Continuous random variables

Tan Do

Vietnamese-German University

Lecture 8

In this lecture

- Continuous random variables: probability density function and cumulative distribution function

Examples of continuous random variables

Example (Metal cylinder production) A company manufactures metal cylinders. These cylinders, which must slide freely within an outer casing, are designed to have a diameter of 50 mm.
However, the company discovers that the cylinders it manufactures can have a diameter anywhere between 49.5 and 50.5 mm.
Let X be the random variable which gives the diameter of a randomly chosen cylinder. Then X takes any value in the interval $[49.5, 50.5]$.

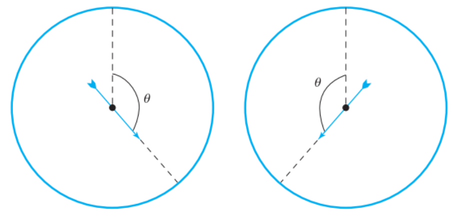
Example (Battery failure times) Suppose that a random variable X is the time to failure of a newly charged battery. Failure can be defined to be the moment at which the battery can no longer supply enough energy to operate a certain appliance.
Then X is a continuous random variable since it takes any value in $[0, \infty)$.

Example (Milk container contents) A machine-filled milk container is labeled as containing 2 liters. However, the actual amount of milk deposited into the container varies between 1.95 and 2.2 liters.

If the random variable X measures the amount of the milk in a randomly chosen container, it is a continuous random variable taking any value in the interval $[1.95, 2.2]$.

Games of chance

Dial-spinning game Suppose that a dial is spun and the angle θ between the dial and a fixed mark is measured ($0^\circ \le \theta \le 180^\circ$).



Let X be the random variable that gives the value of θ obtained from a spin. Then X is a continuous random variable with values in $[0, 180]$.

Suppose that a player spins the dial and then wins an amount corresponding to $\$1000 \times \frac{\theta}{180}$. Then we can consider the amount won by the player to be a random variable taking values in $[0, 1000]$.

Probability density function

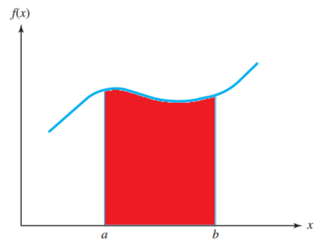
We use probability density function (p.d.f) to describe the probabilistic properties of a continuous random variable.

The probability that the random variable lies between 2 values a and b is given by

$$P(a \leq X \leq b) = \int_a^b f(x)\,dx.$$

It is required that $f(x) \geq 0$ for all x and

$$\int_{\text{state space}} f(x)\,dx = 1.$$



Note that the probability that a continuous random variable X takes a specific value a is always 0 because

$$P(X = a) = \int_a^a f(x)\,dx = 0.$$

Example (Metal cylinder production) Suppose that the diameter of a metal cylinder has a probability density function

$$f(x) = 1.5 - 6(x-50)^2$$

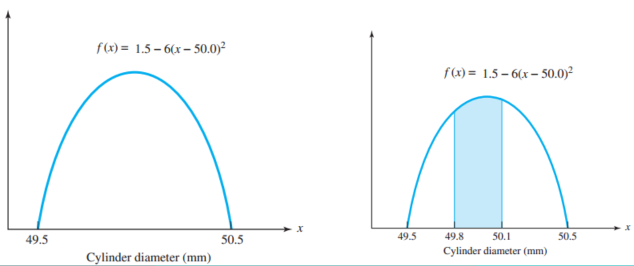
for $x \in [49.5, 50.5]$ and $f(x) = 0$ elsewhere.

This is a valid probability density function since $f(x) \geq 0$ for all x and

$$\int_{49.5}^{50.5} 1.5 - 6(x-50)^2\, dx = \left[1.5x - 2(x-50)^3\right]_{49.5}^{50.5} = 1.$$

The probability that a metal cylinder has a diameter between 49.8 and 50.1 mm is given by

$$\int_{49.8}^{50.1} 1.5 - 6(x-50)^2\, dx = \left[1.5x - 2(x-50)^3\right]_{49.8}^{50.1} = 0.432.$$



Example (Battery failure times) Suppose that the battery failure time, measured in hours, has a probability density function given by

$$f(x) = \frac{2}{(x+1)^3}$$

for $x \geq 0$ and $f(x) = 0$ elsewhere.

