The Storm of the Century! (Statistics of Extremes)

Lee Fawcett

School of Maths & Stats, Newcastle University

lee.fawcett@ncl.ac.uk

Who am I?

- **Lecturer** in the School of Maths & Stats at Newcastle University
- **Applied statistician**
- **Teach**: 1st/2nd/4th year students
- **Research**: Statistical models for extreme weather events
- **Other stuff**: Admissions, recruitment, consultancy, Schools...

Quick survey

- Who **likes** Maths at School?
- Who thinks they are **good** at Maths?
- Who thinks Maths is **pointless**?
  - Algebra
  - Pythagoras
  - Data handling

Quick survey

- Who **likes** Maths at School?
- Who thinks they are **good** at Maths?
- Who thinks Maths is **pointless**?
  - Algebra
  - Pythagoras
  - **Data handling**

Background: Statistics of extremes

So many areas of Statistics are interested in what happens **on average**: Mean, median, mode...

One area that is different is the field of **extreme value theory**.

Why might we be interested in **extremes**, rather than averages?

Background: Statistics of extremes

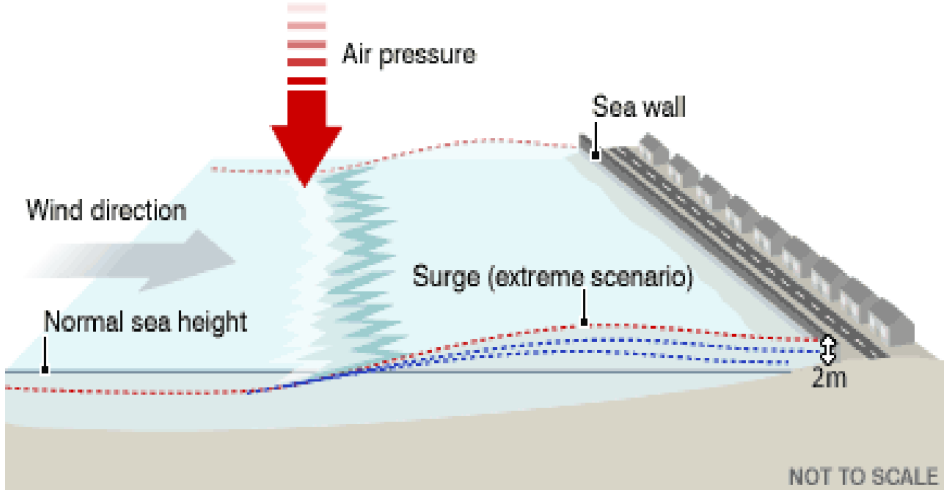
Earth Scientists – for example, **meteorologists**, **seismologists**, **hydrologists** and **oceanographers** – predict and assess the likelihood of storms, earthquakes, volcanoes and changes in sea level.

We can usually deal with averages here – it is the **extremes** that cause devastation!

Such scientists require the expertise of **mathematicians** to help with **predictions** and **calculations**.

In this talk I will focus on how maths/statistics can be used to help oceanographers and civil engineers to predict extreme **sea–surges**.

Sea–surge


Air pressure
Sea wall
Wind direction
Surge (extreme scenario)
Normal sea height
2m
NOT TO SCALE

Earthquakes/sea–surge: Japan, 2011

Boxing Day Tsunami: Indian Ocean, 2004

Boxing Day Tsunami: Indian Ocean, 2004

Boxing Day Tsunami: Indian Ocean, 2004

Hurricane Sandy: Cuba, 2012

Hurricane Sandy: USA, 2012

Hurricane Sandy: USA, 2012

Hurricane Sandy: USA, 2012

Hurricane ????: USA, 2015?

Sea–surge: The Great North Sea Flood, 1953

Sea–surge: The Great North Sea Flood, 1953

- Killed **2,551 people** in the UK, Holland and Belgium
- Estimated as the "**Storm of the Century**"
- Often referred to as "**Europe's Katrina**"

Sea–surge: The Great North Sea Flood, 2025?

