Continuous uncertainty encodings: Densities and CDFs

SIADS 542: Presenting uncertainty – Week 2, Lecture 1

Matthew Kay
Assistant Professor
School of Information
University of Michigan

Roadmap

Week 1: Introduction to uncertainty visualization
Types of uncertainty; small vs large world

Week 2 & 3: Small world uncertainty
Continuous uncertainty encodings
Frequency-framing uncertainty encodings

Week 4: Large world uncertainty

Today

We've established a vocabulary for uncertainty

Today

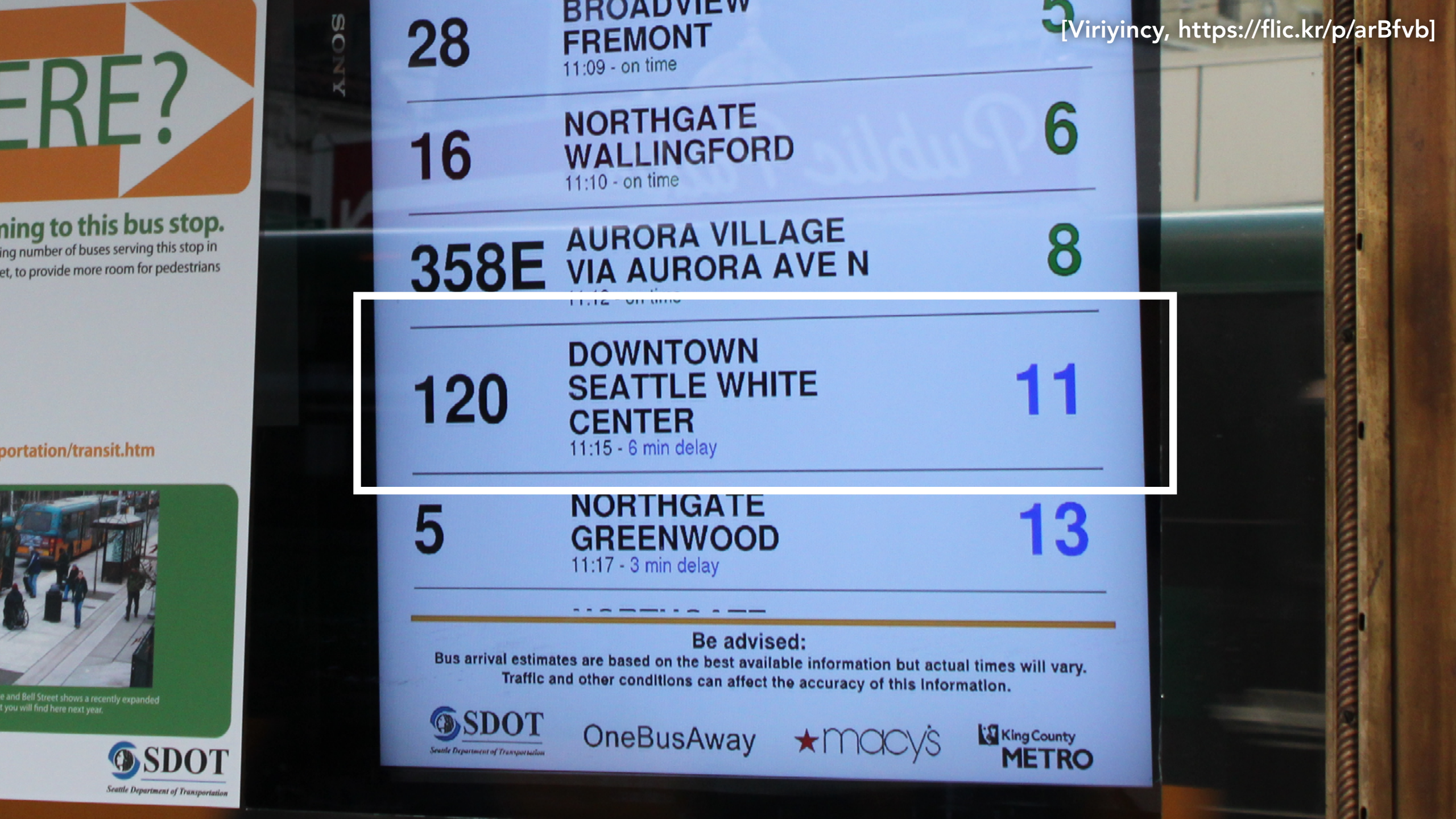
We've established a vocabulary for uncertainty
and a way to formally describe visualizations
of it within the grammar of graphics

Today

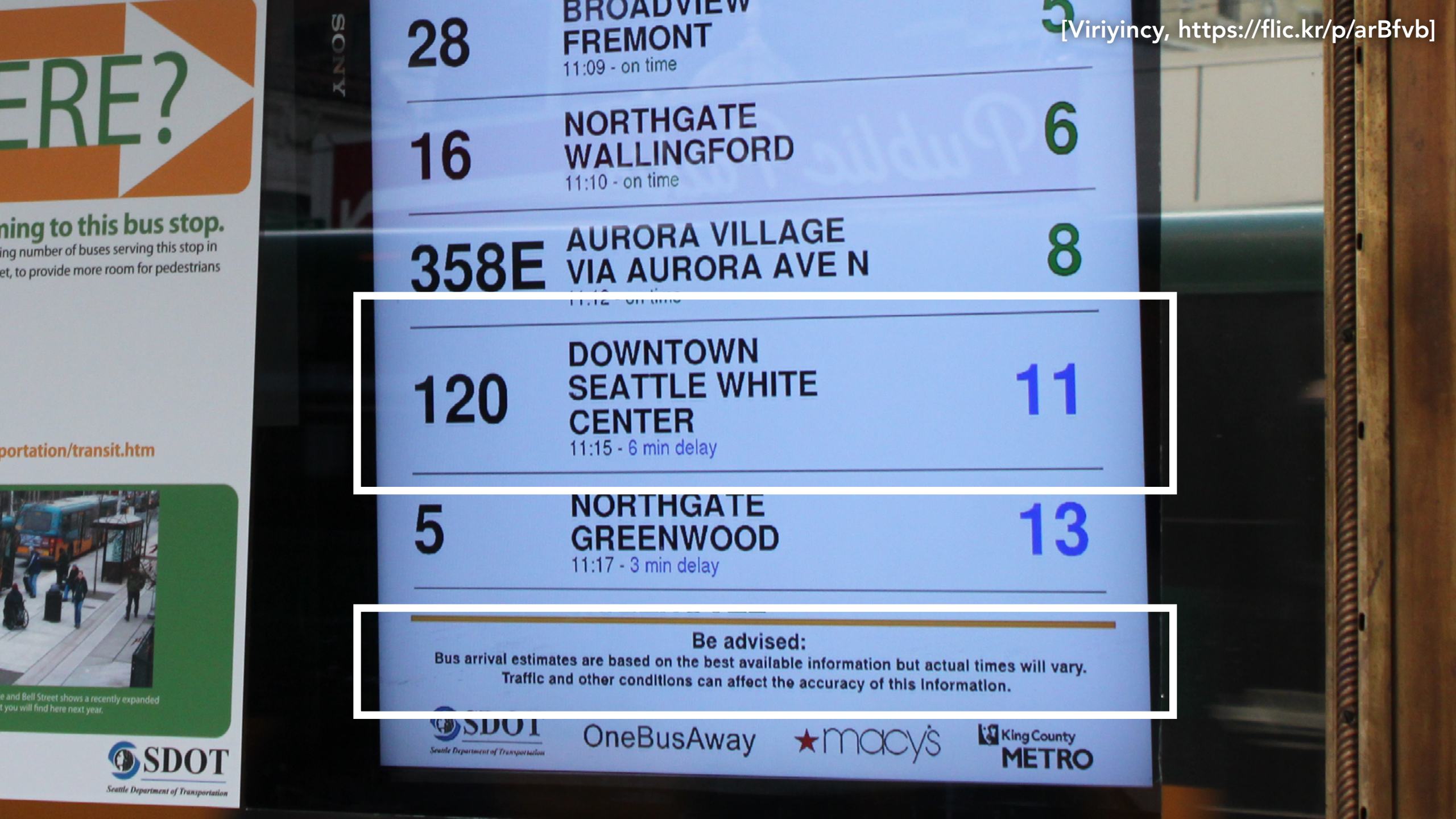
We've established a vocabulary for uncertainty
and a way to formally describe visualizations
of it within the grammar of graphics