This is a valid probability density function since $f(x) \geq 0$ for all x and

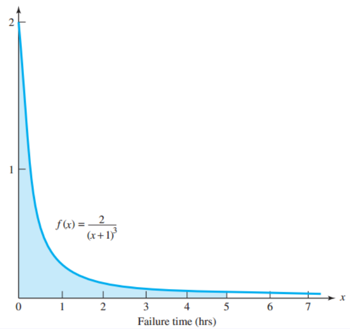
$$\int_0^\infty \frac{2}{(x+1)^3} = \left[\frac{-1}{(x+1)^2}\right]_0^\infty = 0 - (-1) = 1.$$

Suppose we want to calculate the probability that a battery lasts longer than 5 hours, i.e., $P(X > 5)$.
First we calculate

$$P(0 \leq x \leq 5) = \int_0^5 \frac{2}{(x+1)^3} = \left[\frac{-1}{(x+1)^2}\right]_0^5 = \frac{35}{36}.$$

So

$$P(X > 5) = 1 - P(0 \leq X \leq 5) = \frac{1}{36}.$$



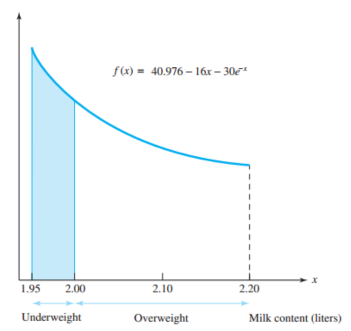
Example (Milk container contents) Suppose that the probability density function of the amount of milk deposited in a milk container is

$$f(x) = 40.976 - 16x - 30e^{-x}$$

for $x \in [1.95, 2.2]$ and $f(x) = 0$ elsewhere.

Exercise Show that this is a valid probability density function.

The figure illustrates the area that corresponds to the probability that the actual amount of milk is less than the advertised 2 liters.



Exercise Calculate the probability that the actual amount of milk is less than the advertised 2 liters.

Games of chance

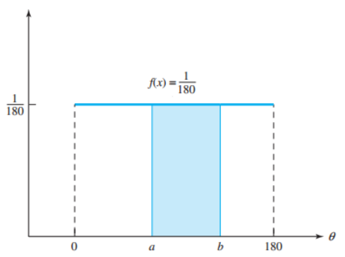
Dial spinning game If the dial is fair, i.e., the values of θ in $[0, 180]$ are all equally likely. This can be achieved with a "flat" probability density function as follows.

Since the total area under the probability density function is 1, its height must be $\frac{1}{180}$, and so

$$f(\theta) = \frac{1}{180}$$

for all $\theta \in [0, 180]$ and $f(\theta) = 0$ elsewhere.

This is an example of the uniform probability density function.



The probability that the angle θ lies between a and b is

$$P(a \leq \theta \leq b) = \int_a^b \frac{1}{180} = \frac{b-a}{180}.$$

That is, $P(a \leq \theta \leq b)$ is proportional to the length of the interval $[a, b]$.

The winnings

$$X = \$1000 \times \frac{\theta}{180}$$

are intuitively "equally likely" to be anywhere in $[0, 1000]$. So the probability density function should be also flat.

Since the total area under the probability density function is 1, the probability density function should be

$$f(x) = \frac{1}{1000}$$

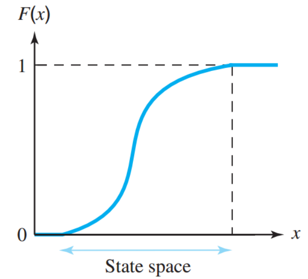
for $0 \leq x \leq 1000$ and $f(x) = 0$ elsewhere.

Cumulative distribution function

The cumulative distribution function for a continuous random variable is defined in exactly the same way as for a discrete random variable:

$$F(x) = P(X \leq x).$$

Note that $F(x)$ is a continuous, increasing function that is 0 prior to and at the beginning of the state space and is 1 at the end of and after the state space.



