Continuous Random Variables

James Heald[1]

[1]Gatsby Computational Neuroscience Unit
University College London

Gatsby Bridging Programme 2023

Objectives

- Introduce the concept and formal definition of a continuous random variable X and a probability density function.
- Learn how to find the probability that a continuous random variable falls in some interval [a, b].
- Learn that if X is continuous, the probability that X takes on any specific value is 0.
- Introduce the concept and formal definition of a cumulative distribution function of a continuous random variable.
- Learn how to find the cumulative distribution function of a continuous random variable X from the probability density function of X.

Discrete vs. continuous random variables

Unlike discrete random variables, which can take on a countable number of possible values (e.g. faces of a die or cards of a deck), continuous random variables can take on an uncountable number of possible values (e.g. all the real numbers in an interval).

Discrete vs. continuous random variables

Unlike discrete random variables, which can take on a countable number of possible values (e.g. faces of a die or cards of a deck), continuous random variables can take on an uncountable number of possible values (e.g. all the real numbers in an interval).

Examples

- the voltage membrane potential of a cell
- the interspike interval of a neuron
- the force generated by a muscle
- the velocity of an eye movement

Continuous random variables

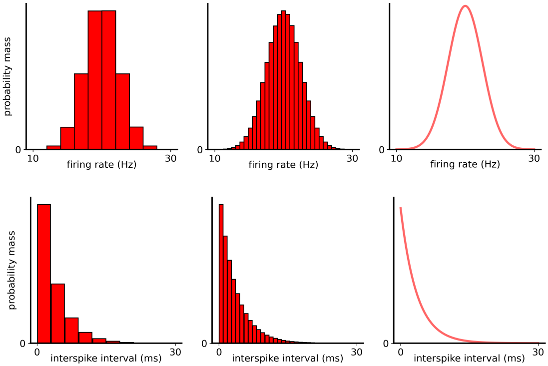
Definition

A random variable X is continuous if:

1. possible values comprise either a single interval on the number line (i.e. for some $a < b$, any number x between a and b is a possible value) or a union of disjoint intervals, and
2. $P(X = c) = 0$ for any number c that is a possible value of X.

Discrete probability distributions in the limit

Continuous random variables can be discretised into bins to form a discrete distribution that can be viewed as a probability histogram. As the bins become narrower, the histogram approaches a smooth curve.



The probability density function

Definition

The **probability density function** (PDF) of a continuous random variable X is a function $f(x)$ defined on the interval $(-\infty, \infty)$ such that for any two numbers a and b with $a \leq b$,

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

That is, the probability that X takes on a value in the interval [a, b] is the area under the graph of the density function above this interval.

The probability density function

Definition

The **probability density function** (PDF) of a continuous random variable X is a function $f(x)$ defined on the interval $(-\infty, \infty)$ such that for any two numbers a and b with $a \leq b$,

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

That is, the probability that X takes on a value in the interval [a, b] is the area under the graph of the density function above this interval.

A valid probability density function $f(x)$ has the following properties:

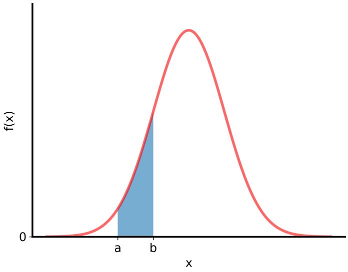
$$f(x) \geq 0 \text{ for all } x \tag{1}$$

$$\int_{-\infty}^{\infty} f(x)dx = 1. \tag{2}$$

Probabilities as integrals

The probability that a continuous random variable X takes on a value in the interval [a, b] is given by the area under the probability density function $f(x)$.

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



Density as probability per unit length

The probability of a small interval δ is approximately the density $\times$ δ:

$$\begin{aligned} P(x \leq X \leq x + \delta) &= \int_x^{x+\delta} f(t)dt \\ &\approx f(x) \times \delta \end{aligned}$$

Density as probability per unit length

The probability of a small interval δ is approximately the density $\times$ δ:

$$P(x \leq X \leq x + \delta) = \int_x^{x+\delta} f(t)dt$$
$$\approx f(x) \times \delta$$

Thus density is probability per unit length (rate of probability accumulation):

$$\frac{P(x \leq X \leq x + \delta)}{\delta} \approx f(x)$$

Each possible value has zero probability

The probability that X takes on a particular value a is 0, as

$$\begin{aligned} P(X = a) &= \int_a^a f(x)dx \\ &= \lim_{\epsilon \to 0} \int_{a-\epsilon}^{a+\epsilon} f(x)dx \\ &= 0. \end{aligned}$$

Each possible value has zero probability

The probability that X takes on a particular value a is 0, as

$$
\begin{aligned}
P(X = a) &= \int_a^a f(x)dx \\
&= \lim_{\epsilon \to 0} \int_{a-\epsilon}^{a+\epsilon} f(x)dx \\
&= 0.
\end{aligned}
$$

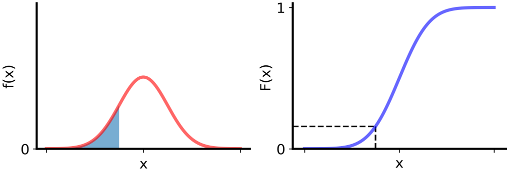
This implies that probabilities don't depend on interval end points:

$$P(a \leq X \leq b) = P(a < X < b) = P(a < X \leq b) = P(a \leq X < b),$$

as $P(X = a) = P(X = b) = 0$.