Now let's dig into the pros/cons of
continuous uncertainty encodings

Continuous uncertainty encodings: An example scenario

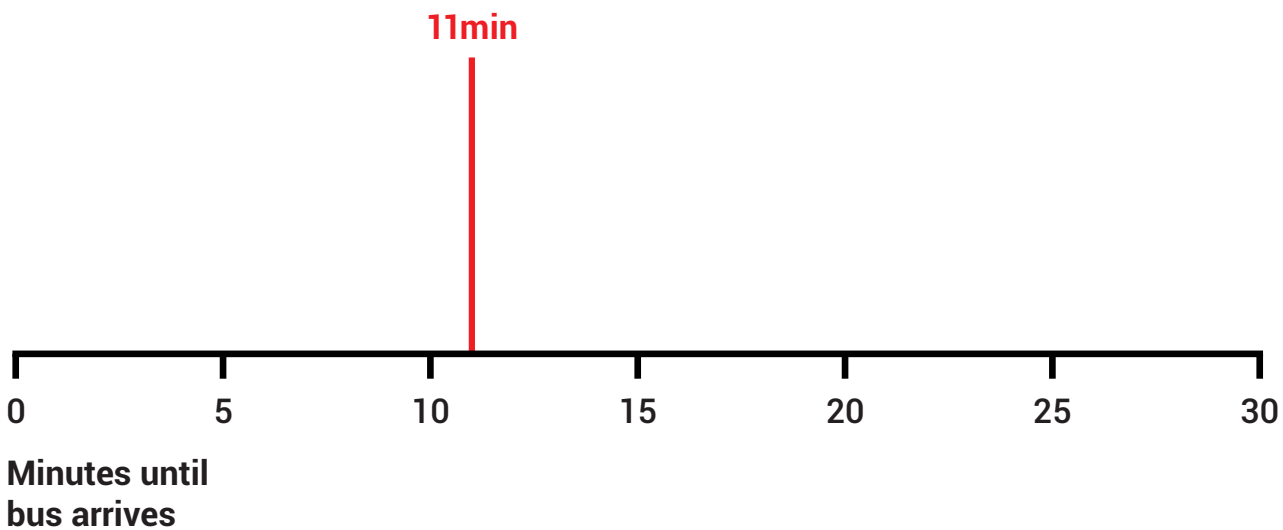


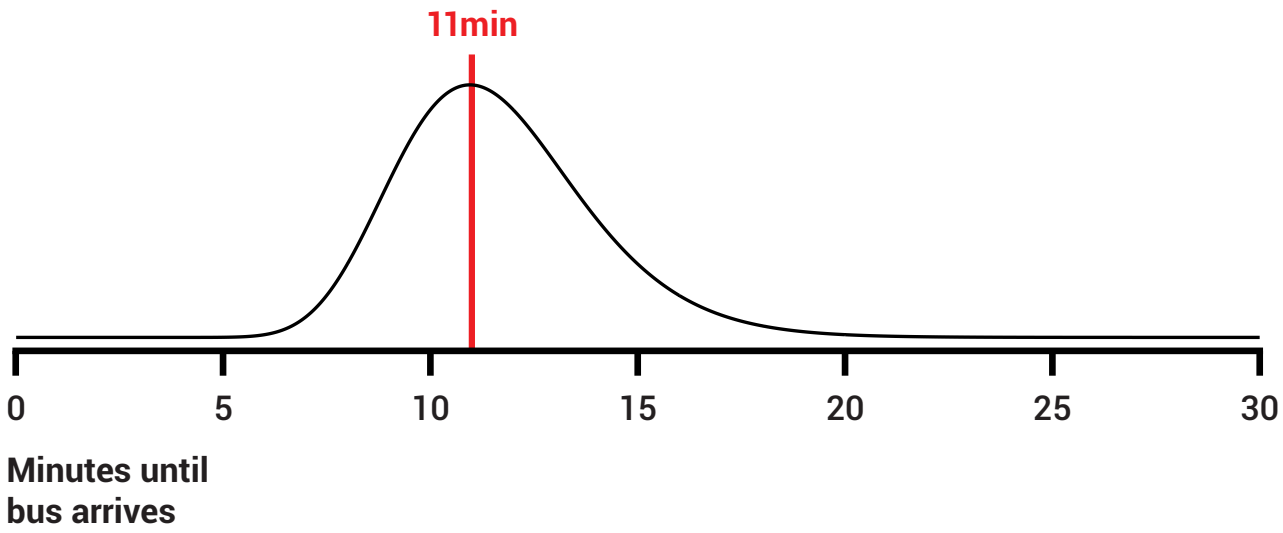
[Viriyincy, https://flic.kr/p/arBfvb]

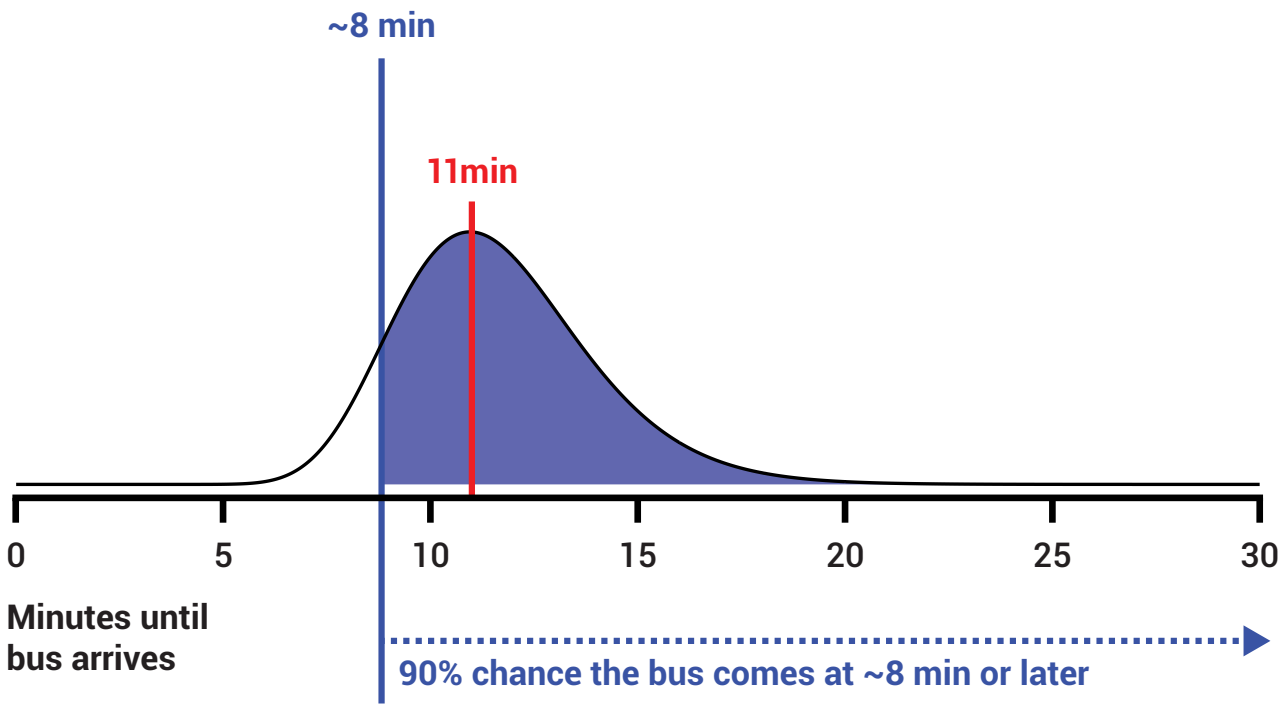


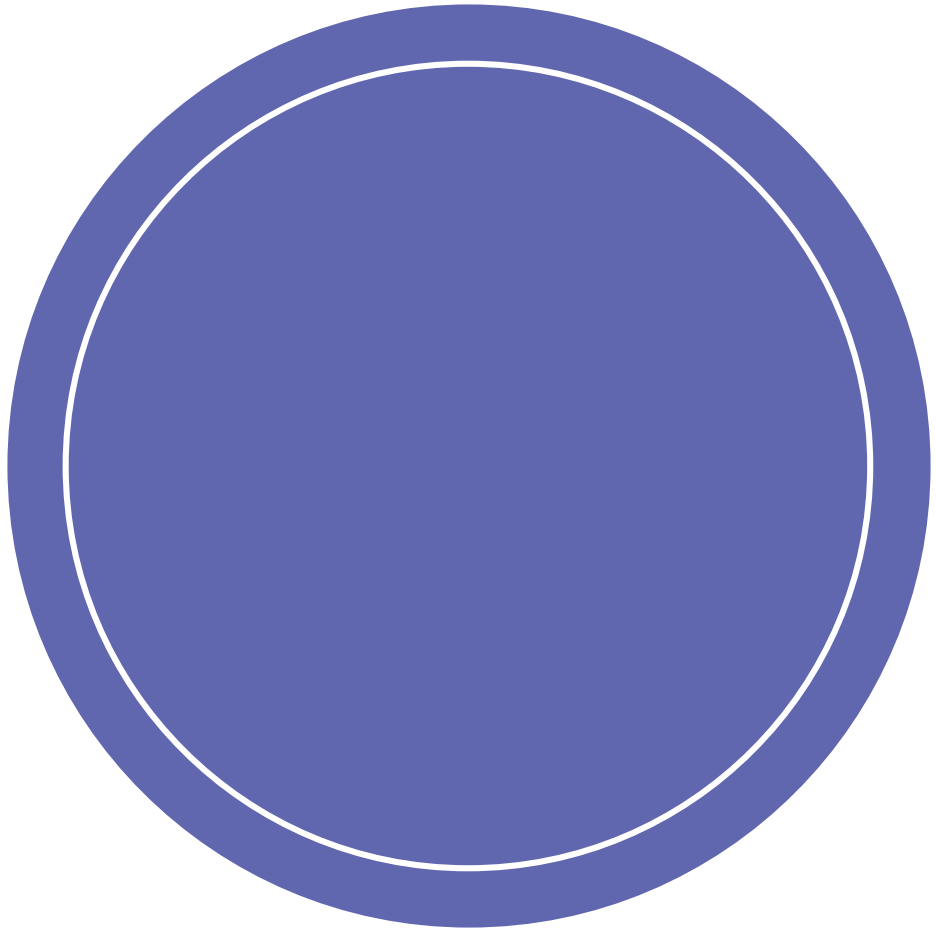
[Viriyincy, https://flic.kr/p/arBfvb]

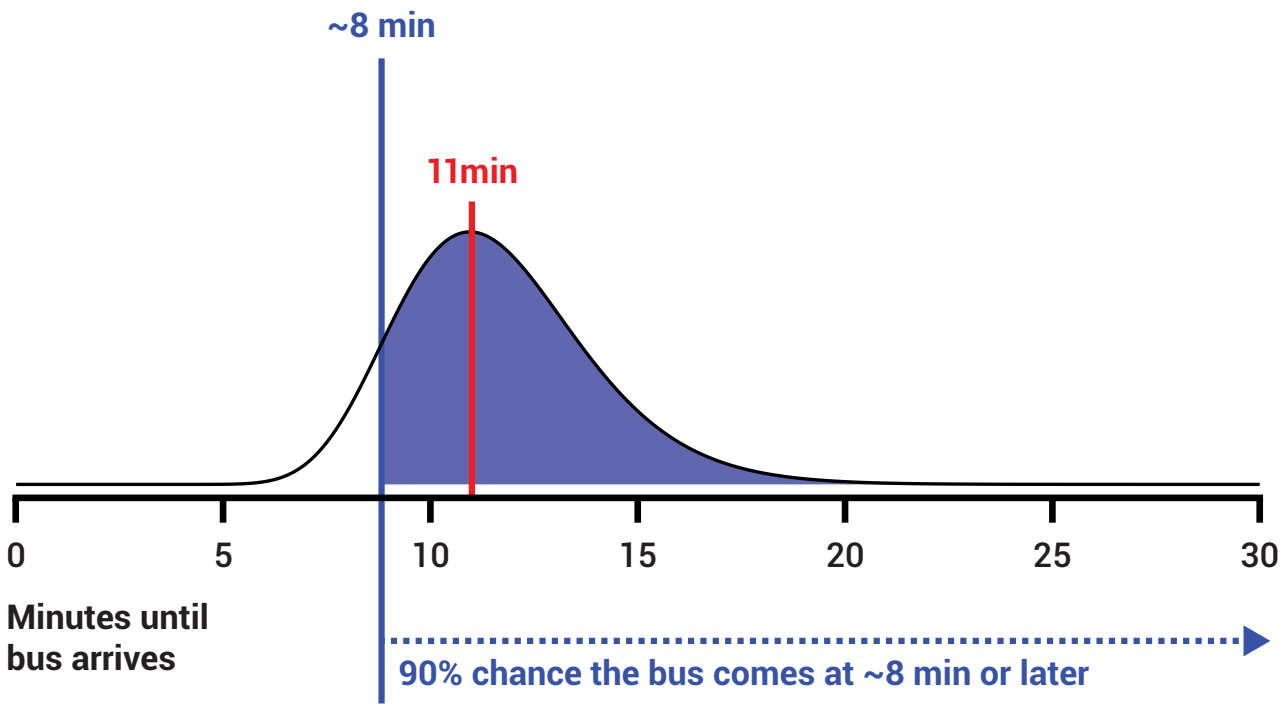
Do I have time to get a coffee?

