Continuous Random Variables: pdf's

Engineering Statistics
Section 4.1

Josh Engwer

TTU

26 February 2016

Discrete & Continuous Random Variables (Definitions)

There are two types of random variables:

Definition

(Discrete Random Variable)

X is a **discrete random variable** $\iff$ Supp(X) is countable.

i.e. The meaningful values of X comprise a subset of integers $\mathbb{Z}$ or rationals $\mathbb{Q}$.

NOTE: Discrete random variables were explored in Chapter 3.

Definition

(Continuous Random Variable)

X is a **continuous random variable** $\iff$ Supp(X) is uncountable.

i.e. The meaningful values of X comprise an interval or union of intervals or $\mathbb{R}$.

Continuous random variables will now be covered here in Chapter 4.

Continuous r.v.'s & Their Supports (Examples)

- Experiment: Randomly choose a point on unit square centered at $(0,0)$.

X	$\equiv$	x-coordinate	$\Longrightarrow$	Supp(X)	$=$	$\left[-\frac{1}{2}, \frac{1}{2}\right]$
Y	$\equiv$	Magnitude of y-coord	$\Longrightarrow$	Supp(Y)	$=$	$\left[0, \frac{1}{2}\right]$
Z	$\equiv$	Distance to $(0,0)$	$\Longrightarrow$	Supp(Z)	$=$	$[0, 1/\sqrt{2}]$
W	$\equiv$	Distance to farthest corner	$\Longrightarrow$	Supp(W)	$=$	$[1/2, \sqrt{2}]$

- Experiment: Randomly select two adults in a busy airport.

X	$\equiv$	Height of 1^{st} adult (in ft)	$\Longrightarrow$	Supp(X)	$=$	$[1.75, 9.00]$
Y	$\equiv$	Height of 1^{st} adult (in cm)	$\Longrightarrow$	Supp(Y)	$=$	$[54, 275]$
Z	$\equiv$	Height of 1^{st} adult (in m)	$\Longrightarrow$	Supp(Z)	$=$	$[0.54, 2.75]$
W	$\equiv$	Height difference of the two adults (in ft)	$\Longrightarrow$	Supp(W)	$=$	$[-7.25, 7.25]$

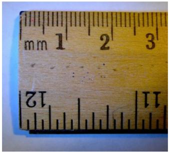
- Experiment: Take note of the time when a call center phone rings.

X	$\equiv$	Elapsed time until next phone call (in secs)	$\Longrightarrow$	Supp(X)	$=$	$(0, \infty)$
Y	$\equiv$	Elapsed time until next phone call (in hours)	$\Longrightarrow$	Supp(Y)	$=$	$(0, \infty)$

Measurements are Never 100% Accurate

A height is reported as, say, 251.6 cm instead of 251.5786800233791427 cm!
Also, there are a finite # of people, hence a finite # of heights to measure!

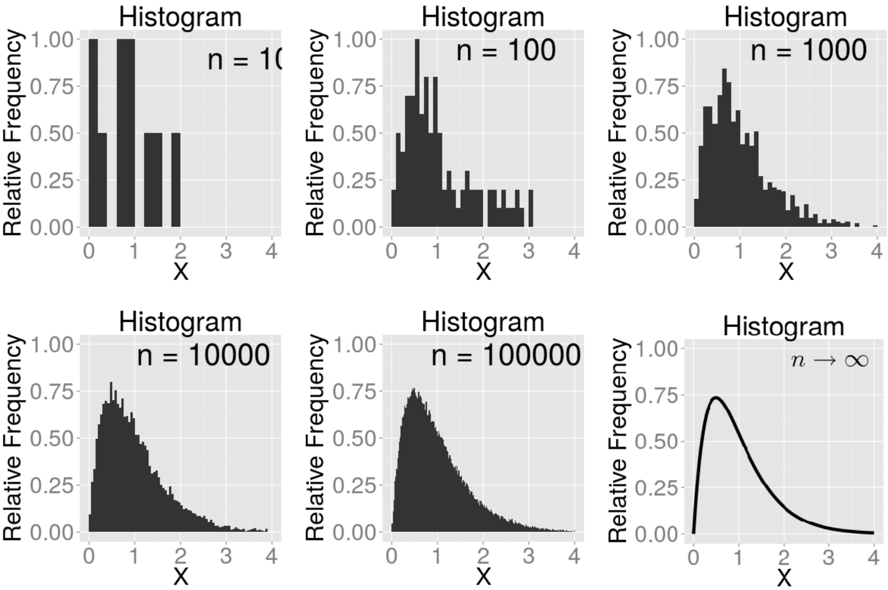
So why is a r.v. for a person's height considered continuous & not discrete??