Given the probability density function $f(x)$, we can calculate the cumulative distribution function $F(x)$ by

$$F(x) = \int_{-\infty}^{x} f(x)\,dx.$$

If $F(x)$ is known, then $f(x)$ is found by

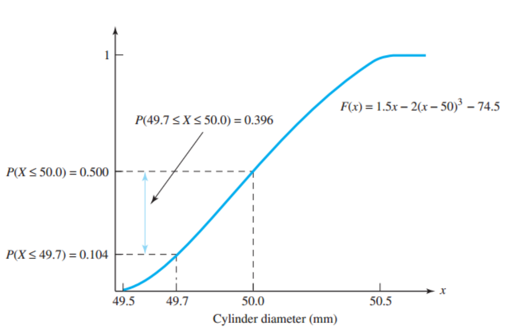
$$f(x) = \frac{dF(x)}{dx} = F'(x).$$

In addition, the cumulative function give a convenient way to calculate $P(a \leq X \leq b)$ since

$$P(a \leq X \leq b) = P(x \leq b) - P(x \leq a) = F(b) - F(a).$$

Example (Metal cylinder production) The cumulative distribution function of the metal cylinder diameters can be constructed from the probability density function as

$$F(x) = P(X \leq x) = \int_{49.5}^{x} 1.5 - 6(y-50)^2 \, dy = \left[1.5y - 2(y-50)^3\right]_{49.5}^{x} = 1.5x - 2(x-50)^3 - 74.5.$$



Exercise Calculate $P(49.7 \leq X \leq 50)$.

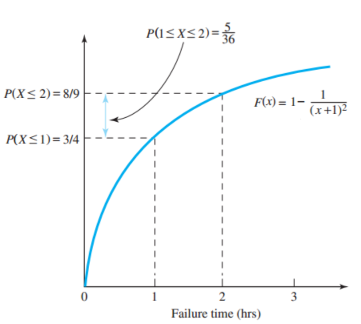
Example (Battery failure times) The cumulative distribution function of the battery failure times is

$$F(x) = P(X \leq x) = \int_0^x \frac{2}{(y+1)^3}\, dy = \left[\frac{-1}{(y+1)^2}\right]_0^x = 1 - \frac{1}{(x+1)^2}$$

for $x \geq 0$.

The probability that a battery lasts between 1 and 2 hours is

$$P(1 \leq X \leq 2) = F(2) - F(1) = \frac{8}{3} - \frac{3}{4} = \frac{5}{36}.$$



Games of chance

Dial spinning game The cumulative distribution function of the angle θ is

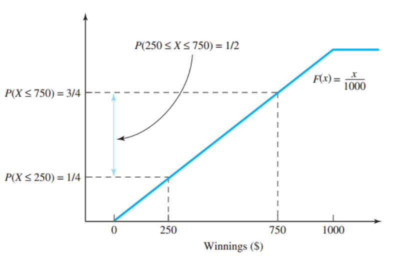
$$F(\theta) = \int_0^\theta f(y)\,dy = \int_0^\theta \frac{1}{180}\,dy = \frac{\theta}{180}$$

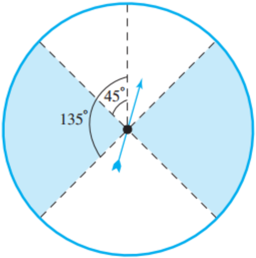
for $0 \le \theta \le 180$.

The cumulative distribution function of the winnings X is

$$F(x) = \int_0^x f(y)\,dy = \int_0^x \frac{1}{1000}\,dy = \frac{x}{1000}$$

for $0 \le x \le 1000$.





Dial region corresponding to winnings between \$250 and \$750

The probability that the winnings are between \$250 and \$750 is

$$P(250 \leq X \leq 750) = F(750) - F(250) = \frac{750}{1000} - \frac{250}{1000} = \frac{1}{2}.$$

This is expected since the corresponding values of θ is $[45°, 135°]$, which make up half of the dial.