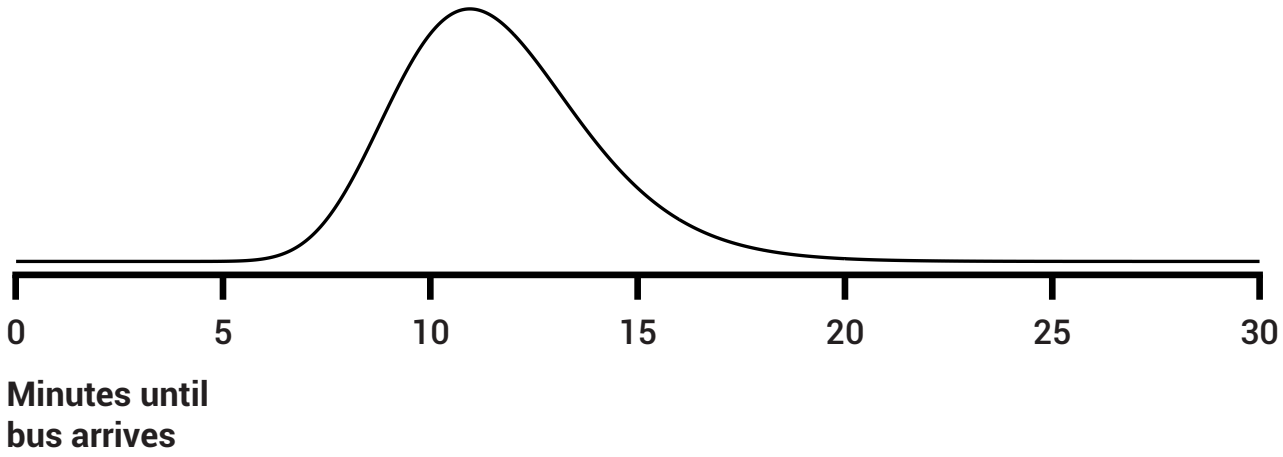

11min
0
5
10
15
20
25
30
Minutes until
bus arrives

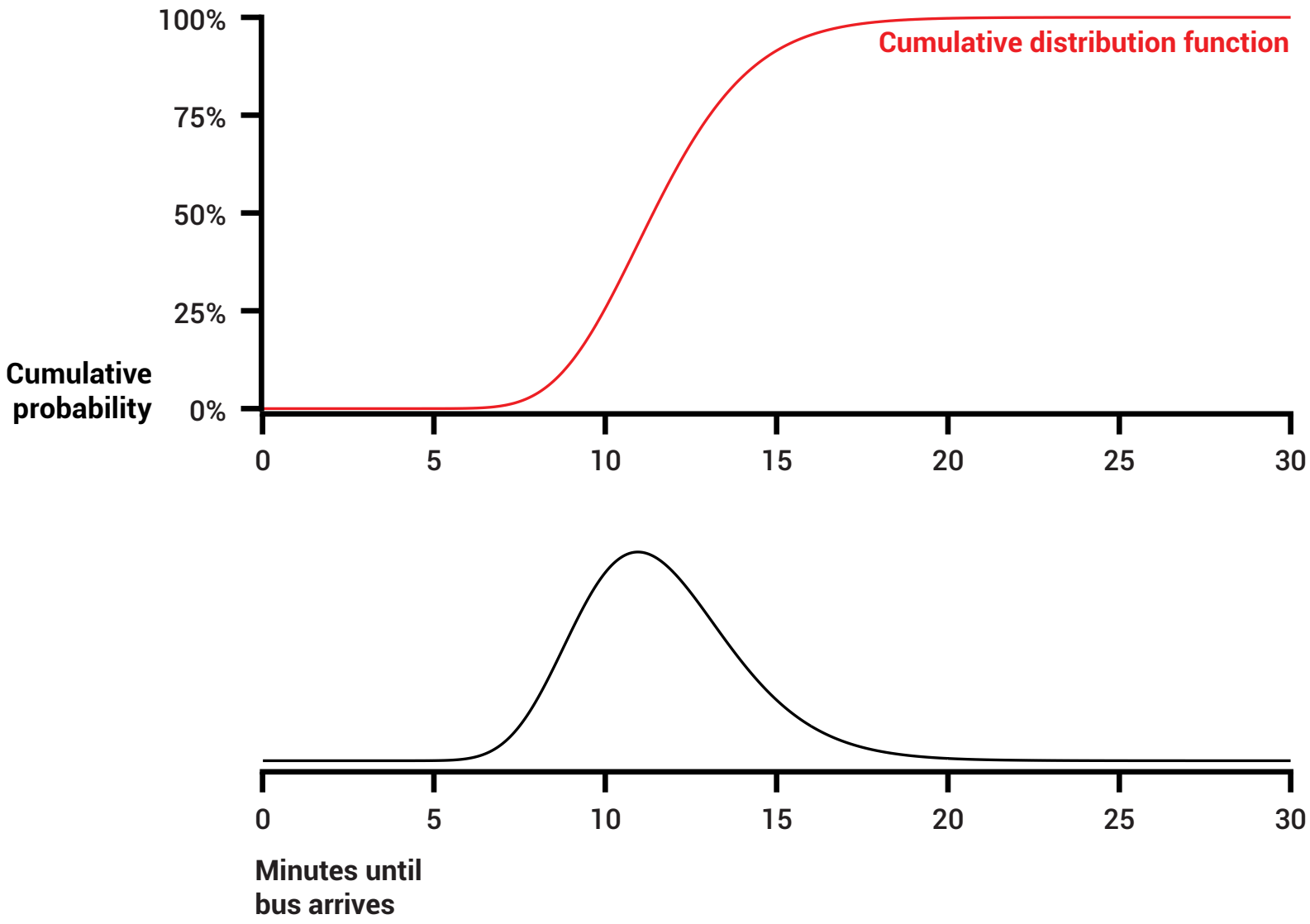

11min
0
5
10
15
20
25
30
Minutes until
bus arrives

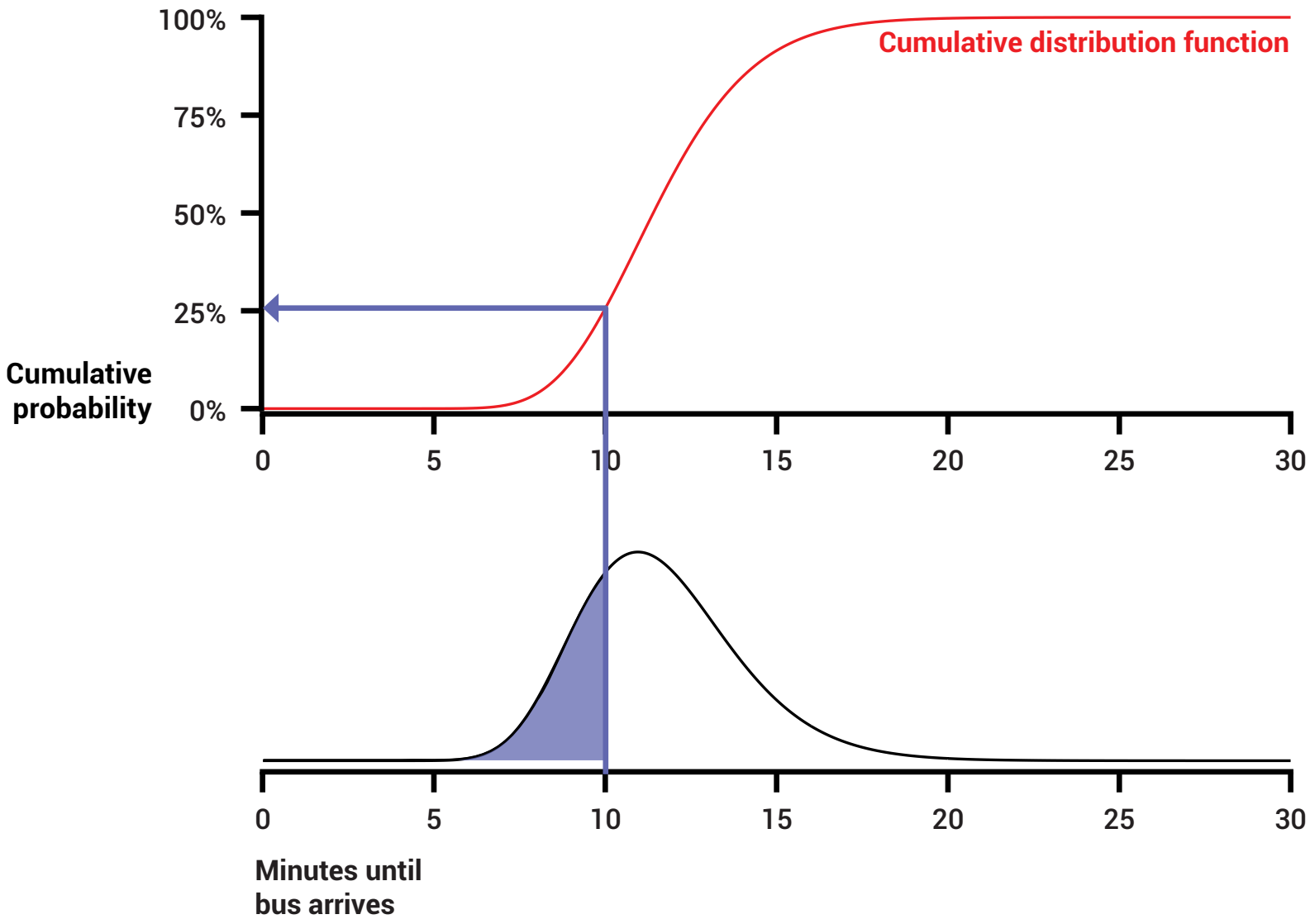

~8 min
11min
0
5
10
15
20
25
30
Minutes until
bus arrives
90% chance the bus comes at ~8 min or later

