Functions of Continuous Random Variables

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

Plan

Explain how to derive the distribution of a function of a continuous random variable

Function of a discrete random variable

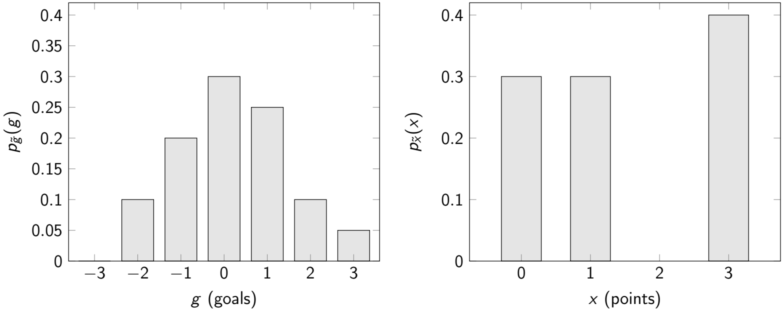
Let h be a deterministic function

If $\tilde{a}$ is a discrete random variable, is $\tilde{b} := h(\tilde{a})$ a discrete random variable?
Yes

How do we compute $p_{\tilde{b}}$ from $p_{\tilde{a}}$?

$$p_{\tilde{b}}(b) = \sum_{\{a \mid h(a)=b\}} p_{\tilde{a}}(a)$$

Converting goal difference to points



$$p_{\tilde{x}}(0) = 0.3 \qquad p_{\tilde{x}}(1) = 0.3 \qquad p_{\tilde{x}}(3) = 0.4$$

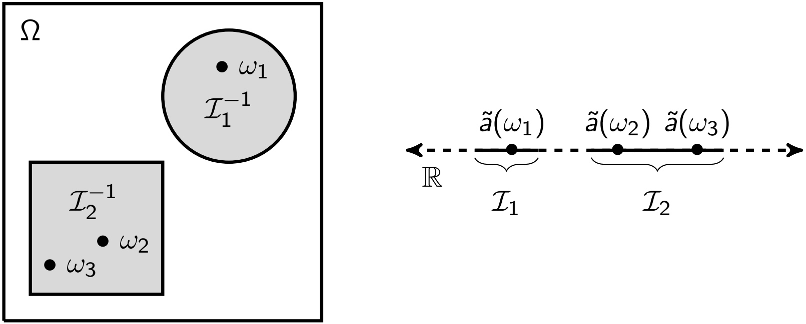
Function of a continuous random variable

Let h be a deterministic function

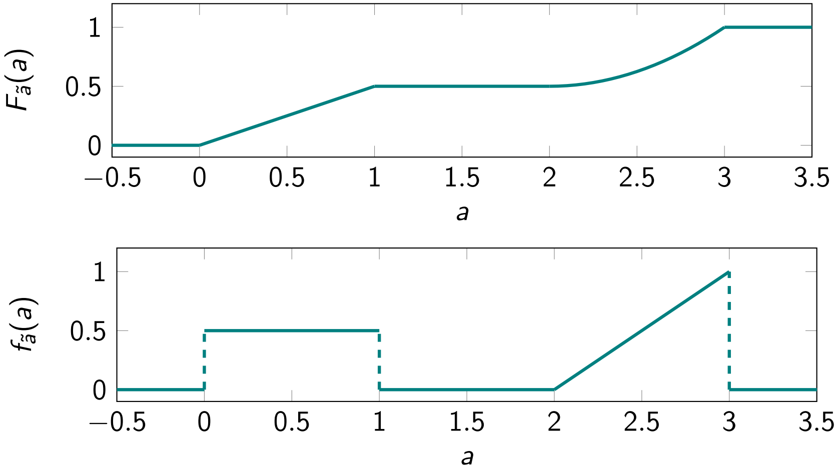
If $\tilde{a}$ is a continuous random variable, is $\tilde{b} := h(\tilde{a})$ a continuous random variable?

Not necessarily, but for most h it is

Continuous random variables



Cdf and pdf


$F_{\tilde{a}}(a)$
1
0.5
0
−0.5
0
0.5
1
1.5
2
2.5
3
3.5
a
$f_{\tilde{a}}(a)$
1
0.5
0
−0.5
0
0.5
1
1.5
2
2.5
3
3.5
a

Function of a random variable

h is a deterministic function

$\tilde{a}$ is a continuous random variable with pdf $f_{\tilde{a}}$

What is the cdf of $\tilde{b} := h(\tilde{a})$?

