## Cumulative Distribution Function

- The cumulative (probability) distribution function or CDF of a random variable **X** is denoted by F(u) or $F_X(u)$
- Definition: $F_X(u) = P\{\mathbf{X} \le u\}$, $-\infty < u < \infty$
- The CDF is a real-valued function defined for all real number values of its argument u
- Cumulative because the value of the CDF at u is the total probability mass from $-\infty$ up to (and including) the point u




## CDF of a discrete random variable

- The CDF of any discrete random variable is a staircase function
- If **X** takes on values $u_1, u_2, \ldots u_n, \ldots$ with probabilities $p(u_1), p(u_2), \ldots, p(u_n), \ldots$ then the CDF has
  - jumps whose heights are $p(u_1), p(u_2), \ldots, p(u_n), \ldots$
  - at locations $u_1, u_2, \ldots u_n, \ldots$ respectively

  and is flat in between the jumps




## CDF of a discrete random variable

- The CDF of any discrete random variable is a staircase function
- We have defined the $u_i$ to be an increasing sequence $u_1 < u_2 < \ldots < u_n < \ldots$
- For all u in the range $u_i \le u < u_{i+1}$, $F_X(u)$ has value

$$\sum_{j=1}^{i} p(u_j)$$

- $F_X(u)$ jumps in value by $p(u_{i+1})$ at $u = u_{i+1}$




## Negative thoughts about the CDF

- The CDF of a discrete random variable **X** is far more cumbersome than the pmf
- The CDF contains no information that we could not have deduced from the pmf
- So, why bother with the CDF at all?
- The concept of the CDF of **X** is a very general one that applies to every random variable, discrete or continuous or mixed…




## The power of positive thinking

- The concept of the CDF of **X** is a very general one that applies to every random variable, discrete or continuous or mixed…
- Every random variable **X** has a CDF that is defined in exactly the same way:

$$F_X(u) = P\{\mathbf{X} \le u\}, \quad -\infty < u < \infty$$

- But, discrete random variables have pmfs, continuous ones have pdfs, etc …




## More power of positive thinking

- Common properties of random variables, i.e. those shared by all random variables, can be stated in terms of the CDF instead of having a different formula for each type of random variable
- The CDF is a unifying concept for all kinds of random variables

## A more balanced viewpoint

- Using the general formulations of random variable properties via the CDF is often a messy job — use the pmf or pdf instead
- Using the CDF as an intermediate step is very helpful in some types of calculations, and often simplifies these calculations
- The CDF should be used sparingly and for stating general concepts only



## CDF properties

- For any random variable **X**,

  $F_X(u) = P\{X \le u\}, \quad -\infty < u < \infty$
- $0 \le F_X(u) \le 1$
- If $a < b$, then $F_X(a) \le F_X(b)$
- $F_X(-\infty) = 0$ and $F_X(+\infty) = 1$
- If CDF is discontinuous at t, $F_X(t^+) > F_X(t^-)$, and $F_X(t)$ equals $F_X(t^+)$, the limiting value of $F_X(u)$ as $u \to t$ from the right



## CDF properties (continued)

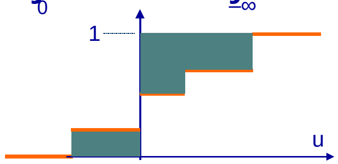
- If $F_X(u)$ is discontinuous at u = t, then $F_X(t^+) - F_X(t^-)$ = jump in the value of $F_X(u)$ $= P\{X = t\}$
- If $F_X(u)$ is continuous at u = v, then $F_X(v) = F_X(v^+) = F_X(v^-)$ $\quad P\{X = v\} = 0$
- $P\{X \le u\} = F_X(u) = F_X(u^+)$
- $P\{X > u\} = 1 - F_X(u)$
- $P\{X < u\} = F_X(u^-)$
- $P\{X \ge u\} = 1 - F_X(u^-)$



## Expectation of a discrete RV

- E[**X**] = area above CDF to right of 0 – area below CDF to left of 0

$$= \int_0^\infty [1 - F_X(u)]\,du - \int_{-\infty}^0 F_X(u)\,du$$



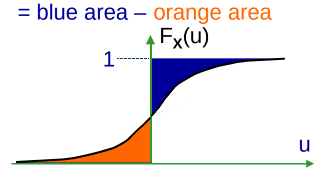



## Expectation of an arbitrary RV

- The expected value E[**X**] of an arbitrary random variable **X** can be defined as

$$E[X] = \int_0^\infty [1 - F_X(u)]\,du - \int_{-\infty}^0 F_X(u)\,du$$

= blue area – orange area






## Great expectations? Infinite ones?

- The expected value E[**X**] of an arbitrary random variable **X** can be defined as

$$E[X] = \int_0^\infty [1 - F_X(u)]\,du - \int_{-\infty}^0 F_X(u)\,du$$

- If both integrals are finite, then E[**X**] is a finite number
- If exactly one integral is infinite, some authors say that E[**X**] is $+\infty$ or $-\infty$
- If both integrals are infinite, then all the authors agree that E[**X**] is undefined