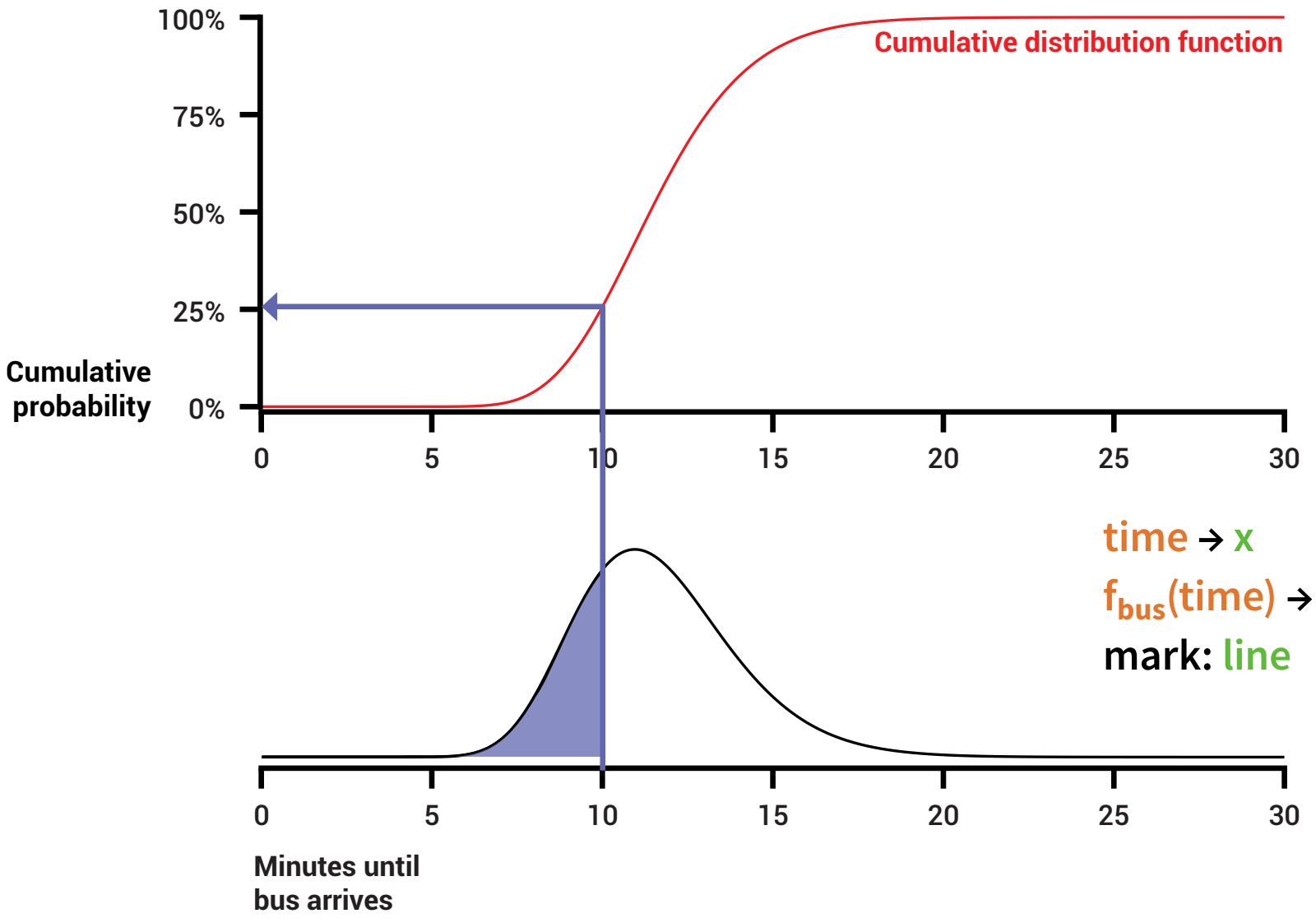


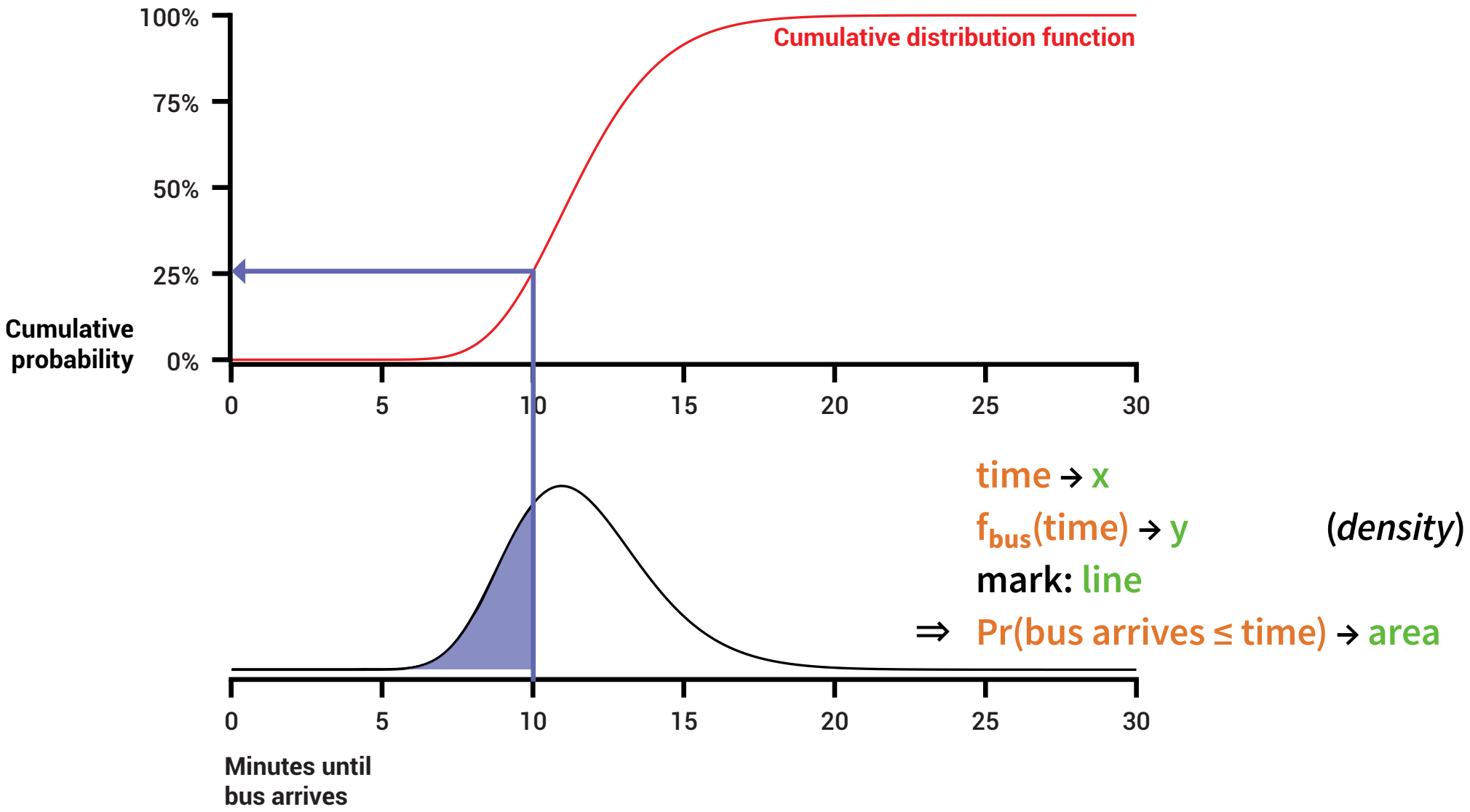

~8 min
11min
0
5
10
15
20
25
30
Minutes until
bus arrives
90% chance the bus comes at ~8 min or later

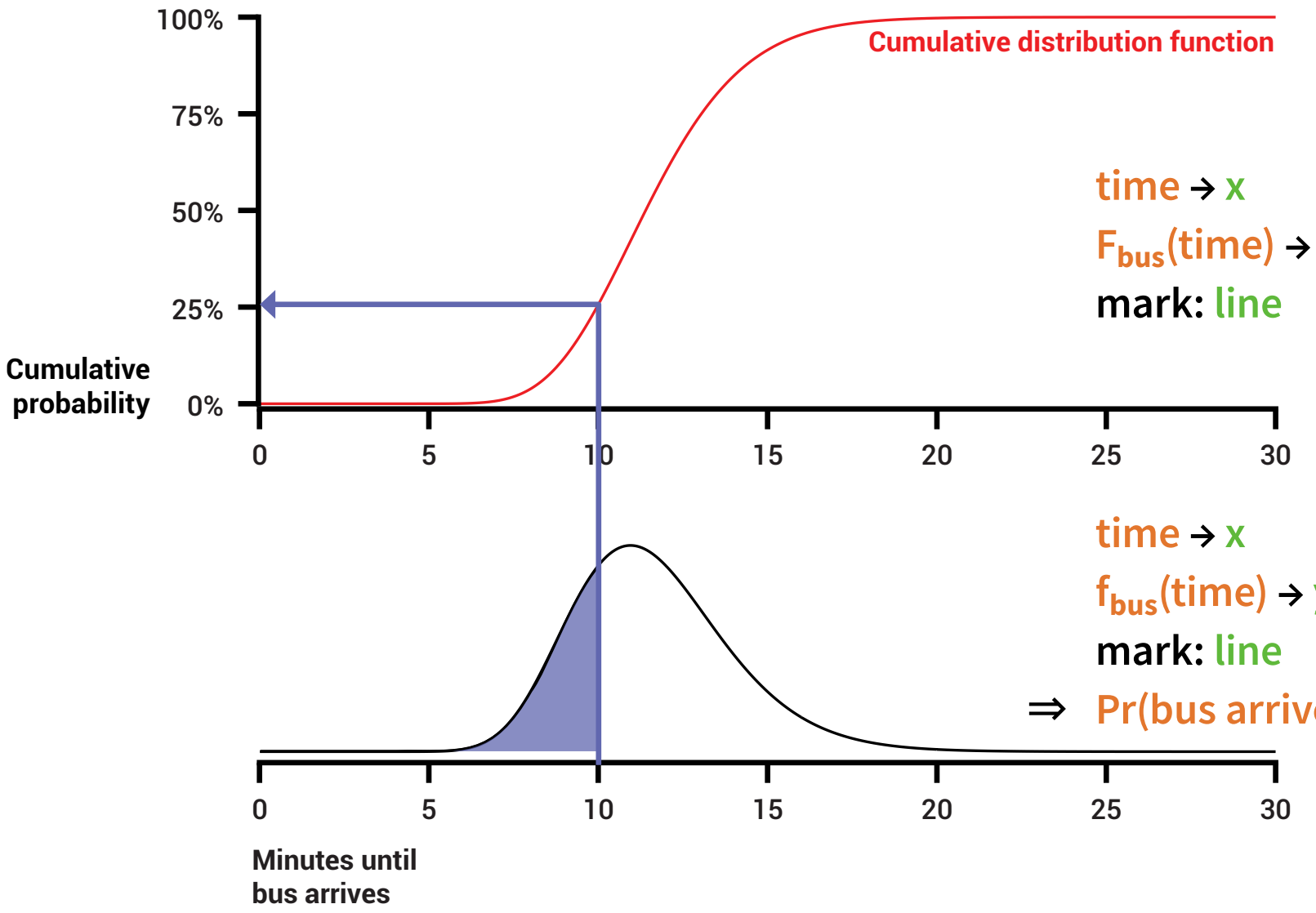

0
5
10
15
20
25
30
Minutes until
bus arrives


100%
75%
50%
25%
0%
Cumulative probability
Cumulative distribution function
0
5
10
15
20
25
30
0
5
10
15
20
25
30
Minutes until bus arrives


