3.3 Continuous Random Variables

Recall A discrete random variable is a function Y: sample space $\to \mathbb{R}$ that can take on at most countably infinitely many values. A continuous random variable, on the other hand, can take on uncountably many values (usually in an interval or union of intervals), e.g. physical characteristics (time, length, position, the duration of a randomly selected phone call). Examples of continuous data are in Section 3.1.

A continuous random variable is required to have a continuous c.d.f.:

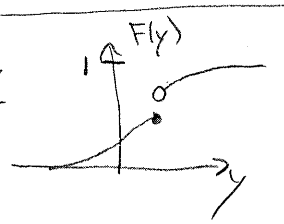
Def A random variable $Y: S \to \mathbb{R}$ on a sample space S is a continuous random variable if its c.d.f. $F: \mathbb{R} \to \mathbb{R}$, $F(y) := P(Y \le y)$ is a continuous function, that is, for every $y_0 \in \mathbb{R}$, we have $\lim_{y \to y_0} F(y) = F(y_0)$.

Rem We do not require the random variable to take on uncountably many values. That is a consequence of the requirement that the c.d.f. is continuous.

<u>Examples</u> of continuous random variables that we will discuss are

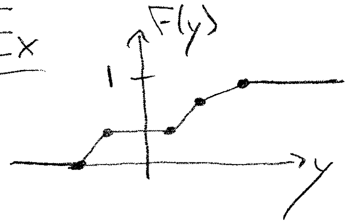
- uniform random variables
- exponential random variables
- gamma random variables
- chi-square random variables
- normal random variables, and
- beta random variables.

<u>Ex</u>



$\Rightarrow Y$ is <u>not</u> a continuous random variable

<u>Ex</u>



$\Rightarrow Y$ is a continuous random variable

<u>Rem</u> Most continuous random variables of importance in probability theory are even "absolutely continuous", which means there exists a non-negative integrable function $f:\mathbb{R}\to[0,\infty)$ such that $F(y)=\int_{-\infty}^{y} f(t)\,dt$.

In this case, f is called a <u>probability density function</u> or <u>p.d.f.</u> for the random variable. Its central importance is that $P(a\le Y\le b)=\int_a^b f(t)\,dt$.

We will later see sufficient conditions on F that guarantee the existence of a p.d.f. and apply to all our main examples.

We first recall the properties of c.d.f.'s (for both discrete and continuous random variables).

<u>Thm</u> (Properties of any C.D.F.) Let $F: \mathbb{R} \to \mathbb{R}$ be the c.d.f. of a random variable (not necessarily continuous). Then:

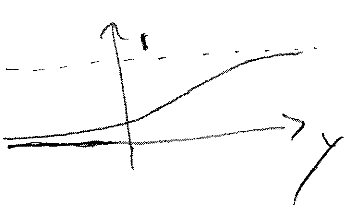
1) F is weakly increasing, i.e. if $y_1 < y_2$ then $F(y_1) \le F(y_2)$.

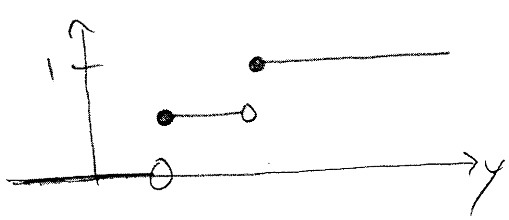
2) $\lim_{y \to -\infty} F(y) = 0$

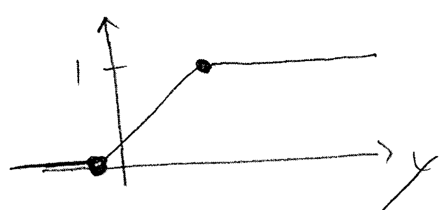
3) $\lim_{y \to \infty} F(y) = 1$

4) F is right continuous, i.e. for every $y_0 \in \mathbb{R}$
$$\lim_{y \to y_0^+} F(y) = F(y_0).$$

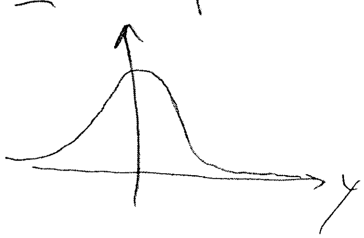
<u>Ex</u> Graphs of possible c.d.f.'s are