## Keeping things together …

- The expected value E[**X**] of an arbitrary random variable **X** can be defined as
  $$E[\mathbf{X}] = \int_0^{\infty} [1 - F_X(u)]\,du \;-\; \int_{-\infty}^{0} F_X(u)\,du$$
- Never, ever, break up the first integral into two parts as the integral of 1 from 0 to ∞ minus the integral of $F_X(u)$ from 0 to ∞
- Always think of the integrand in the first integral as P{**X** > u}



## Continuous CDFs

- The CDF of a discrete random variable **X** is a staircase function with jumps at each value $u_1, u_2, \ldots u_n, \ldots$ taken on by **X**
- The CDF is not continuous at $u_1, u_2, \ldots u_n$, but is continuous in between these jumps
- For the next several lectures, we shall only consider CDFs that are continuous at all values of u, $-\infty < u < \infty$



## Continuous CDFs (continued)

- Supose that $F_X(u)$ is continuous at all u, $-\infty < u < \infty$
- We no longer need to fuss about $F_X(u^+)$ versus $F_X(u^-)$:
  $F_X(u^+) = F_X(u^-)$ for all values of u
- This implies that
  $F_X(u^+) - F_X(u^-) = P\{\mathbf{X} = u\} = 0$ for all u
- Every event {**X** = u} has zero probability!



## All the probability disappeared?

- Supose that $F_X(u)$ is continuous at all u, $-\infty < u < \infty$
- Every event {**X** = u} has zero probability
- No, the nonzero probabilities are assigned to the intervals of the line, not to individual real numbers
- For a more detailed discussion of this point, see Lecture 4 (Slides 33-41) in this series of lectures



## All the probability is still there!

- Supose that $F_X(u)$ is continuous at all u, $-\infty < u < \infty$
- Every event {**X** = u} has zero probability
- Events of the form {**X** ≤ u} have nonzero value $F_X(u)$ for many choices of u
- For a < b, events of the form {a ≤ **X** ≤ b} have probability $F_X(b) - F_X(a) \geq 0$
- Events of the form {a ≤ **X** ≤ a + δa} where δa is very small can have nonzero prob.



## Meaningless questions

- Supose that $F_X(u)$ is continuous at all u, $-\infty < u < \infty$
- Every event {**X** = u} has zero probability
- It is meaningless to ask for the probability of the event {**X** = 3.14159628…}
- Event {**X** = u} is physically unobservable
- At best, a measurement will tell us that the event {u – δu ≤ **X** ≤ u + δu} occurred
- This event we can talk about!

## An exercise

- Supose that $F_X(u)$ is continuous at all u, $-\infty < u < \infty$
- Exercise: Let s denote any number in the range $0 < s < 1$. Show that for any continuous CDF $F_X(u)$, there must be at least one number t such that $F_X(t) = s$
- Can there be many t such that $F_X(t) = s$? What can be said about the set of numbers: $\{t: F_X(t) = s\}$?

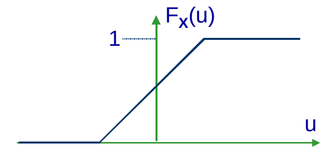
## Continuous random variables

- Definition: A continuous random variable **X** is one whose CDF $F_X(u)$ is
  - continuous at all u, $-\infty < u < \infty$
  - differentiable at all u (except possibly at a set of points $u_1 < u_2 < \ldots u_n < \ldots$)
- More precisely, any finite-length interval contains at most finitely many points where $F_X(u)$ is not differentiable

## Continuous but not differentiable?

- Do there exist continuous functions that are not differentiable?
- Aren't all continuous functions always differentiable?