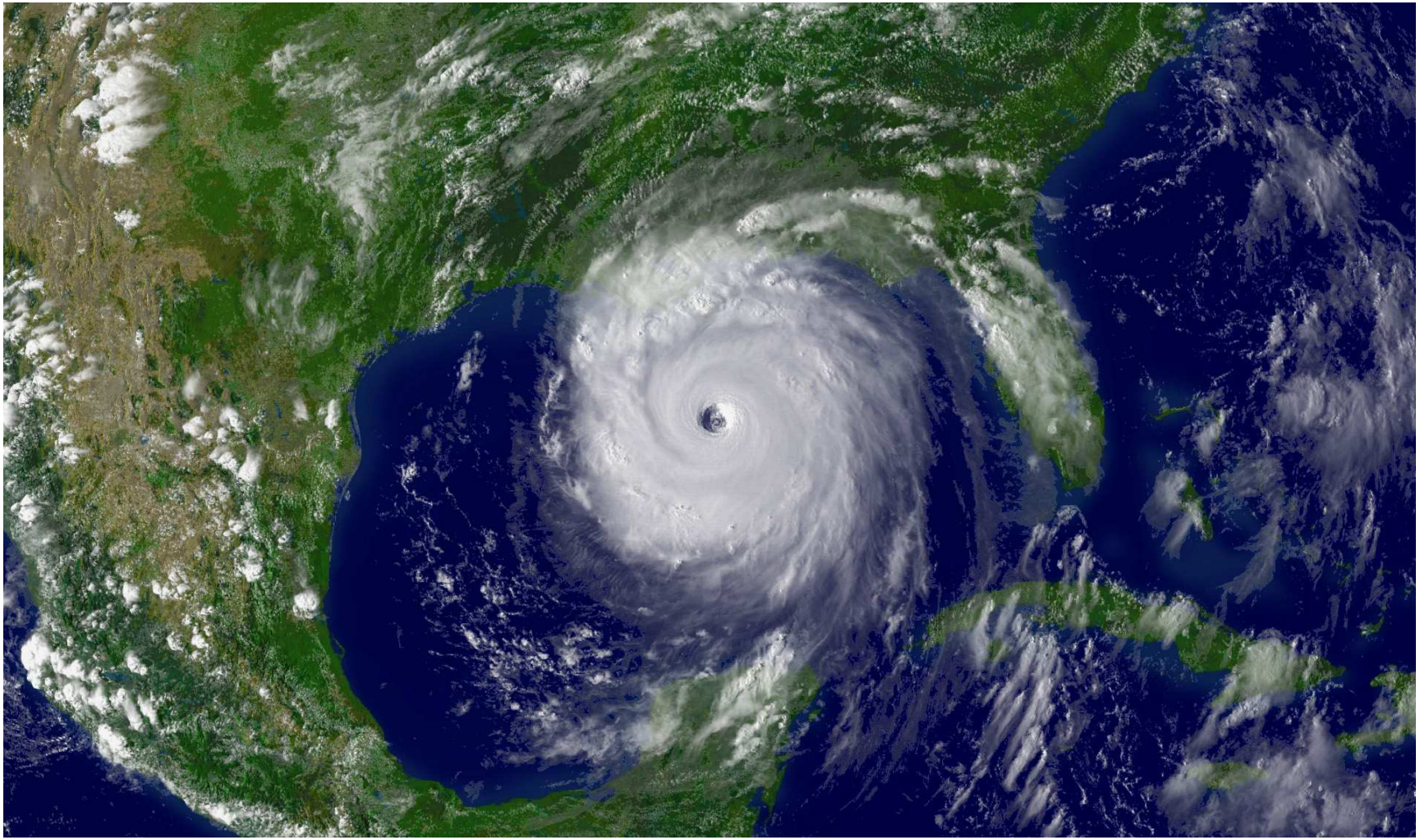
Extreme rainfall: NE England, June 2012

Extreme rainfall: NE England, June 2012

Extreme rainfall: NE England, June 2012

Extreme rainfall: NE England, June 2012

Hurricane Katrina: New Orleans, August 2005



Hurricane Katrina: New Orleans, August 2005

Some facts

- Category 5 Hurricane (Category 3 when it made landfall)
- Killed 1833 people
- Caused $108 bn worth of damage
- Most damage/loss of life caused by **storm surge**
- Storm surge reached 14.4 feet above sea level
- Lowest air pressure 902 mb
- Political controversy
- Billed as the “storm of the century”

Hurricane Katrina: New Orleans, August 2005

Some facts

- Category 5 Hurricane (Category 3 when it made landfall)
- Killed 1833 people
- Caused $108 bn worth of damage
- Most damage/loss of life caused by **storm surge**
- Storm surge reached 14.4 feet above sea level
- Lowest air pressure 902 mb
- Political controversy
- Billed as the “storm of the century”