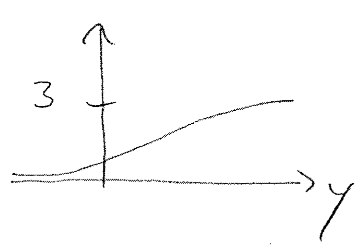


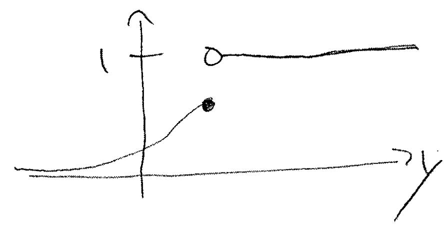




<u>Ex</u> Graphs of functions which cannot be c.d.f.'s are







Recall the Fundamental Theorem of Calculus Part I.

Thm If f is continuous on $[a,b]$, then the function $G:[a,b]\to\mathbb{R}$ defined by

$$G(x) := \int_a^x f(t)\,dt$$

is continuous on $[a,b]$ and differentiable on (a,b).

Moreover, $\frac{d}{dx}G(x) = \frac{d}{dx}\int_a^x f(t)\,dt = f(x)$.

We now give sufficient conditions on a c.d.f. to guarantee the existence of a p.d.f.:

Prop Let X be a continuous random variable and $F:\mathbb{R}\to\mathbb{R}$ its c.d.f. Suppose:

(i) F' exists everywhere except at finitely many places.

(ii) F' has at most finitely many discontinuities, and all are jump discontinuities.

Let

$$f(x) := \begin{cases} F'(x) & \text{if } F'(x) \text{ exists} \\ 0 & \text{if } F'(x) \text{ does not exist and } x \notin \text{image}(X) \\ 1 & \text{if } F'(x) \text{ does not exist and } x \in \text{image}(X) \end{cases}$$

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

In other words, f is a p.d.f. for X.

<u>Pr</u> Let $a_1, a_2, \ldots, a_n$ be the discontinuities of F' listed in increasing order (among the a_i's we include the inputs where F' is not defined).

First we verify $F(x) = \int_{-\infty}^{x} f(t)\,dt$ for $x \in (-\infty, a_1)$.

For $x \in (-\infty, a_1)$ we have

$$\begin{aligned}
\int_{-\infty}^{x} f(t)\,dt &= \lim_{M \to -\infty} \int_{M}^{x} f(t)\,dt \\
&= \lim_{M \to -\infty} \int_{M}^{x} F'(t)\,dt \\
&= \lim_{M \to -\infty} \left(F(x) - F(M)\right) \quad \text{by FTC2} \\
&= F(x) - \underbrace{\lim_{M \to -\infty} F(M)}_{=0 \text{ because } F \text{ is a c.d.f.}} \\
&= F(x) \quad \checkmark
\end{aligned}$$

Next we verify $F(a_1) = \int_{-\infty}^{a_1} f(t)\,dt$ (case $x = a_1$)

$$\int_{-\infty}^{a_1} f(t)\,dt = \underbrace{\int_{-\infty}^{a_1-1} f(t)\,dt}_{\substack{\shortparallel \\ F(a_1-1) \text{ by above}}} + \int_{a_1-1}^{a_1} f(t)\,dt$$

$$= F(a_1-1) + \lim_{N\to a_1^-} \int_{a_1-1}^{N} f(t)\,dt$$

$$= F(a_1-1) + \lim_{N\to a_1^-} \int_{a_1-1}^{N} F'(t)\,dt$$

$$= F(a_1-1) + \lim_{N\to a_1^-} \left(F(N) - F(a_1-1)\right)$$

$$= F(a_1-1) - F(a_1-1) + \underbrace{\lim_{N\to a_1^-} F(N)}_{=F(a_1) \text{ because } F \text{ is continuous.}}$$

$$= F(a_1) \checkmark$$

So far, we have verified that $F(x) = \int_{-\infty}^{x} f(t)\,dt$ for $x \in (-\infty, a_1]$. Next we show this equality also holds true for $x \in (a_1, a_2)$. Let $x \in (a_1, a_2)$. Then

