what is the probability that it will be working after an additional 8 years?

Exercise 10.3. Suppose that the time (in minutes) required to check out a book at the library can be represented by an exponentially distributed random variable with parameter $\lambda = \frac{2}{11}$.

(a) What is the probability that it will take at least 5 minutes to check out a book?
(b) What is the probability that it will take at least 11 minutes to check out a book given that you have already waited for 6 minutes?

Exercise 10.4. Let X be an exponential random variable with mean $\mathbb{E}X = 1$. Define a new random variable $Y = e^X$. Find the PDF of Y, $f_Y(y)$.

Exercise 10.5. Suppose that X has an exponential distribution with parameter $\lambda = 1$. For any $c > 0$ show that $Y = \frac{X}{c}$ is exponential with parameter $\lambda = c$.

Exercise 10.6. Let X be a uniform random variable over $[0, 1]$. Define a new random variable $Y = e^X$. Find the probability density function of Y, $f_Y(y)$.

Exercise 10.7. An insurance company insures a large number of homes. The insured value, X, of a randomly selected home is assumed to follow a distribution with density function

$$f_X(x) = \begin{cases} \frac{8}{x^3} & x > 2, \\ 0 & \text{otherwise.} \end{cases}$$

(A) Given that a randomly selected home is insured for at most 4, calculate the probability that it is insured for less than 3.
(B) Given that a randomly selected home is insured for at least 3, calculate the probability that it is insured for less than 4.

Exercise 10.8. A hospital is to be located along a road of infinite length. If population density is exponentially distributed along the road, where should the station be located to minimize the expected distance to travel to the hospital? That is, find an a to minimize $\mathbb{E}|X - a|$, where X is exponential with rate λ.

10.4. Selected solutions

Solution to Exercise 10.1(A): We have

$$\begin{aligned}\mathbb{P}(X>3) &= 1-\mathbb{P}(0<X<3)\\ &= 1-\int_0^3 \frac{1}{2}e^{-\frac{x}{2}}dx = e^{-\frac{3}{2}} \approx 0.2231.\end{aligned}$$

Solution to Exercise 10.1(B): There are two ways to do this. The longer one is to explicitly find $\mathbb{P}(X>5 \mid X>2)$. The shorter one is to remember that the exponential distribution is memoryless and to observe that $\mathbb{P}(X>t+3 \mid X>t) = P(X>3)$, so the answer is the same as the answer to part (a).

Solution to Exercise 10.2: e^{-1}

Solution to Exercise 10.3(A): Recall that by Equation 10.1.1 $\mathbb{P}(X>a) = e^{-\lambda a}$, and therefore

$$\mathbb{P}(X>5) = e^{-\frac{10}{11}}.$$

Solution to Exercise 10.3(B): We use the memoryless property

$$\begin{aligned}\mathbb{P}(X>11 \mid X>6) &= \mathbb{P}(X>6+5 \mid X>6)\\ &= \mathbb{P}(X>5) = e^{-\frac{10}{11}}.\end{aligned}$$

Solution to Exercise 10.4: Since $\mathbb{E}[X] = 1/\lambda = 1$ then we know that $\lambda = 1$. Then its PDF and CDF are

$$\begin{aligned}f_X(x) &= e^{-x}, x \geqslant 0\\ F_X(x) &= 1-e^{-x}, x \geqslant 0.\end{aligned}$$

Thus

$$F_Y(y) = \mathbb{P}(Y \leqslant y) = \mathbb{P}\left(e^X \leqslant y\right) = \mathbb{P}(X \leqslant \log y) = F_X(\log y),$$

and so

$$F_Y(y) = 1-e^{-\log y} = 1-\frac{1}{y}, \text{ when } \log(y) \geqslant 0,$$

taking derivative we get

$$f_Y(y) = \frac{dF_y(y)}{dy} = \frac{1}{y^2} \text{ when } y \geqslant 1.$$

Solution to Exercise 10.5: Since X is exponential with parameter 1, then its PDF and CDF are

$$\begin{aligned}f_X(x) &= e^{-x}, x \geqslant 0\\ F_X(x) &= 1-e^{-x}, x \geqslant 0.\end{aligned}$$

Thus

$$F_Y(y) = \mathbb{P}(Y \leqslant y) = \mathbb{P}\left(\frac{X}{c} \leqslant y\right) = \mathbb{P}(X \leqslant cy) = F_X(cy),$$

and so

$$F_Y(y) = 1-e^{-cy}, \text{ when } cy \geqslant 0,$$

taking derivatives we get

$$f_Y(y) = \frac{dF_y(y)}{dy} = ce^{-cy}, \text{ when } y \geqslant 0.$$

Note that this is the PDF of an exponential with parameter $\lambda = c$.

Solution to Exercise 10.6: Since X is uniform over $[0,1]$, then its PDF and CDF are

$$f_X(x) = 1, 0 \leqslant x < 1$$
$$F_X(x) = x, 0 \leqslant x < 1.$$

Thus

$$F_y(y) = \mathbb{P}(Y \leqslant y) = \mathbb{P}\left(e^X \leqslant y\right) = \mathbb{P}(X \leqslant \log y) = F_X(\log y)$$

and so

$$F_Y(y) = \log y, \text{ when } 0 \leqslant \log y < 1.$$

taking derivatives we get

$$f_Y(y) = \frac{dF_y(y)}{dy} = \frac{1}{y}, \text{ when } 1 < y < e^1.$$

Solution to Exercise 10.7(A): Using the definition of conditional probability

$$\mathbb{P}(X < 3 \mid X < 4) = \frac{\mathbb{P}(X < 3)}{\mathbb{P}(X < 4)}.$$

Since

$$\mathbb{P}(X < 4) = \int_2^4 \frac{8}{x^3} dx = -\frac{4}{x^2}\Big|_2^4 = -\frac{1}{4} + 1 = \frac{3}{4}$$

and

$$\mathbb{P}(X < 3) = \mathbb{P}(2 < X < 3) = \int_2^3 \frac{8}{x^3} dx = -\frac{4}{x^2}\Big|_2^3 = -\frac{4}{9} + 1 = \frac{5}{9}.$$

Thus the answer is

$$\frac{\frac{5}{9}}{\frac{3}{4}} = \frac{20}{27} \approx 0.74074074$$

Solution to Exercise 10.7(B): Using the definition of conditional probability

$$\mathbb{P}(X < 4 \mid X > 3) = \frac{\mathbb{P}(3 < X < 4)}{\mathbb{P}(X > 3)}.$$

Since

$$\mathbb{P}(X > 3) = \int_3^\infty \frac{8}{x^3} dx = -\frac{4}{x^2}\Big|_3^\infty = \frac{4}{9},$$
$$\mathbb{P}(3 < X < 4) = \int_3^4 \frac{8}{x^3} dx = -\frac{4}{x^2}\Big|_3^4 = \frac{7}{36}.$$

Thus the answer is

$$\frac{\frac{7}{36}}{\frac{4}{9}} = \frac{7}{16} \approx 0.4375$$