Hurricane Katrina: New Orleans, August 2005

Hurricane Katrina: New Orleans, August 2005

Hurricane Katrina: New Orleans, August 2005

Hurricane Katrina: New Orleans, August 2005

Hurricane Katrina: New Orleans, August 2005

Data application: Sea surges at New Orleans

Below are the annual maximum sea–surge levels observed at New Orleans over a 50 year period *before Hurricane Katrina*:

8.5	8.9	9.1	8.9	8.4	9.7	9.1	9.6	8.7	9.3
9.6	9.3	8.7	9.0	8.8	8.9	8.9	12.2	7.8	7.7
8.3	8.1	7.3	6.8	6.7	7.3	7.6	8.2	8.6	9.8
9.5	7.4	7.3	10.2	10.3	10.4	8.8	9.7	10.0	10.8
11.1	12.7	11.5	11.8	12.6	13.0	10.5	10.5	10.0	9.4

Data application: Sea surges at New Orleans

Suppose you are the mathematician on the team of scientists working on a new section of flood defence system in New Orleans.

One of the civil engineers asks you to work out some exceedance probabilities for her.

In particular, she wants to know the probability that, this year, the annual maximum sea surge at New Orleans will

1. exceed **8.75 feet**;
2. exceed **11.25 feet**;
3. exceed **14 feet**.

Part A: The relative frequency approach

You will have used this approach to probability in your maths at School.

It helps if the data are re–written in ascending order:

6.7	6.8	7.3	7.3	7.3	7.4	7.6	7.7	7.8	8.1
8.2	8.3	8.4	8.5	8.6	8.7	8.7	8.8	8.8	8.9
8.9	8.9	8.9	9.0	9.1	9.1	9.3	9.3	9.4	9.5
9.6	9.6	9.7	9.7	9.8	10.0	10.0	10.2	10.3	10.4
10.5	10.5	10.8	11.1	11.5	11.8	12.2	12.6	12.7	13.0

Then,

$$\boldsymbol{P}(\text{sea–surge exceeds 8.75 feet}) =$$

Part A: The relative frequency approach

You will have used this approach to probability in your maths at School.

It helps if the data are re–written in ascending order:

6.7	6.8	7.3	7.3	7.3	7.4	7.6	7.7	7.8	8.1
8.2	8.3	8.4	8.5	8.6	8.7	8.7	**8.8**	**8.8**	**8.9**
8.9	**8.9**	**8.9**	**9.0**	**9.1**	**9.1**	**9.3**	**9.3**	**9.4**	**9.5**
9.6	**9.6**	**9.7**	**9.7**	**9.8**	**10.0**	**10.0**	**10.2**	**10.3**	**10.4**
10.5	**10.5**	**10.8**	**11.1**	**11.5**	**11.8**	**12.2**	**12.6**	**12.7**	**13.0**

Then,

$$\boldsymbol{P}(\text{sea–surge exceeds 8.75 feet}) = \frac{\mathbf{33}}{50} = \frac{66}{100} = 0.66.$$

Part A: The relative frequency approach

Using the relative frequency approach, work out the other two exceedance probabilities:

- $\boldsymbol{P}$(sea–surge exceeds 11.25 feet)
- $\boldsymbol{P}$(sea–surge exceeds 14 feet)

Part A: The relative frequency approach

6.7	6.8	7.3	7.3	7.3	7.4	7.6	7.7	7.8	8.1
8.2	8.3	8.4	8.5	8.6	8.7	8.7	8.8	8.8	8.9
8.9	8.9	8.9	9.0	9.1	9.1	9.3	9.3	9.4	9.5
9.6	9.6	9.7	9.7	9.8	10.0	10.0	10.2	10.3	10.4
10.5	10.5	10.8	11.1	11.5	11.8	12.2	12.6	12.7	13.0

$$\boldsymbol{P}(\text{sea–surge exceeds 11.25 feet}) =$$

Part A: The relative frequency approach

6.7	6.8	7.3	7.3	7.3	7.4	7.6	7.7	7.8	8.1
8.2	8.3	8.4	8.5	8.6	8.7	8.7	8.8	8.8	8.9
8.9	8.9	8.9	9.0	9.1	9.1	9.3	9.3	9.4	9.5
9.6	9.6	9.7	9.7	9.8	10.0	10.0	10.2	10.3	10.4
10.5	10.5	10.8	11.1	**11.5**	**11.8**	**12.2**	**12.6**	**12.7**	**13.0**

$$\boldsymbol{P}(\text{sea–surge exceeds 11.25 feet}) = \frac{\mathbf{6}}{50} = 0.12$$