$$\int_{-\infty}^{x} f(t)\,dt = \int_{-\infty}^{a_1} f(t)\,dt + \int_{a_1}^{x} f(t)\,dt = F(a_1) + \lim_{M\to a_1^+} \int_{M}^{x} f(t)\,dt$$

$$= F(a_1) + \lim_{M\to a_1^+} \left(F(x) - F(M)\right) \quad \text{by FTC 2.}$$

$$= F(a_1) + F(x) - F(a_1) \quad \text{because } F \text{ is continuous.}$$

$$= F(x) \checkmark$$

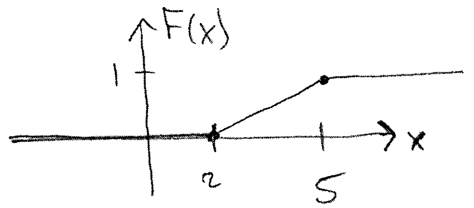
Thus, we now have $F(x) = \int_{-\infty}^{x} f(t)\,dt$ for all $x \in (-\infty, a_2)$.

Now the argument above for a_1 can be repeated for $a_2, a_3, \ldots,$ and a_n.

Finally we conclude $F(x) = \int_{-\infty}^{x} f(t)\,dt$ for all $x \in (-\infty, \infty)$.

$\square$

Ex (of Prop) Suppose a random variable X has image $(2, 5]$ and suppose its c.d.f. $F : \mathbb{R} \to \mathbb{R}$ is

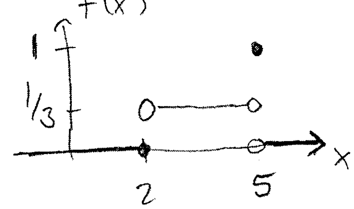


. Then F is continuously differentiable everywhere except at $x = 2$ and at $x = 5$.

$$F'(x) = \begin{cases} 0 & \text{for } x \in (-\infty, 2) \\ \frac{1}{3} & \text{for } x \in (2, 5) \\ 0 & \text{for } x \in (5, \infty) \end{cases}$$ by looking at the graph.

So the Proposition says to define

$$f(x) = \begin{cases} 0 & \text{for } x \in (-\infty, 2) \\ 0 & \text{for } x = 2 \\ \frac{1}{3} & \text{for } x \in (2, 5) \\ 1 & \text{for } x = 5 \\ 0 & \text{for } x \in (5, \infty) \end{cases}$$



Prop! and look at picture.

Then we have $F(x) = \int_{-\infty}^{x} f(t)\,dt$ and $F'(x) = f(x)$ where F' is defined.

Rem Why did we set $f(2)=0$ but $f(5)=1$?

Answer: because the (unclosed) support of the density should be the image of the random variable:

$$\text{support}(f) := \{x \in \text{dom } f \mid f(x) \neq 0\}$$

The support of a random variable X is by definition its image. Thus we see that the support of a random variable is equal to the support of its density or mass function.

Note: In some parts of math, the support of a function means the closure of

$$\{x \in \text{dom } f \mid f(x) \neq 0\},$$

unlike in probability theory.

In the Proposition 1 we saw, given a continuous random variable whose c.d.f. $F: \mathbb{R} \to \mathbb{R}$ has a continuous first derivative except at finitely many places with jump discontinuities, how to construct a density using $f(x) := F'(x)$ where F' is defined. But is the density necessarily F'? Yes, except at the points of discontinuity:

<u>Prop 2</u> Let $X: S \to \mathbb{R}$ be a continuous random variable and $F: \mathbb{R} \to \mathbb{R}$ its c.d.f. Suppose X has a density function that is continuous except at finitely many jump discontinuities, i.e. suppose there exists a function $f: \mathbb{R} \to \mathbb{R}$ such that

(i) $\int_{-\infty}^{x} f(t)\,dt = F(x)$

(ii) f is continuous except at finitely many jump discontinuities.

Then for every $x_0 \in \mathbb{R}$ where f is continuous we have F is differentiable at x_0 and $F'(x_0) = f(x_0)$.

