The Probability Density Function

Probability and Statistics for Data Science

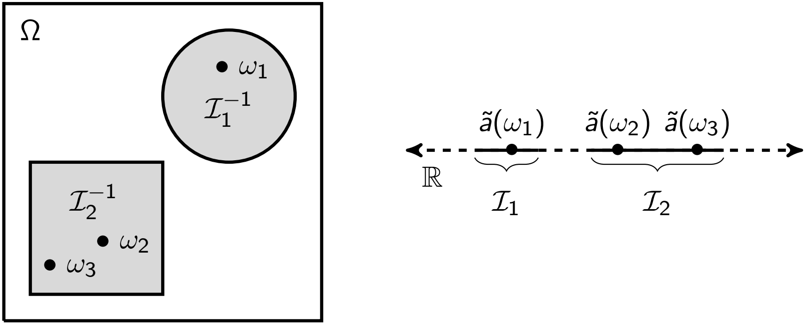
Carlos Fernandez-Granda

These slides are based on the book Probability and Statistics for Data Science by Carlos Fernandez-Granda, available for purchase here. A free preprint, videos, code, slides and solutions to exercises are available at https://www.ps4ds.net

Goal

Define probability density and describe its properties

Continuous random variables



Continuous random variables

We describe continuous random variables in terms of the probability that they belong to any interval

How do we encode this information?

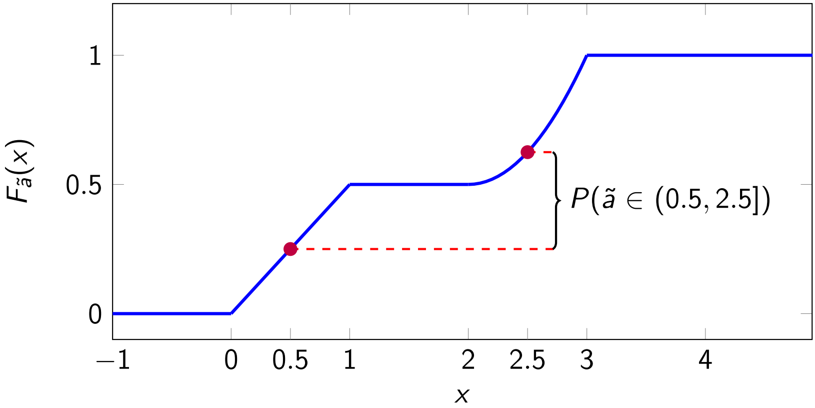
Cumulative distribution function

The cumulative distribution function (cdf) of a random variable $\tilde{a}$ is

$$F_{\tilde{a}}(a) := \mathrm{P}(\tilde{a} \leq a)$$

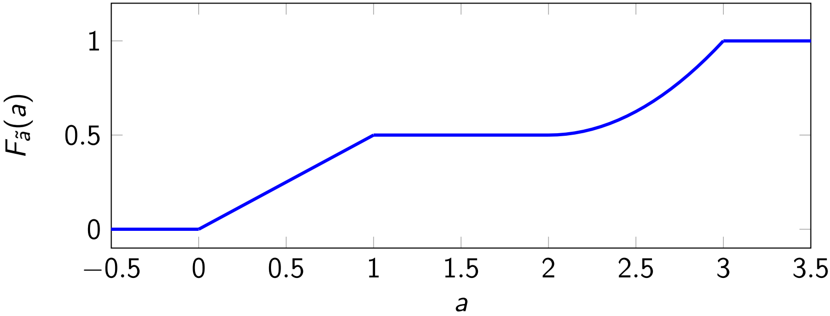
Probability that $\tilde{a}$ is less than or equal to a, for all $a \in \mathbb{R}$