100%
75%
50%
25%
0%
Cumulative probability
Cumulative distribution function
0
5
10
15
20
25
30
0
5
10
15
20
25
30
Minutes until bus arrives


100%
75%
50%
25%
0%
Cumulative probability
Cumulative distribution function
0
5
10
15
20
25
30
time → x
f_{bus}(time) → y
mark: line
(*density*)
0
5
10
15
20
25
30
Minutes until bus arrives


100%
75%
50%
25%
0%
Cumulative probability
Cumulative distribution function
0
5
10
15
20
25
30
Minutes until bus arrives
time → x
f_{bus}(time) → y (*density*)
mark: line
⇒ Pr(bus arrives ≤ time) → area



time → x

F_{bus}(time) → y (*CDF*)

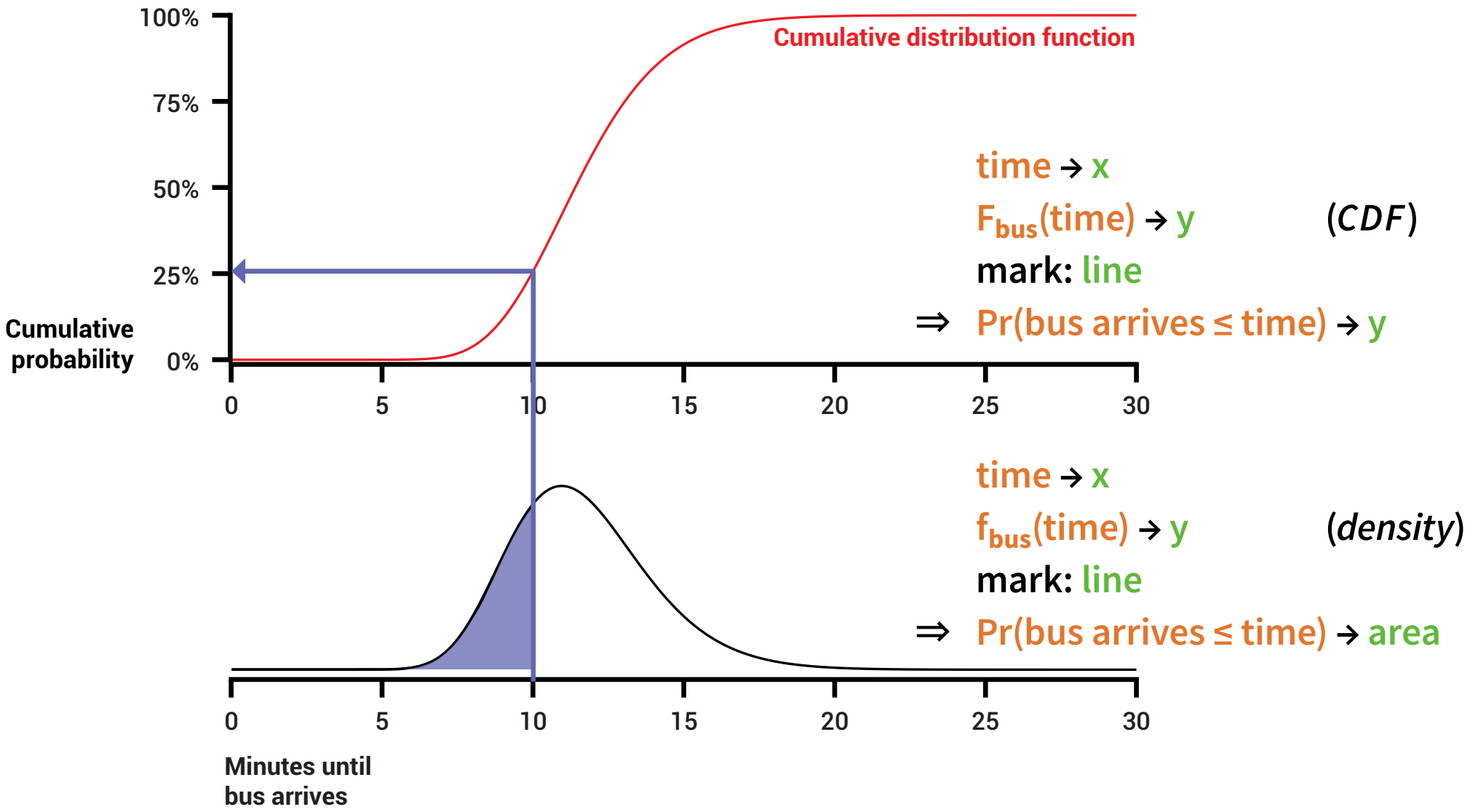
mark: line

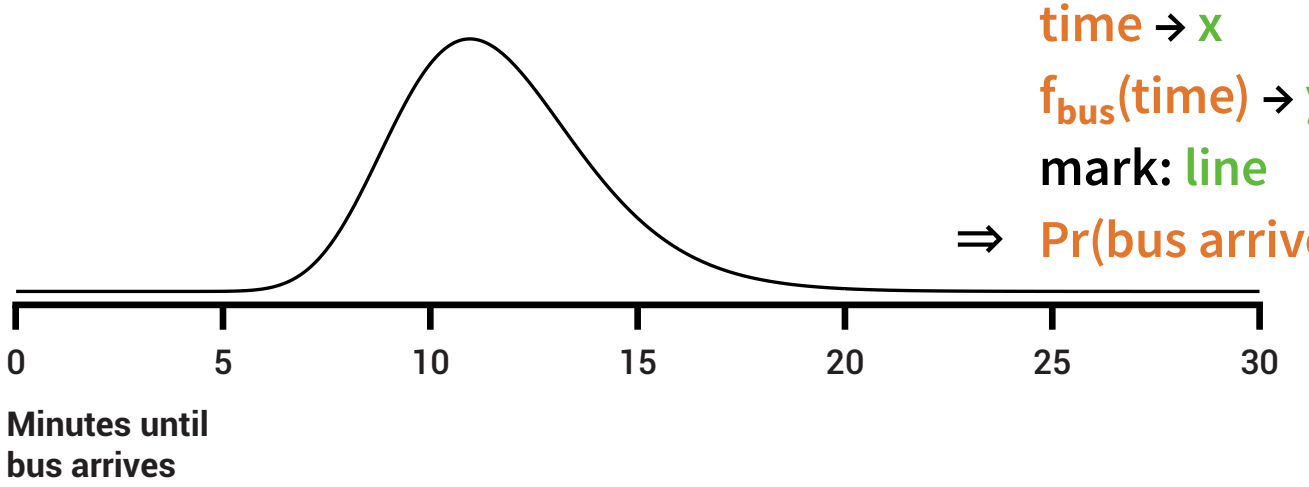
time → x

f_{bus}(time) → y (*density*)

mark: line

⇒ Pr(bus arrives ≤ time) → area


100%
75%
50%
25%
0%
Cumulative probability
Cumulative distribution function
0
5
10
15
20
25
30
time → x
F_{bus}(time) → y (CDF)
mark: line
⇒ Pr(bus arrives ≤ time) → y
time → x
f_{bus}(time) → y (density)
mark: line
⇒ Pr(bus arrives ≤ time) → area
Minutes until bus arrives



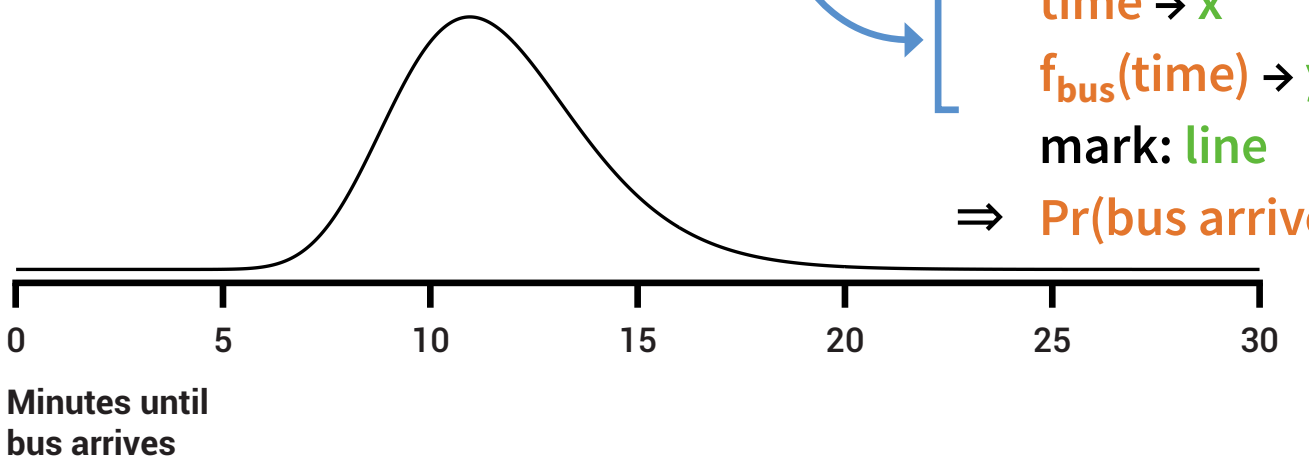
time → x

f_{bus}(time) → y (*density*)