## Continuous RV ≠ discrete RV

- Definition: A continuous random variable **X** is one whose CDF $F_X(u)$ is
  - continuous at all u, $-\infty < u < \infty$
  - differentiable at all u (except possibly at a set of points $u_1 < u_2 < \ldots u_n < \ldots$)
- Note that the staircase CDF of a discrete random variable satisfies the second condition but not the first

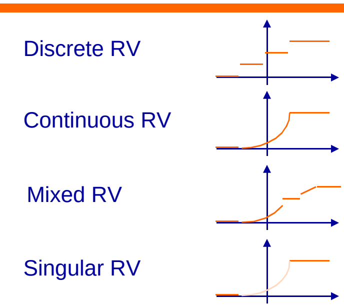
  

## Four kinds of random variables

- Discrete RVs have staircase CDFs
- Continuous RVs have continuous CDFs that are differentiable for almost all u
- Mixed RVs have piecewise differentiable CDFs (with positive slopes) and jump dscontinuities (see e.g. Fig. 4.1 of Ross)
- Singular RVs have continuous CDFs that are non-differentiable almost everywhere

## Diagrams of CDFs



Department of Electrical and Computer Engineering
University of Illinois at Urbana-Champaign 

## A sigh of relief…

- In this course, we consider only discrete and continuous random variables with a very occasional look at mixed random variables
- We do not consider singular random variables at all



## Don't drink and derive…

- Definition: A continuous random variable **X** is one whose CDF $F_X(u)$ is
  - continuous at all u, $-\infty < u < \infty$
  - differentiable at all u (except possibly at a set of points $u_1 < u_2 < \ldots u_n < \ldots$)
- Rhetorical question: What's the point of having a differentiable function if you don't differentiate it?



## The derivative of the CDF

- The derivative of the CDF of a continuous random variable **X** exists for almost all real numbers u
- Definition: The probability density function (pdf) of a continuous random variable **X** is

$$f_X(u) = \begin{cases} \frac{d}{du}F_X(u), & \text{if } F_X(u) \text{ is differentiable} \\ \text{any number} \geq 0 & \text{if CDF is non-diff.} \end{cases}$$



## Stupidest definition you've seen?

- Definition: The probability density function (pdf) of a continuous random variable **X** is

$$f_X(u) = \begin{cases} \frac{d}{du}F_X(u), & \text{if } F_X(u) \text{ is differentiable} \\ \text{any number} \geq 0 & \text{if CDF is non-diff.} \end{cases}$$

- What kind of definition is this where we can set $f_X(u)$ equal to any nonnegative number that we choose?

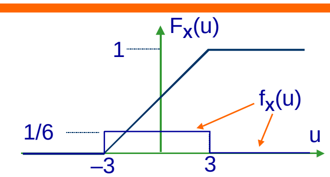


## It a'int as bad as it sounds. Honest!

- We are allowed to set the value of $f_X(u)$ to any nonnegative number only at those few isolated points where the CDF is not differentiable
- Furthermore, the arbitrarily chosen value assigned to the pdf at these isolated points makes no difference whatsoever in any probability calculations

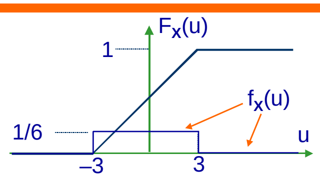


## What should we choose?




- The pdf is as shown by the blue line between –3 and 3 and for u > 3, u < –3
- The derivative of the CDF is undefined at u = ±3

## Possible choices for $f_X(\pm 3)$




- $f_X(u) = 1/6$ for $-3 < u < 3$
- $f_X(u) = 0$ for $|u| > 3$
- $f_X(\pm 3) = 1/6$ or 0 = left (or right) derivative or your SSN as a nine-digit decimal, or …



## What values can X take on?

- For every real number u, P{**X** = u} = 0
- But, what is the set of all possible values that **X** can take on?
- Simple answer: **X** can take on value 3 (say) if and only if $f_X(3) > 0$
- If the CDF is non-differentiable at 3, and 3 is one of the values taken on by **X**, then we assign a positive value to $f_X(3)$



## Convention in this course

- The following convention is used in this course
- The set of all possible values that **X** cannot take on is indicated by the word elsewhere, as in $f_X(u) = 0$ elsewhere
- The set of all possible values that **X** can take on is indicated by defining $f_X(u)$ explicitly (by means of some formula, say) at those values



## Example of convention

$$F_X(u) = \begin{cases} 0 \text{ for } u \le 0, \\ u^3 \text{ for } 0 < u < 1, \\ 1 \text{ for } u \ge 1 \end{cases}$$

- $f_X(u)$ has value $3u^2$ for $0 \le u < 1$
- CDF not differentiable at u = 1. $f_X(1) = ?$
- $f_X(u) = 3u^2$ for $0 \le u \le 1$, and 0 elsewhere means that **X** takes on values in [0, 1]
- $f_X(u) = 3u^2$ for $0 < u < 1$, and 0 elsewhere means that **X** takes on values in (0, 1)