<u>Pr</u> Suppose the hypotheses and suppose f is continuous at x_0. Since f is continuous at x_0 and since f has only finitely many discontinuities, there exists a small enough $\varepsilon > 0$ such that f is continuous on the closed interval $[x_0 - \varepsilon, x_0 + \varepsilon]$.

Let $x \in (x_0 - \varepsilon, x_0 + \varepsilon)$. Then

$$F(x) = \int_{-\infty}^{x} f(t)\,dt \qquad \text{(by assumption)}$$

$$= \int_{-\infty}^{x_0-\varepsilon} f(t)\,dt + \int_{x_0-\varepsilon}^{x} f(t)\,dt$$

$$= F(x_0-\varepsilon) + \int_{x_0-\varepsilon}^{x} f(t)\,dt \quad \Big\} \text{ Thus } F(x) \text{ is differentiable for } x \in (x_0-\varepsilon, x_0+\varepsilon) \text{ by FTC1}$$

$$\Rightarrow \quad \frac{d}{dx}F(x) = \frac{d}{dx}\left(F(x_0-\varepsilon) + \int_{x_0-\varepsilon}^{x} f(t)\,dt\right)$$

$$\frac{d}{dx}F(x) = 0 + \frac{d}{dx}\int_{x_0-\varepsilon}^{x} f(t)\,dt$$

$$\frac{d}{dx}F(x) = f(x) \qquad \text{by FTC1.}$$

Now take $x = x_0$ to get $F'(x_0) = f(x_0)$.

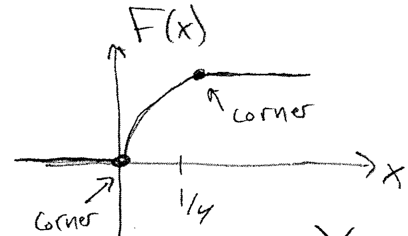
□

<u>Rem</u> Propositions 1 and 2 can be generalized to allow other kinds of discontinuities of F' and many more discontinuities, in fact countably infinitely many which are spread out enough to only have finitely many on any

finite length interval.

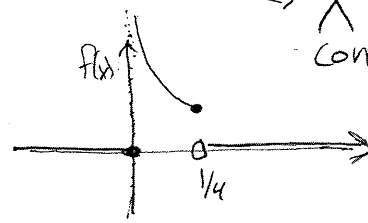
The following example shows another kind of discontinuity of F' for which the conclusion of Proposition 1 holds.

<u>Ex</u> $F(x) = \begin{cases} 0 & \text{for } x \in (-\infty, 0] \\ 2x^{1/2} & \text{for } x \in (0, 1/4] \\ 1 & \text{for } x \in (1/4, \infty) \end{cases}$



$\Rightarrow X$ is a continuous r.v.

$f(x) = \begin{cases} 0 & \text{for } x \in (-\infty, 0] \\ x^{-1/2} & \text{for } x \in (0, 1/4] \\ 0 & \text{for } x \in (1/4, \infty) \end{cases}$



$(0, 1/4] = \text{support}(f) = \text{support}(X) = \text{image}(X)$

f has a vertical asymptote at $x=0$. !

But $F(x) = \int_{-\infty}^{x} \frac{1}{t^{1/2}}\,dt$ is still continuous at $x=0$!

This density function is unbounded, unlike a mass function, which takes values in $[0,1]$.

<u>Rem</u> Density functions of continuous random variables have two key properties.

<u>Thm</u> (Properties of Densities of Continuous R.V.'s).

If $f(x)$ is a probability density function (p.d.f.) of a continuous random variable, then

(i) $f(x) \geq 0$ for all $x \in \mathbb{R}$.

(ii) $\int_{-\infty}^{\infty} f(t)\,dt = 1$

<u>Pr</u> (i) F is weakly increasing, so $f = F' \geq 0$ where defined.

(ii) $\int_{-\infty}^{\infty} f(t)\,dt = \lim_{N\to\infty} \int_{-\infty}^{N} f(t)\,dt \underset{\text{def of density}}{=} \lim_{N\to\infty} F(N) \underset{\text{property of c.d.f.}}{=} 1$ $\square$