mark: line

⇒ Pr(bus arrives ≤ time) → area

explicit
encodings

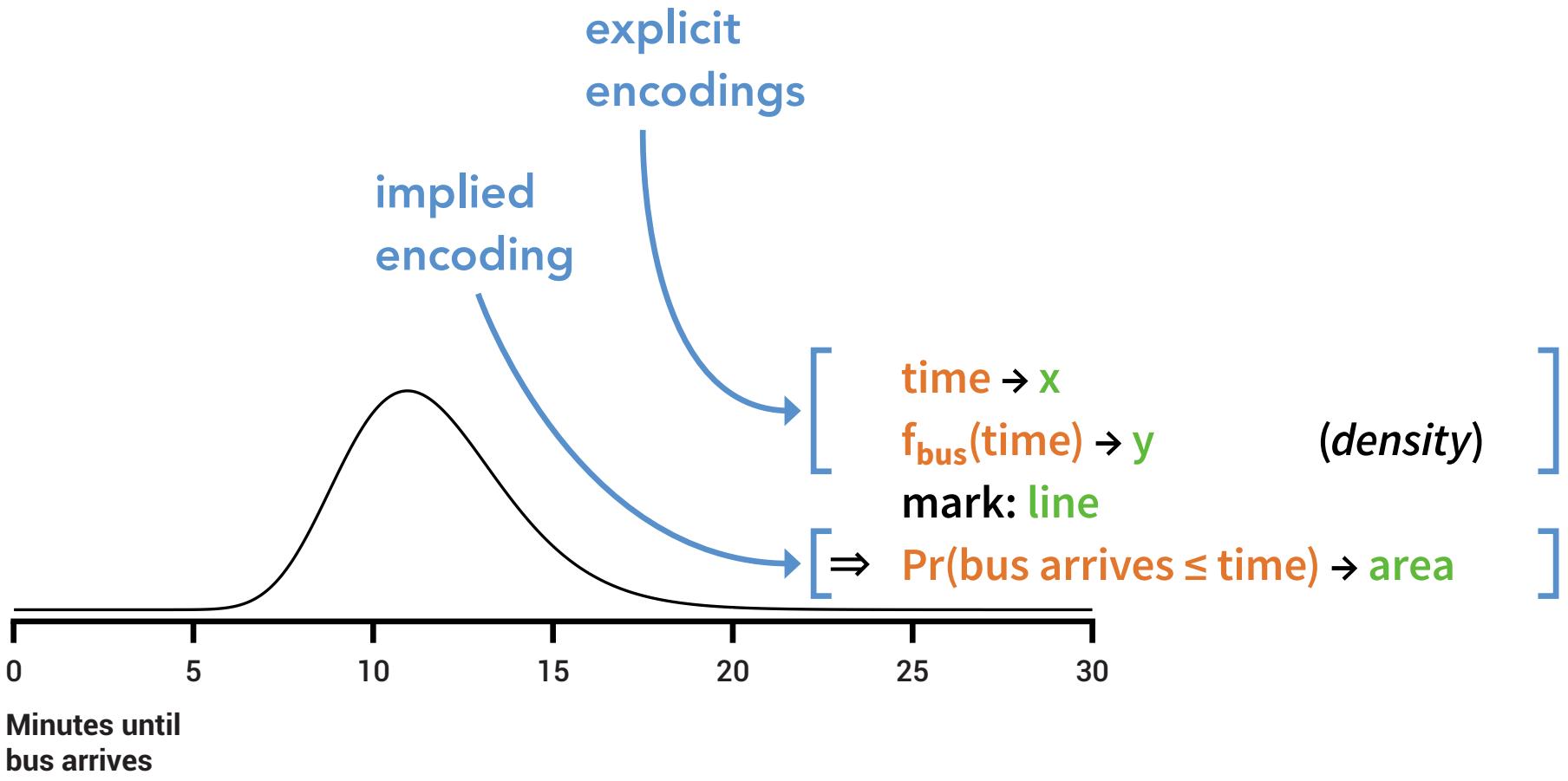


time → x

f_{bus}(time) → y (*density*)

mark: line

⇒ Pr(bus arrives ≤ time) → area

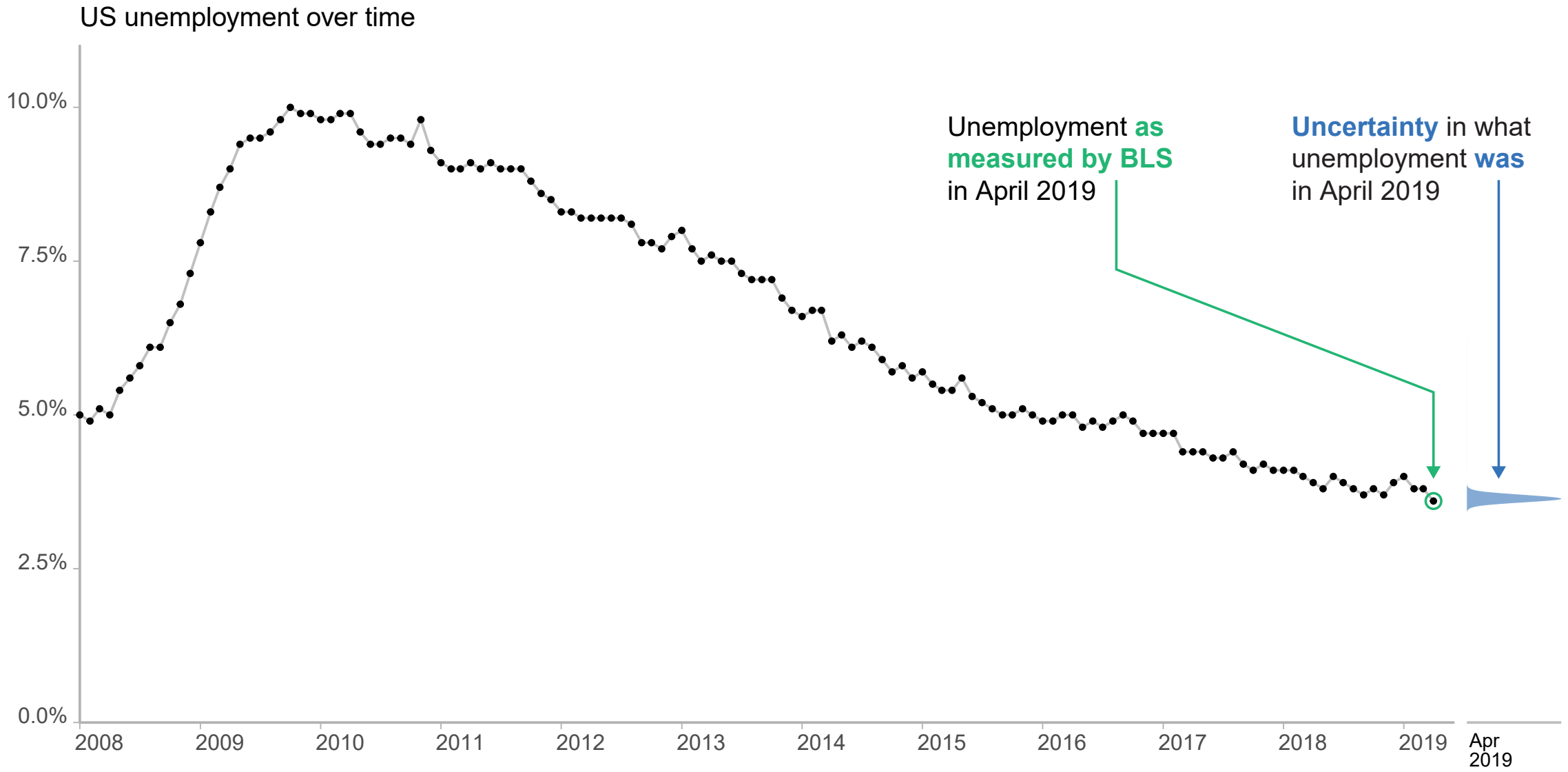

explicit
encodings
implied
encoding
time → x
f_{bus}(time) → y
(density)
mark: line
⇒ Pr(bus arrives ≤ time) → area
0
5
10
15
20
25
30
Minutes until
bus arrives

Continuous uncertainty encodings

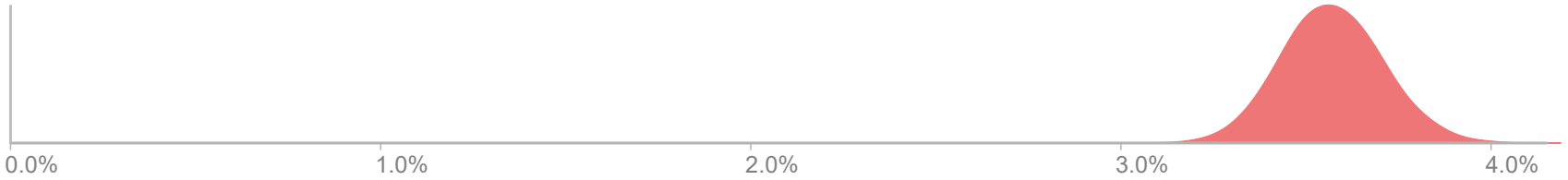
Convey probability/confidence as continuous values

Typically involve mapping density (**f(x)**) or CDF (**F(x)**) onto some visual channel

More examples...


US unemployment over time
10.0%
7.5%
5.0%
2.5%
0.0%
Unemployment as measured by BLS in April 2019
Uncertainty in what unemployment was in April 2019
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
Apr 2019

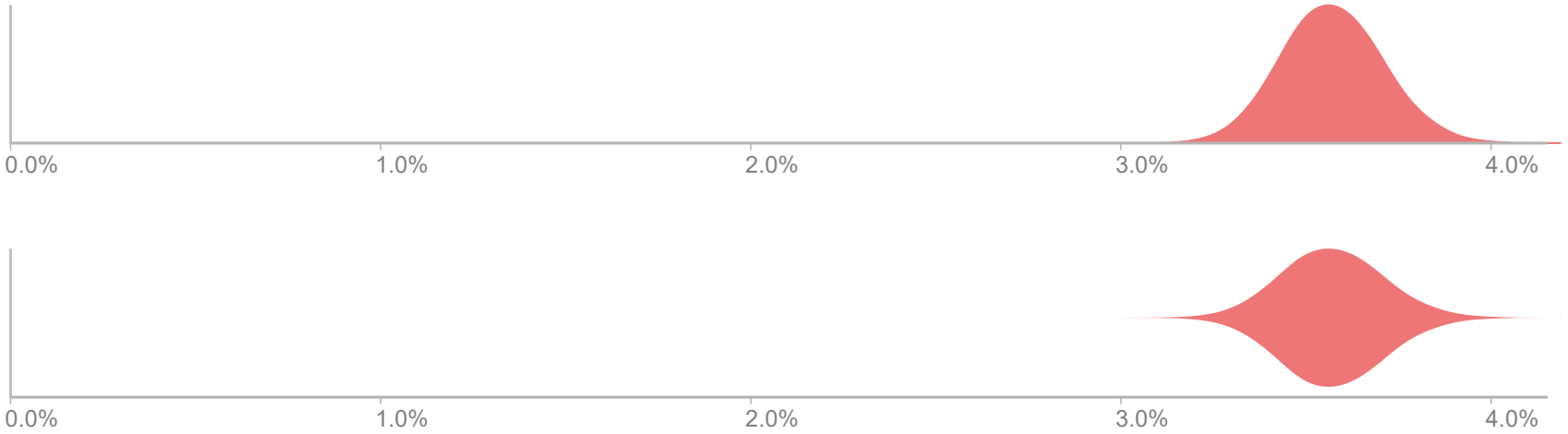
Uncertainty in what US unemployment will be in May 2019: Continuous encodings