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

To compute the pdf $f_{\tilde{b}}$ we differentiate $F_{\tilde{b}}$

Current and voltage

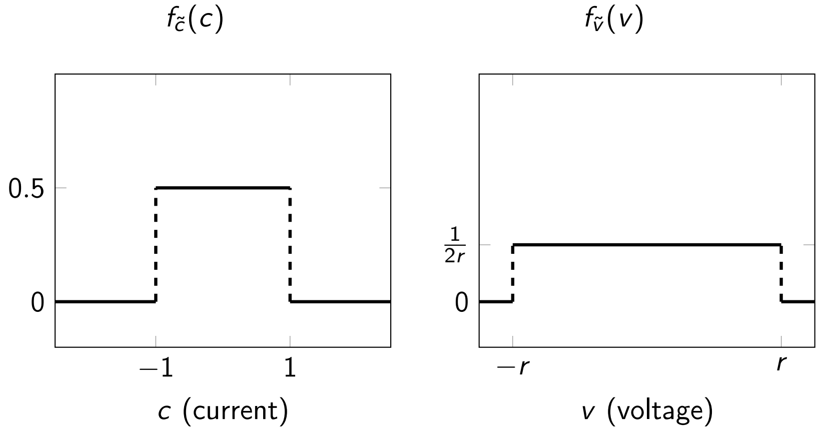
Current $\tilde{c}$ with pdf $f_{\tilde{c}}$

Voltage $\tilde{v} = r\tilde{c}$

$$\begin{aligned} F_{\tilde{v}}(v) &= \mathrm{P}(\tilde{v} \leq v) \\ &= \mathrm{P}(r\tilde{c} \leq v) \\ &= F_{\tilde{c}}\left(\frac{v}{r}\right) \end{aligned}$$

$$\begin{aligned} f_{\tilde{b}}(b) &= \frac{\mathrm{d}F_{\tilde{v}}(v)}{\mathrm{d}v} \\ &= \frac{1}{r} f_{\tilde{c}}\left(\frac{v}{r}\right) \end{aligned}$$

Current and voltage



Current and power

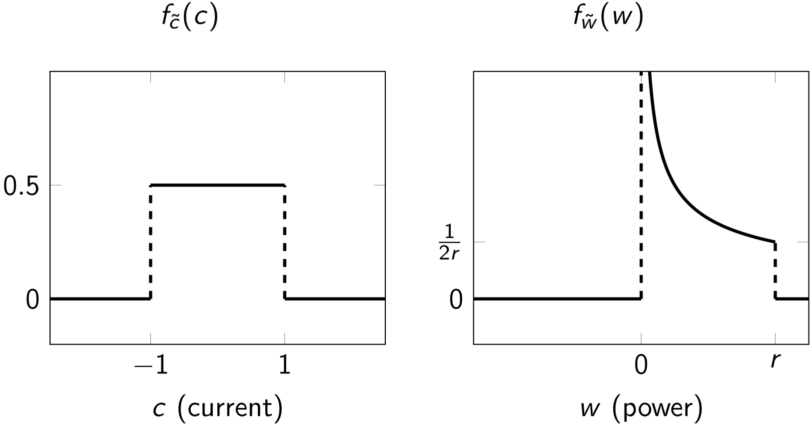
Current $\tilde{c}$ with pdf $f_{\tilde{c}}$

Power $\tilde{w} = r\tilde{c}^2$

$$\begin{aligned}
F_{\tilde{w}}(w) &= \mathrm{P}(\tilde{w} \leq w) \\
&= \mathrm{P}\left(r\tilde{c}^2 \leq w\right) \\
&= \mathrm{P}\left(-\sqrt{\frac{w}{r}} \leq \tilde{c} \leq \sqrt{\frac{w}{r}}\right) \\
&= F_{\tilde{c}}\left(\sqrt{\frac{w}{r}}\right) - F_{\tilde{c}}\left(-\sqrt{\frac{w}{r}}\right)
\end{aligned}$$

$$\begin{aligned}
f_{\tilde{w}}(w) &= \frac{\mathrm{d}}{\mathrm{d}w}\left(F_{\tilde{c}}\left(\sqrt{\frac{w}{r}}\right) - F_{\tilde{c}}\left(-\sqrt{\frac{w}{r}}\right)\right) \\
&= \frac{1}{2\sqrt{rw}}\left(f_{\tilde{c}}\left(\sqrt{\frac{w}{r}}\right) + f_{\tilde{c}}\left(-\sqrt{\frac{w}{r}}\right)\right) \qquad \text{if } w \geq 0
\end{aligned}$$

Current and power



Weird question

What happens if we feed a random variable $\tilde{a}$ into its own cdf?

What is the distribution of $\tilde{b} := F_{\tilde{a}}(\tilde{a})$?

This is crucial to understand p values!

Probability integral transform

We assume that $F_{\tilde{a}}$ is invertible

For $0 \leq b \leq 1$

$$\begin{aligned}
F_{\tilde{b}}(b) &= \mathrm{P}\left(\tilde{b} \leq b\right) \\
&= \mathrm{P}\left(F_{\tilde{a}}(\tilde{a}) \leq b\right) \\
&= \mathrm{P}\left(\tilde{a} \leq F_{\tilde{a}}^{-1}(b)\right) \\
&= F_{\tilde{a}}\left(F_{\tilde{a}}^{-1}(b)\right) \\
&= b
\end{aligned}$$

Uniform distribution in $[0, 1]$

What have we learned?

How to derive the distribution of functions of continuous random variables