## Remarks on convention

- $f_X(u) = 3u^2$ for $0 \le u \le 1$, and 0 elsewhere means that **X** takes on values in [0, 1]
- $f_X(u) = 3u^2$ for $0 < u < 1$, and 0 elsewhere means that **X** takes on values in (0, 1)
- Note that in both definitions, $f_X(0) = 0$
- In the first definition, the value of $f_X(u)$ at u = 0 is explicitly defined by a formula (even though the formula gives a 0 value)
- In second definition, u = 0 is in elsewhere



## The pdf versus the pmf

- A discrete random variable defines a set of point masses on the axis: total mass = 1
- In contrast, a continuous random variable defines a spread of the total probability mass of 1 along the axis
- There is no mass at any point
- The pdf of a continuous random variable tells the density of the mass at each point

## So, what's a pdf mean, anyway?

- The probability density function (pdf) of a continuous random variable tells us the density of the mass at each point on the axis
- pdf is measured in units of mass/length
- The concept of the pdf is analogous to the more common concepts of the mass density, charge density, etc



## pdfs are not what you think they are

- The probability density function (pdf) of a continuous random variable is not, by itself, a probability
- Example: $f_X(u) = 3u^2$ for $0 \le u \le 1$, and 0 elsewhere
- This pdf has value 3 at $u = 1$
- This means the probability mass is four times as dense at $u = 1$ as at $u = 1/2$



## Multiply by a length to get a mass

- The probability density function (pdf) of a continuous random variable is not, by itself, a probability
- If $f_X(u)$ is positive at the point a, then
  $$P\{a \le \mathbf{X} \le a + \delta a\} \approx f_X(a) \bullet \delta a$$
- Note that $\delta a$ is the length of the interval
- The approximation is better and better as $\delta a$ becomes smaller and smaller



## Why does this work?

$$P\{a \le \mathbf{X} \le a + \delta a\} \approx f_X(a) \bullet \delta a$$

- $f_X(a)$ = derivative of the CDF at point a
  $= \text{limit}_{\Delta a \to 0} [F_X(a+\Delta a) - F_X(a)]/\Delta a$
- But this means that for small values of $\delta a$,
  $P\{a \le \mathbf{X} \le a + \delta a\} = F_X(a+ \delta a) - F_X(a)$
  $\approx f_X(a) \bullet \delta a$
- Furthermore, the smaller the value of $\delta a$, the better the approximation



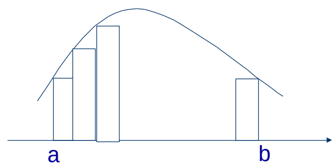
## More generally, …

- If $f_X(u)$ is positive at the point u, then
  $$P\{u \le \mathbf{X} \le u + \delta u\} \approx f_X(u) \bullet \delta u$$
- Let $n \bullet \delta a = b - a$. Then, $P\{a < \mathbf{X} < b\}$
  $= P\{a \le \mathbf{X} \le a+\delta a\} + P\{a+\delta a \le \mathbf{X} \le a+2 \bullet \delta a\}$
  $+ \ldots + P\{a+(n-1) \bullet \delta a \le \mathbf{X} \le a+n \bullet \delta a\}$
- $P\{a < \mathbf{X} < b\} \approx f_X(a) \bullet \delta a + f_X(a+ \delta a) \bullet \delta a + \ldots$
  = area under $f_X(u)$ from a to b in the limit as $\delta a$ goes to 0



## Graphical interpretation

- $P\{a < \mathbf{X} < b\} \approx f_X(a) \bullet \delta a + f_X(a+ \delta a) \bullet \delta a + \ldots$
  = area under $f_X(u)$ from a to b in the limit as $\delta a$ goes to 0
- Each rectangle has base $\delta a$



## Probability = area under the curve

- $P\{a < \mathbf{X} < b\} = P\{a \le \mathbf{X} \le b\} = P\{a \le \mathbf{X} < b\} = P\{a < \mathbf{X} \le b\}$ are all equal to

  the area under the curve $f_X(u)$ from a to b
- The area can be expressed as an integral, but it is of the utmost importance that you remember this notion as area under the curve





## Summary

- A continuous random variable is one whose CDF is continuous everywhere and differentiable everywhere except for a few discrete points
- The pdf is the derivative of the CDF
- The pdf is not a probability; its units are mass/length
- Probabilities are the areas under the pdf