Density plot

$f_{unemp}(x) \rightarrow y$

Uncertainty in what US unemployment will be in May 2019: Continuous encodings



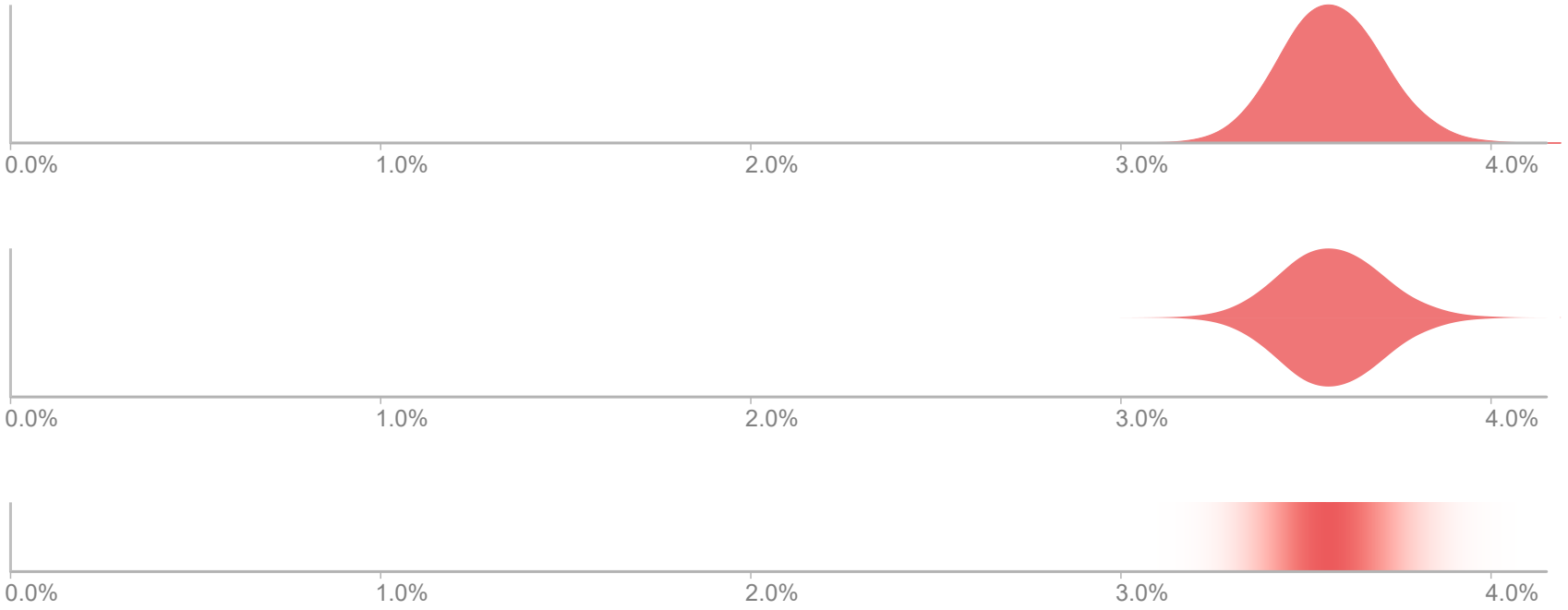
Density plot

$f_{unemp}(x) \rightarrow y$

Violin plot

$f_{unemp}(x) \rightarrow$ thickness

Uncertainty in what US unemployment will be in May 2019: Continuous encodings



Density plot

$f_{unemp}(x) \rightarrow y$

Violin plot

$f_{unemp}(x) \rightarrow$ thickness

Gradient plot

$f_{unemp}(x) \rightarrow$ opacity

Uncertainty in what US unemployment will be in May 2019: Continuous encodings

Density plot

$f_{unemp}(x) \rightarrow y$

Violin plot

$f_{unemp}(x) \rightarrow$ thickness

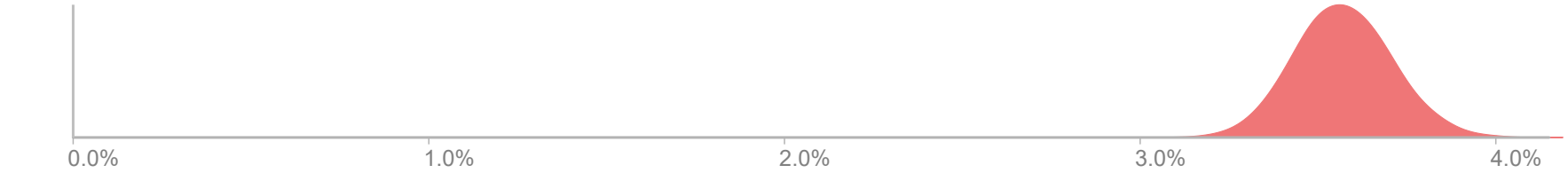
Gradient plot

$f_{unemp}(x) \rightarrow$ opacity

CDF

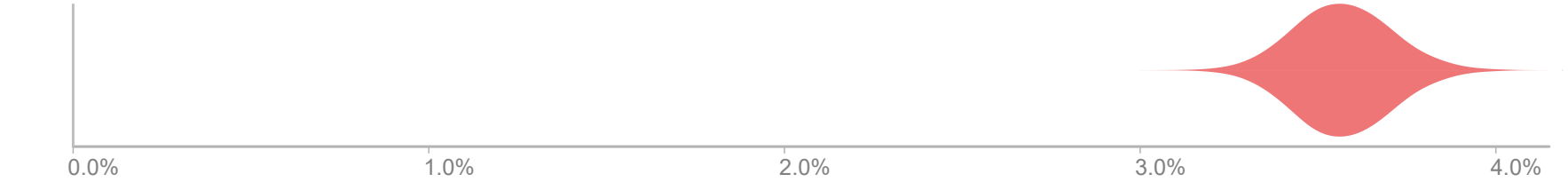
$F_{unemp}(x) \rightarrow$ height

Uncertainty in what US unemployment will be in May 2019: Continuous encodings



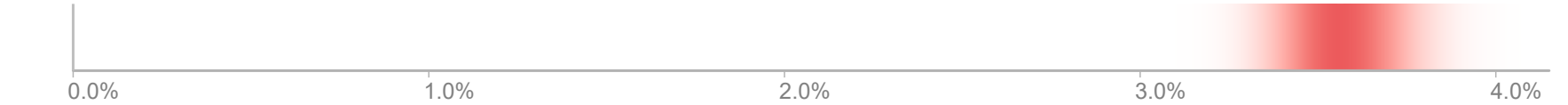
Density plot

$f_{unemp}(x) \rightarrow y$



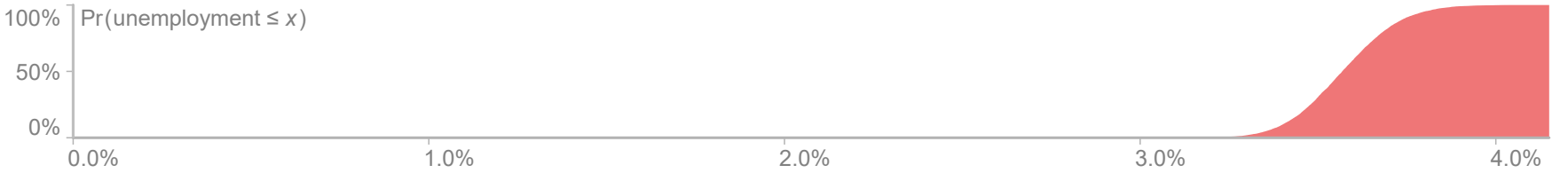
Violin plot

$f_{unemp}(x) \rightarrow$ thickness



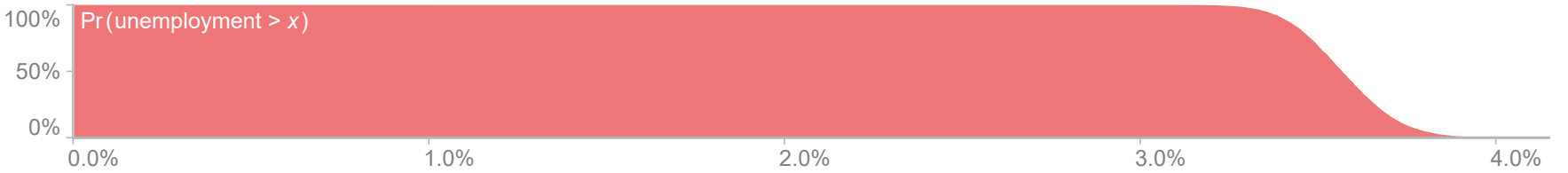
Gradient plot

$f_{unemp}(x) \rightarrow$ opacity



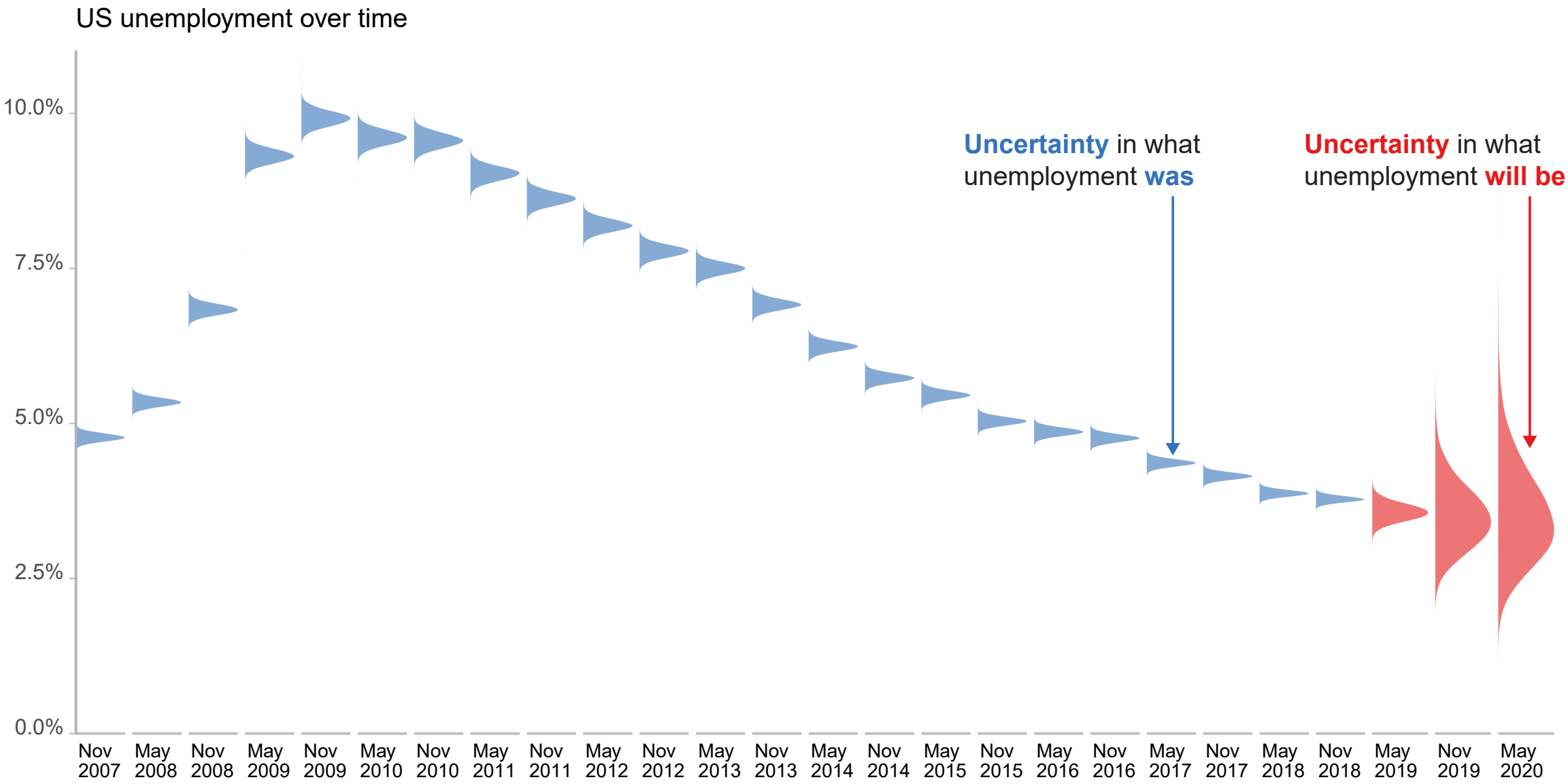
CDF

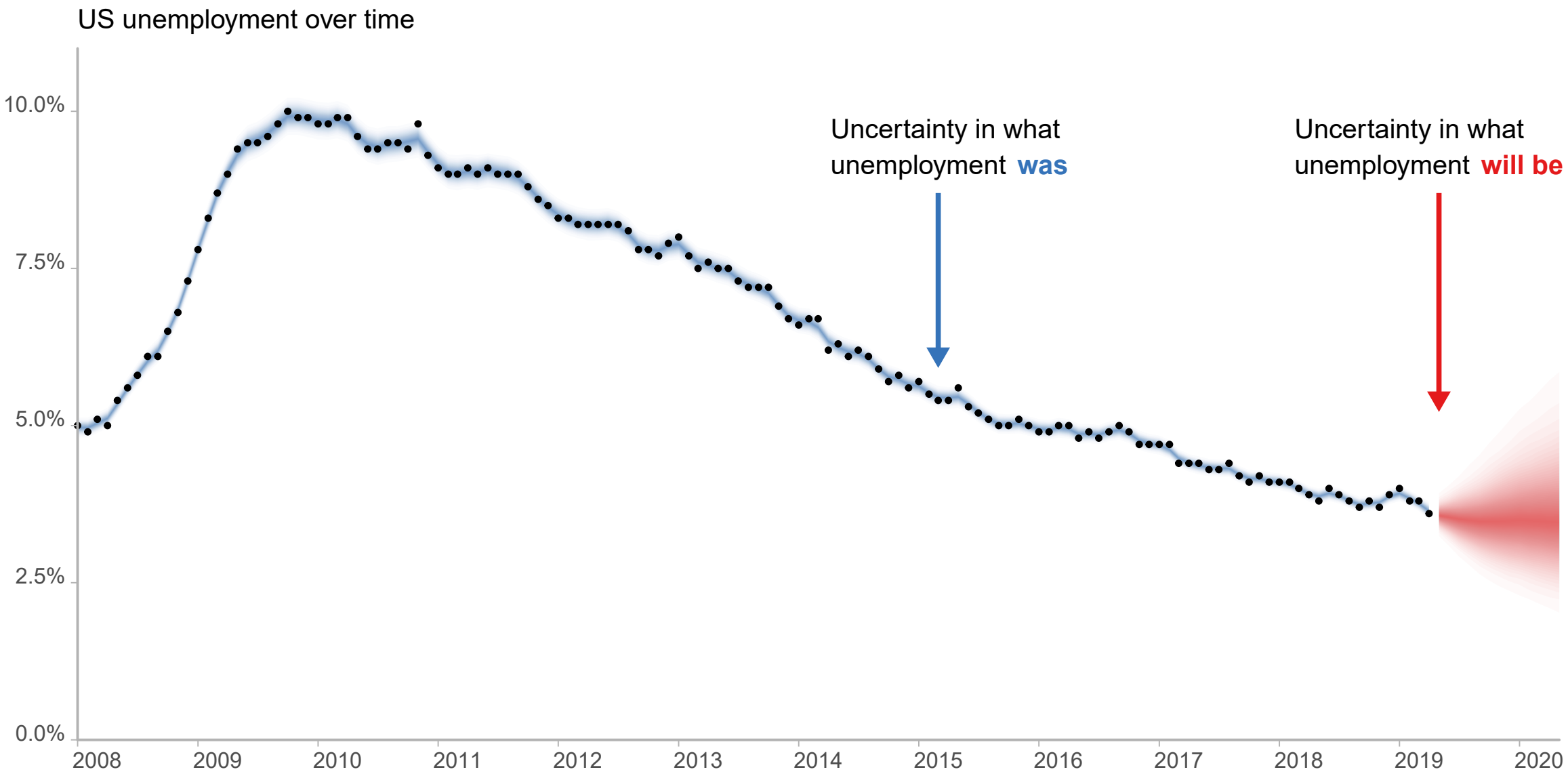
$F_{unemp}(x) \rightarrow$ height

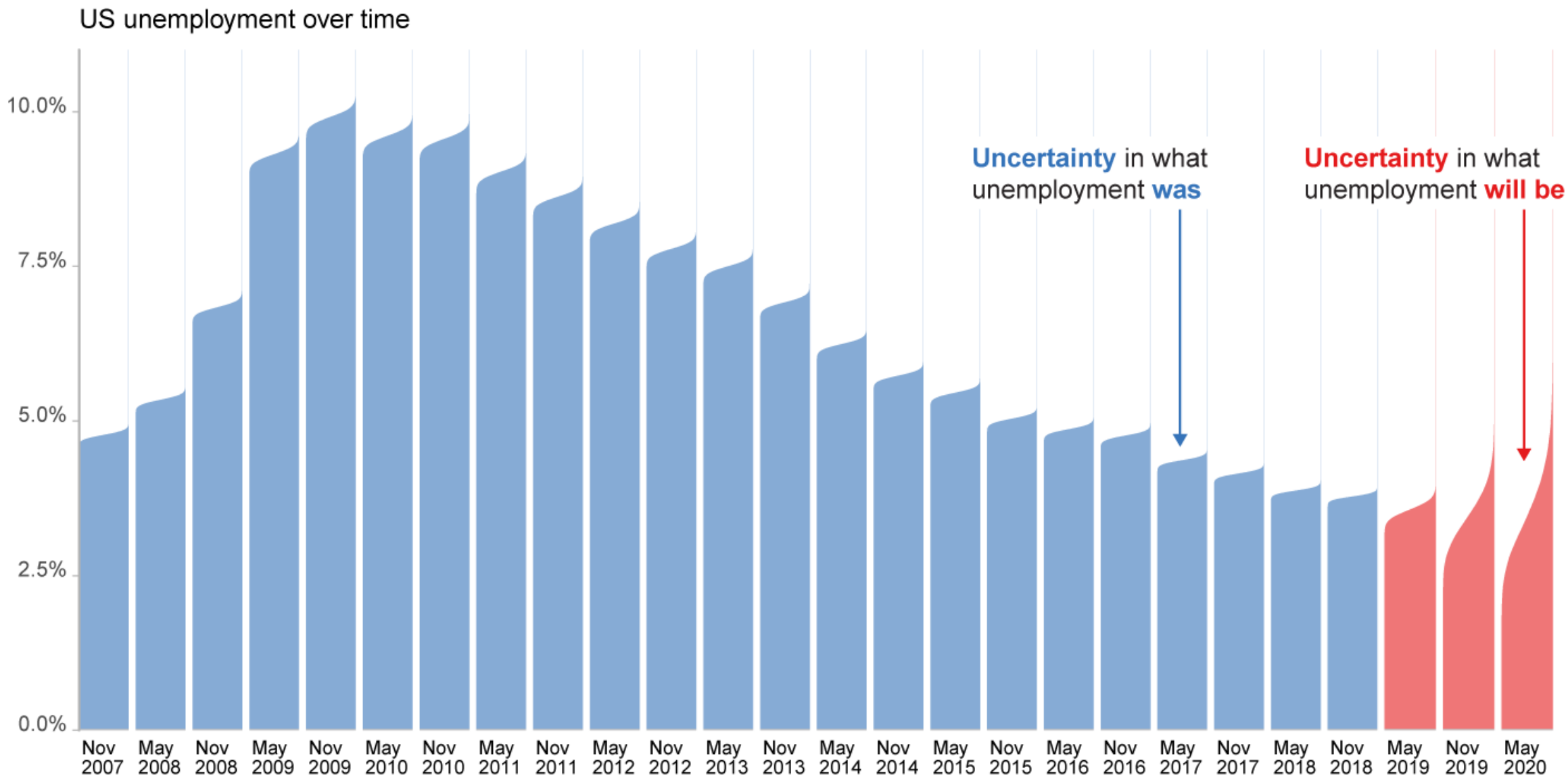