Probability of an interval



Probability density

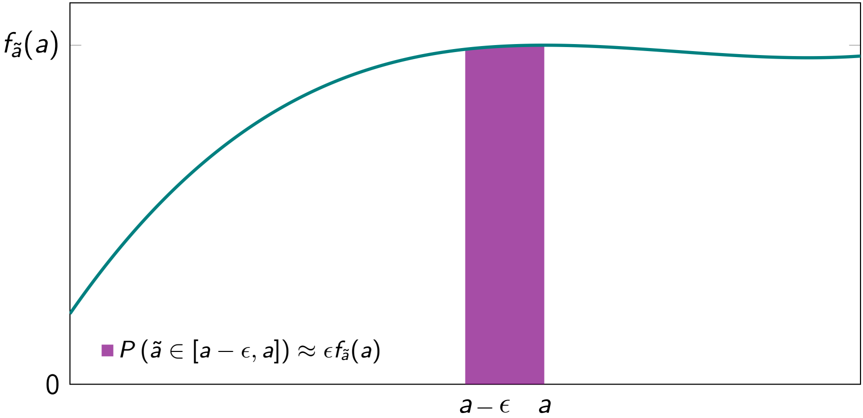
The cdf is a global quantity



How can we characterize local behavior?

Use density!

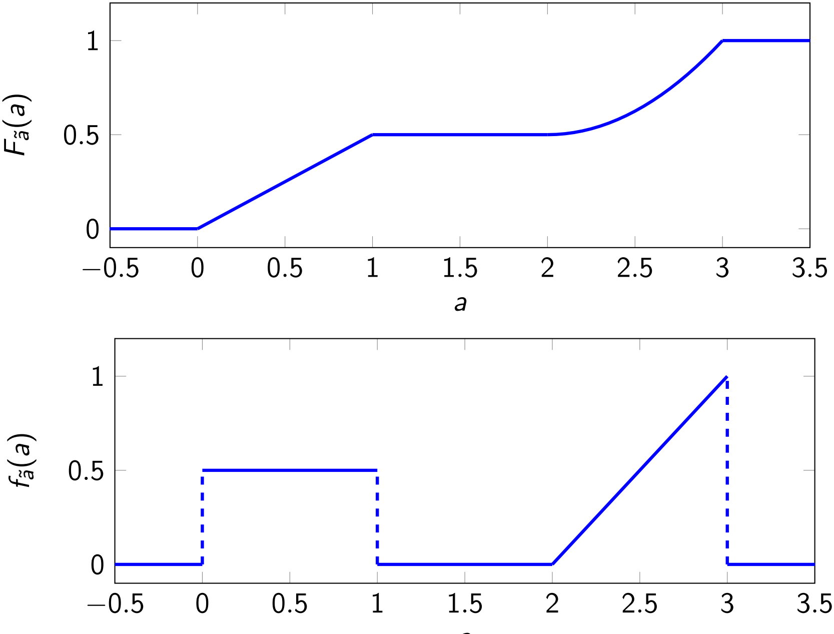
Probability density



The density is the derivative of the cdf

$$
\begin{aligned}
f_{\tilde{a}}(a) &= \lim_{\epsilon \to 0} \frac{\mathrm{P}(a - \epsilon \leq \tilde{a} \leq a)}{\epsilon} \\
&= \lim_{\epsilon \to 0} \frac{F(a) - F(a - \epsilon)}{\epsilon} \\
&= \frac{\mathrm{d}F_{\tilde{a}}(a)}{\mathrm{d}a}
\end{aligned}
$$

The pdf is the derivative of the cdf



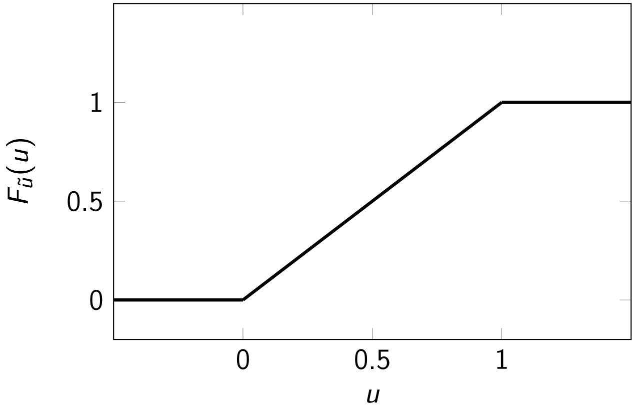
Probability density function

Let $\tilde{a} : \Omega \to \mathbb{R}$ be a random variable with cdf $F_{\tilde{a}}$