Well, it turns out that in practice:

- Very large populations are well-approximated by continuous distributions.
- It's easier to mathematically work with continuous r.v.'s than discrete r.v.'s.

Very large pop's are well-approx'd by continuous dist's



Probability Density Function (pdf) of a Continuous r.v.

Definition

(Probability Density Function of a Continuous Random Variable)

Let X be a **continuous** random variable.
Then, its **probability density function (pdf)**, which is denoted $f_X(x)$, is defined to be a function of the possible values of X such that:

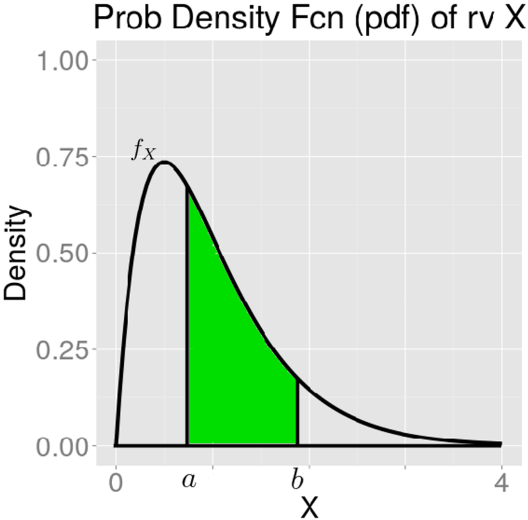
Non-negativity on its Support: $f_X(x) \geq 0 \quad \forall x \in \text{Supp}(X)$

Zero outside of its Support: $f_X(x) = 0 \quad \forall x \notin \text{Supp}(X)$

Universal Integral of Unity: $$\int_{\text{Supp}(X)} f_X(x)\, dx = 1$$

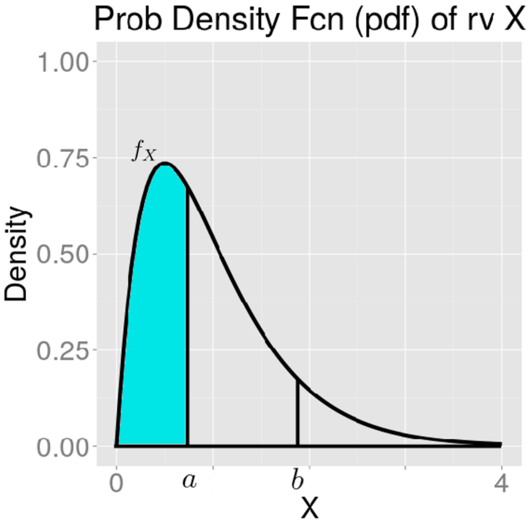
The graph of $f_X(x)$ is often called the **density curve** for random variable X.

Probability of a Continuous r.v. via its pdf



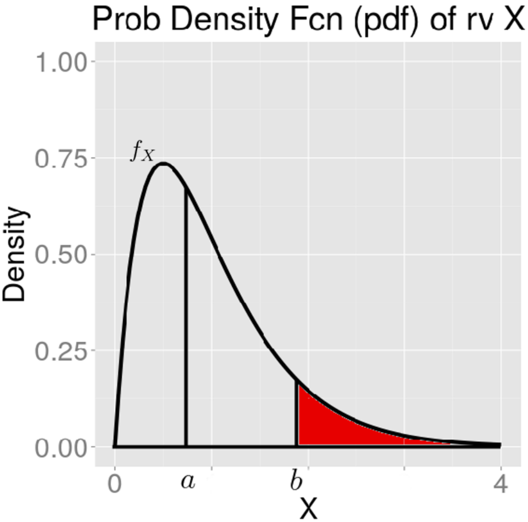
$$\mathbb{P}(a \leq X \leq b) = \int_a^b f_X(x)\, dx = (\text{Green area under the curve})$$