Complementary CDF

$1 - F_{unemp}(x) \rightarrow$ height


US unemployment over time
10.0%
7.5%
5.0%
2.5%
0.0%
Uncertainty in what unemployment was
Uncertainty in what unemployment will be
Nov 2007
May 2008
Nov 2008
May 2009
Nov 2009
May 2010
Nov 2010
May 2011
Nov 2011
May 2012
Nov 2012
May 2013
Nov 2013
May 2014
Nov 2014
May 2015
Nov 2015
May 2016
Nov 2016
May 2017
Nov 2017
May 2018
Nov 2018
May 2019
Nov 2019
May 2020


US unemployment over time
10.0%
7.5%
5.0%
2.5%
0.0%
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
Uncertainty in what unemployment was
Uncertainty in what unemployment will be


US unemployment over time
10.0%
7.5%
5.0%
2.5%
0.0%
Uncertainty in what unemployment was
Uncertainty in what unemployment will be
Nov 2007
May 2008
Nov 2008
May 2009
Nov 2009
May 2010
Nov 2010
May 2011
Nov 2011
May 2012
Nov 2012
May 2013
Nov 2013
May 2014
Nov 2014
May 2015
Nov 2015
May 2016
Nov 2016
May 2017
Nov 2017
May 2018
Nov 2018
May 2019
Nov 2019
May 2020

Summing up

Continuous uncertainty encodings

Usually involve densities or CDFs

Important to consider implied encodings, tradeoffs in resolution / effectiveness

Continuous uncertainty encodings: Densities and CDFs

SIADS 542: Presenting uncertainty – Week 2, Lecture 1

Matthew Kay
Assistant Professor
School of Information
University of Michigan