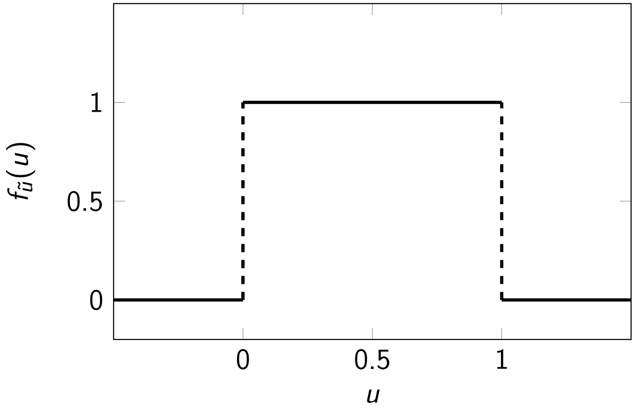
If $F_{\tilde{a}}$ is differentiable, the probability density function (pdf) of $\tilde{a}$ is

$$f_{\tilde{a}}(a) := \frac{\mathrm{d}F_{\tilde{a}}(a)}{\mathrm{d}a}$$

Uniform distribution



Uniform distribution



Uniform distribution

A uniform random variable $\tilde{u}$ on the interval $[a, b]$ has pdf

$$f_{\tilde{u}}(u) = \begin{cases} \frac{1}{b-a}, & \text{if } a \leq u \leq b \\ 0, & \text{otherwise} \end{cases}$$

Can a pdf be larger than one?

Using pdf to compute probabilities

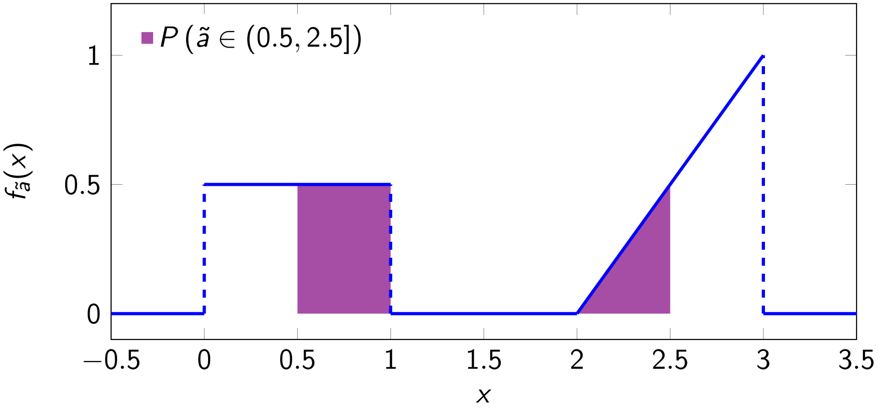
For an interval

$$\begin{aligned} \mathrm{P}\left(a < \tilde{a} \leq b\right) &= F_{\tilde{a}}\left(b\right) - F_{\tilde{a}}\left(a\right) \\ &= \int_{a}^{b} f_{\tilde{a}}\left(a\right) \, \mathrm{d}a \end{aligned}$$

For any countable union of disjoint intervals, $B = \cup_i \mathcal{I}_i$

$$\begin{aligned} \mathrm{P}\left(\tilde{a} \in B\right) &= \mathrm{P}\left(\tilde{a} \in \cup_i \mathcal{I}_i\right) \\ &= \sum_{i=1}^{n} \mathrm{P}\left(\tilde{a} \in \mathcal{I}_i\right) \\ &= \sum_{i=1}^{n} \int_{\mathcal{I}_i} f_{\tilde{a}}\left(a\right) \, \mathrm{d}a \\ &= \int_{B} f_{\tilde{a}}\left(a\right) \, \mathrm{d}a \end{aligned}$$

Using pdf to compute probabilities



Properties

Are pdfs always nonnegative?

Yes, because the cdf is nondecreasing

$$\int_{\mathbb{R}} f_{\tilde{a}}(a)\,\mathrm{d}a = \mathrm{P}\left(\tilde{a} \in \mathbb{R}\right) = 1$$

What functions are valid pdfs?

Any nonnegative function $f : \mathbb{R} \to \mathbb{R}$

$$\int_{\mathbb{R}} f(a)\, \mathrm{d}a = 1$$

can be interpreted as the pdf of a continuous random variable

We can reverse engineer the underlying probability space

What have we learned?

Definition and properties of probability density function