Probability of a Continuous r.v. via its pdf



$$\mathbb{P}(X \le a) = \int_{-\infty}^{a} f_X(x)\, dx = \text{(Blue area under the curve)}$$

Probability of a Continuous r.v. via its pdf



$$\mathbb{P}(X \geq b) = \int_b^\infty f_X(x)\, dx = (\text{Red area under the curve})$$

Probability of a Continuous r.v. (Properties)

Theorem

*Let X be a **continuous** r.v. with pdf $f_X(x)$. Let scalars $a < b$. Then:*

$$\int_{-\infty}^{\infty} f_X(x)\,dx = 1$$

$$\mathbb{P}(X = a) = 0$$

$$\begin{aligned}
\mathbb{P}(a < X < b) &= \mathbb{P}(a \le X \le b) \\
\mathbb{P}(a < X \le b) &= \mathbb{P}(a \le X \le b) \\
\mathbb{P}(a \le X < b) &= \mathbb{P}(a \le X \le b)
\end{aligned}$$

$$\begin{aligned}
\mathbb{P}(X > b) &= \mathbb{P}(X \ge b) \\
\mathbb{P}(X < a) &= \mathbb{P}(X \le a)
\end{aligned}$$

PROOF: $\int_{-\infty}^{\infty} f_X(x)\,dx = \int_{\text{Supp}(X)} f_X(x)\,dx + \int_{[\text{Supp}(X)]^c} f_X(x)\,dx = 1 + 0 = 1$

Probability of a Continuous r.v. (Properties)

Theorem

*Let X be a **continuous** r.v. with pdf $f_X(x)$. Let scalars $a < b$. Then:*

$$\int_{-\infty}^{\infty} f_X(x)\,dx = 1$$

$$\mathbb{P}(X = a) = 0$$

$$\begin{aligned}
\mathbb{P}(a < X < b) &= \mathbb{P}(a \le X \le b) \\
\mathbb{P}(a < X \le b) &= \mathbb{P}(a \le X \le b) \\
\mathbb{P}(a \le X < b) &= \mathbb{P}(a \le X \le b)
\end{aligned}$$

$$\begin{aligned}
\mathbb{P}(X > b) &= \mathbb{P}(X \ge b) \\
\mathbb{P}(X < a) &= \mathbb{P}(X \le a)
\end{aligned}$$

PROOF: $\mathbb{P}(X = a) = \int_a^a f_X(x)\,dx = \Big[F_X(x)\Big]_{x=a}^{x=a} \stackrel{FTC}{=} F_X(a) - F_X(a) = 0$

$F_X(x)$ is the anti-derivative of $f_X(x)$ and will be explored in the next section.

Probability of a Continuous r.v. (Properties)

Theorem

*Let X be a **continuous** r.v. with pdf $f_X(x)$. Let scalars $a < b$. Then:*

$$\int_{-\infty}^{\infty} f_X(x)\,dx = 1$$

$$\mathbb{P}(X = a) = 0$$

$$\begin{aligned}
\mathbb{P}(a < X < b) &= \mathbb{P}(a \leq X \leq b) \\
\mathbb{P}(a < X \leq b) &= \mathbb{P}(a \leq X \leq b) \\
\mathbb{P}(a \leq X < b) &= \mathbb{P}(a \leq X \leq b)
\end{aligned}$$

$$\begin{aligned}
\mathbb{P}(X > b) &= \mathbb{P}(X \geq b) \\
\mathbb{P}(X < a) &= \mathbb{P}(X \leq a)
\end{aligned}$$