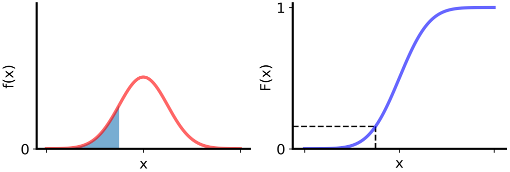
The cumulative distribution function

The cumulative distribution function (CDF) $F(x)$ is the area under the probability density function $f(x)$ to the left of x.



The cumulative distribution function

The cumulative distribution function (CDF) $F(x)$ is the area under the probability density function $f(x)$ to the left of x.



The CDF is a monotonically-increasing continuous function $F : \mathbb{R} \mapsto [0, 1]$ satisfying $\lim_{x \to -\infty} F(x) = 0$ and $\lim_{x \to \infty} F(x) = 1$.

The cumulative distribution function

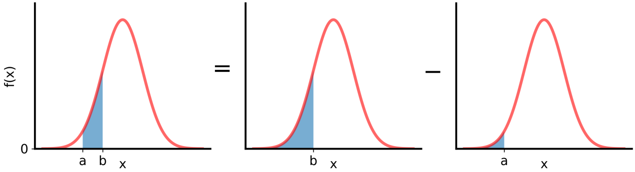
Definition

Let X be a continuous random variable with probability density function $f(x)$, then the **cumulative distribution function** $F(x)$ is defined as

$$\begin{aligned} F(x) &= P(X \leq x) \\ &= \int_{-\infty}^{x} f(t)dt. \end{aligned}$$

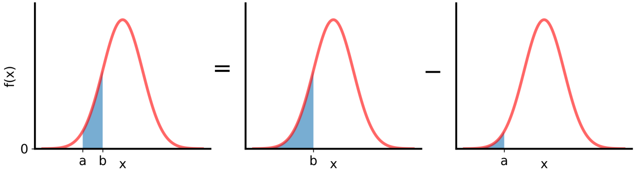
Computing probabilities using the CDF

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

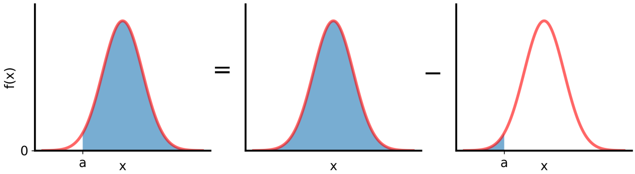


Computing probabilities using the CDF

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



$$\begin{aligned} P(X > a) &= F(\infty) - F(a) \\ &= 1 - F(a). \end{aligned}$$



Obtaining the PDF from the CDF

At every x at which the derivative $F'(x)$ exists, $F'(x) = f(x)$.

TODO: show Gaussians of different widths.

Obtaining the PDF from the CDF

At every x at which the derivative $F'(x)$ exists, $F'(x) = f(x)$.

Examples

When X has a uniform distribution, for $a < x < b$:

$$F'(x) = \frac{d}{dx}\left(\frac{x-a}{b-a}\right) = \frac{1}{b-a} = f(x)$$

TODO: show Gaussians of different widths.

Percentiles of a continuous distribution

Definition

Let p be a number between 0 and 1. The **(100p)th percentile** of the distribution of a continuous random variable X, denoted by $\eta(p)$, is defined by

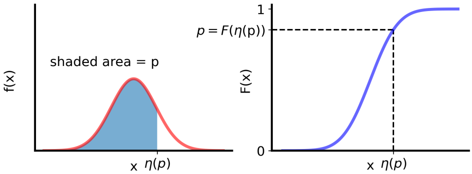
$$p = F(\eta(p)) = \int_{-\infty}^{\eta(p)} f(x)dx$$

Percentiles of a continuous distribution

Definition

Let p be a number between 0 and 1. The **(100p)th percentile** of the distribution of a continuous random variable X, denoted by $\eta(p)$, is defined by

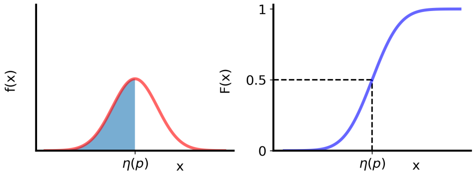
$$p = F(\eta(p)) = \int_{-\infty}^{\eta(p)} f(x)dx$$



Median

Definition

The **median** of a continuous distribution, denoted by $\tilde{\mu}$, is the 50th percentile, so $\tilde{\mu}$ satisfies $F(\tilde{\mu}) = 0.5$. That is, half the are area under the probability density function is to the left of $\tilde{\mu}$ and half is to the right of $\tilde{\mu}$.