PROOF:

$$\mathbb{P}(X < a) = \mathbb{P}(X \leq a) - \mathbb{P}(X = a) = \mathbb{P}(X \leq a) - 0 = \mathbb{P}(X \leq a)$$

Probability of a Continuous r.v. (Properties)

Theorem

*Let X be a **continuous** r.v. with pdf $f_X(x)$. Let scalars $a < b$. Then:*

$$\int_{-\infty}^{\infty} f_X(x)\,dx = 1$$

$$\mathbb{P}(X = a) = 0$$

$$\begin{aligned}
\mathbb{P}(a < X < b) &= \mathbb{P}(a \le X \le b) \\
\mathbb{P}(a < X \le b) &= \mathbb{P}(a \le X \le b) \\
\mathbb{P}(a \le X < b) &= \mathbb{P}(a \le X \le b)
\end{aligned}$$

$$\begin{aligned}
\mathbb{P}(X > b) &= \mathbb{P}(X \ge b) \\
\mathbb{P}(X < a) &= \mathbb{P}(X \le a)
\end{aligned}$$

PROOF: The rest follow in a similar fashion. QED

It's easier to mathematically work with continuous r.v.'s

Consider the **discrete** r.v. X with the following **pmf** & support:

$$p_X(k) = \frac{6}{\pi^2 k^2}, \quad \text{Supp}(X) = \{1, 2, 3, 4, \cdots\}$$

$$\mathbb{P}(X > 2) = \sum_{k=3}^{\infty} p_X(k) = \frac{6}{\pi^2} \sum_{k=3}^{\infty} \frac{1}{k^2} = \left(\begin{array}{c} \text{Requires Fourier Analysis} \\ \text{to sum!! (Diff Eqn II)} \end{array} \right)$$

Consider the **continuous** r.v. X with the following **pdf** & support:

$$f_X(x) = \frac{1}{x^2}, \quad \text{Supp}(X) = [1, \infty)$$

$$\begin{aligned} \mathbb{P}(X > 2) &= \int_2^{\infty} f_X(x)\, dx = \int_2^{\infty} \frac{1}{x^2}\, dx = \left[-\frac{1}{x} \right]_{x=2}^{x \to \infty} \\ &\overset{FTC}{=} \lim_{x \to \infty} \left(-\frac{1}{x} \right) - \left(-\frac{1}{(2)} \right) = 0 + \frac{1}{2} = \boxed{\frac{1}{2}} \end{aligned}$$

Nonelementary Integrals

Of course from CalcII some integrals cannot be exactly computed by hand:

Definition

A **nonelementary integral** is an integral whose antiderivative cannot be expressed in a finite closed form.

Here's a small list of nonelementary integrals (there are many, many more):

$$\int e^{x^2}\,dx \qquad \int \frac{e^x}{x}\,dx \qquad \int \sqrt{x}e^{-x}\,dx$$

$$\int \sin(x^2)\,dx \qquad \int \cos(e^x)\,dx \qquad \int e^{\cos x}\,dx$$

$$\int \sqrt{1+x^4}\,dx \qquad \int \ln(\ln x)\,dx \qquad \int \frac{x}{e^x-1}\,dx$$

$$\int \frac{1}{\ln x}\,dx \qquad \int \frac{\sin x}{x}\,dx \qquad \int \sin(\sin x)\,dx$$

$$\int x^x\,dx \qquad \int \frac{1}{x^x}\,dx \qquad \int \arctan(\ln x)\,dx$$

In statistics, most pdf's result in elementary integrals that can be computed. For the few pdf's that lead to nonelementary integrals, a table will be provided.

Textbook Logistics for Section 4.1

- Difference(s) in Notation:

CONCEPT	TEXTBOOK NOTATION	SLIDES/OUTLINE NOTATION
Probability of Event	$P(A)$	$\mathbb{P}(A)$
Support of a r.v.	"All possible values of X"	Supp(X)
pdf of a r.v.	$f(x)$	$f_X(x)$

- Skip **uniform random variables** (pg 144)
  - Uniform random variables are a special type of continuous rv.
  - Uniform random variables will be formally covered in Section 4.3

Fin

Fin.