Mean and variance

The expected value (mean) of a continuous random variable X is:

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f(x) \mathrm{d}x.$$

Mean and variance

The expected value (mean) of a continuous random variable X is:

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f(x) \mathrm{d}x.$$

The expected value of a function $g(x)$ of X is:

$$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f(x) \mathrm{d}x.$$

Mean and variance

The expected value (mean) of a continuous random variable X is:

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f(x) \mathrm{d}x.$$

The expected value of a function $g(x)$ of X is:

$$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f(x) \mathrm{d}x.$$

The variance of X is:

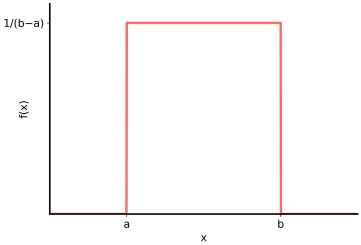
$$\begin{aligned} \mathrm{Var}[X] &= \mathbb{E}[(x - \mathbb{E}[X])^2] \\ &= \int_{-\infty}^{\infty} (x - \mathbb{E}[X])^2 f(x) \mathrm{d}x. \end{aligned}$$

Example: the uniform distribution

Definition

X is said to have a **uniform distribution** on the interval [a, b] if the PDF of X is:

$$f(x; a, b) = \begin{cases} \frac{1}{b-a} & \text{if } a \leq x \leq b \\ 0 & \text{otherwise.} \end{cases}$$



Example: the uniform distribution

When X has a uniform distribution, its expected value is:

$$\begin{aligned}\mathbb{E}[X] &= \int_{-\infty}^{\infty} x f(x) \mathrm{d}x \\ &= \int_{a}^{b} x \frac{1}{b-a} \mathrm{d}x \\ &= \frac{a+b}{2}.\end{aligned}$$

Example: the uniform distribution

When X has a uniform distribution, its expected value is:

$$\begin{aligned}\mathbb{E}[X] &= \int_{-\infty}^{\infty} x f(x) \mathrm{d}x \\ &= \int_a^b x \frac{1}{b-a} \mathrm{d}x \\ &= \frac{a+b}{2}.\end{aligned}$$

Its variance is:

$$\begin{aligned}\mathrm{Var}[X] &= \int_{-\infty}^{\infty} (x - \mathbb{E}[X])^2 f(x) \mathrm{d}x \\ &= \int_a^b \left(x - \frac{a+b}{2}\right)^2 \frac{1}{b-a} \mathrm{d}x \\ &= \frac{(b-a)^2}{12}.\end{aligned}$$

Sampling using the CDF

The inverse transform sampling algorithm is a procedure for sampling a continuous random variable using the inverse of its cumulative distribution function.

Sampling using the CDF

The inverse transform sampling algorithm is a procedure for sampling a continuous random variable using the inverse of its cumulative distribution function.

Recall that $F : \mathbb{R} \mapsto [0, 1]$.

Sampling using the CDF

The inverse transform sampling algorithm is a procedure for sampling a continuous random variable using the inverse of its cumulative distribution function.

Recall that $F : \mathbb{R} \mapsto [0, 1]$.

To draw a sample $x \sim f(x)$:

1. Sample $u \sim \mathrm{U}(0, 1)$ (using a pseudo-random number generator)
2. Let $x = F^{-1}(u)$

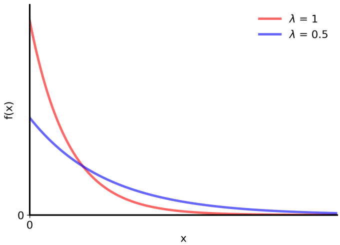
Example: the exponential distribution

Definition

X is said to have an **exponential distribution** on the interval $[0, \infty)$ if the PDF of X is:

$$f(x;\lambda) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{otherwise,} \end{cases}$$

where λ is a rate parameter that governs the rate of decay of $f(x)$.



Example: the exponential distribution

For $x \geq 0$, the PDF of the exponential distribution is:

$$f(x) = \lambda e^{-\lambda x},$$

Example: the exponential distribution

For $x \geq 0$, the PDF of the exponential distribution is:

$$f(x) = \lambda e^{-\lambda x},$$

which implies that the CDF is:

$$F(x) = 1 - e^{-\lambda x} = u,$$

Example: the exponential distribution

For $x \geq 0$, the PDF of the exponential distribution is:

$$f(x) = \lambda e^{-\lambda x},$$

which implies that the CDF is:

$$F(x) = 1 - e^{-\lambda x} = u,$$

and the inverse of the CDF is:

$$F^{-1}(u) = -\frac{\log(1-u)}{\lambda} = x.$$

Example: the exponential distribution

For $x \geq 0$, the PDF of the exponential distribution is:

$$f(x) = \lambda e^{-\lambda x},$$

which implies that the CDF is:

$$F(x) = 1 - e^{-\lambda x} = u,$$

and the inverse of the CDF is:

$$F^{-1}(u) = -\frac{\log(1-u)}{\lambda} = x.$$

Hence, to sample $x \sim f(x)$:

1. Sample $u \sim \mathrm{U}(0, 1)$
2. Let $x = -\frac{\log(1-u)}{\lambda}$

The normal (Gaussian) distribution