Part A: The relative frequency approach

6.7	6.8	7.3	7.3	7.3	7.4	7.6	7.7	7.8	8.1
8.2	8.3	8.4	8.5	8.6	8.7	8.7	8.8	8.8	8.9
8.9	8.9	8.9	9.0	9.1	9.1	9.3	9.3	9.4	9.5
9.6	9.6	9.7	9.7	9.8	10.0	10.0	10.2	10.3	10.4
10.5	10.5	10.8	11.1	11.5	11.8	12.2	12.6	12.7	13.0

$$\boldsymbol{P}(\text{sea–surge exceeds 14 feet}) =$$

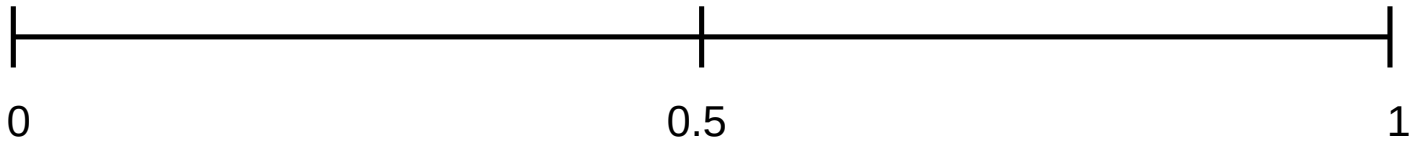
Part A: The relative frequency approach

6.7	6.8	7.3	7.3	7.3	7.4	7.6	7.7	7.8	8.1
8.2	8.3	8.4	8.5	8.6	8.7	8.7	8.8	8.8	8.9
8.9	8.9	8.9	9.0	9.1	9.1	9.3	9.3	9.4	9.5
9.6	9.6	9.7	9.7	9.8	10.0	10.0	10.2	10.3	10.4
10.5	10.5	10.8	11.1	11.5	11.8	12.2	12.6	12.7	13.0

$$\boldsymbol{P}(\text{sea–surge exceeds 14 feet}) = \frac{\mathbf{0}}{50} = 0$$

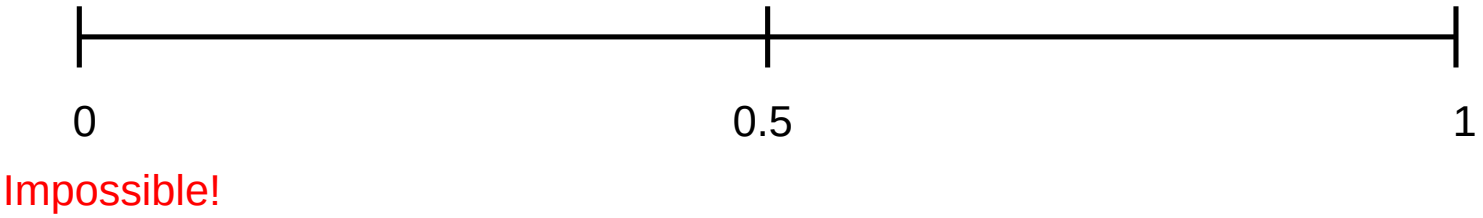
Part A: The relative frequency approach

Think about the **probability scale**:



Part A: The relative frequency approach

Think about the **probability scale**:

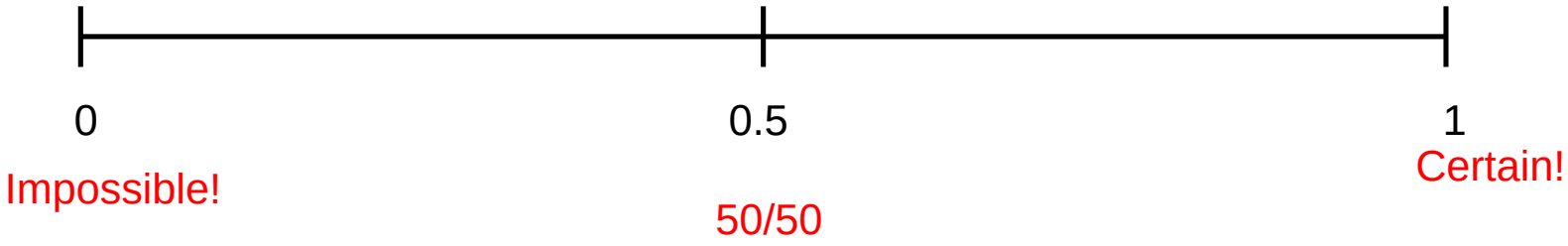


Part A: The relative frequency approach

Think about the **probability scale**:

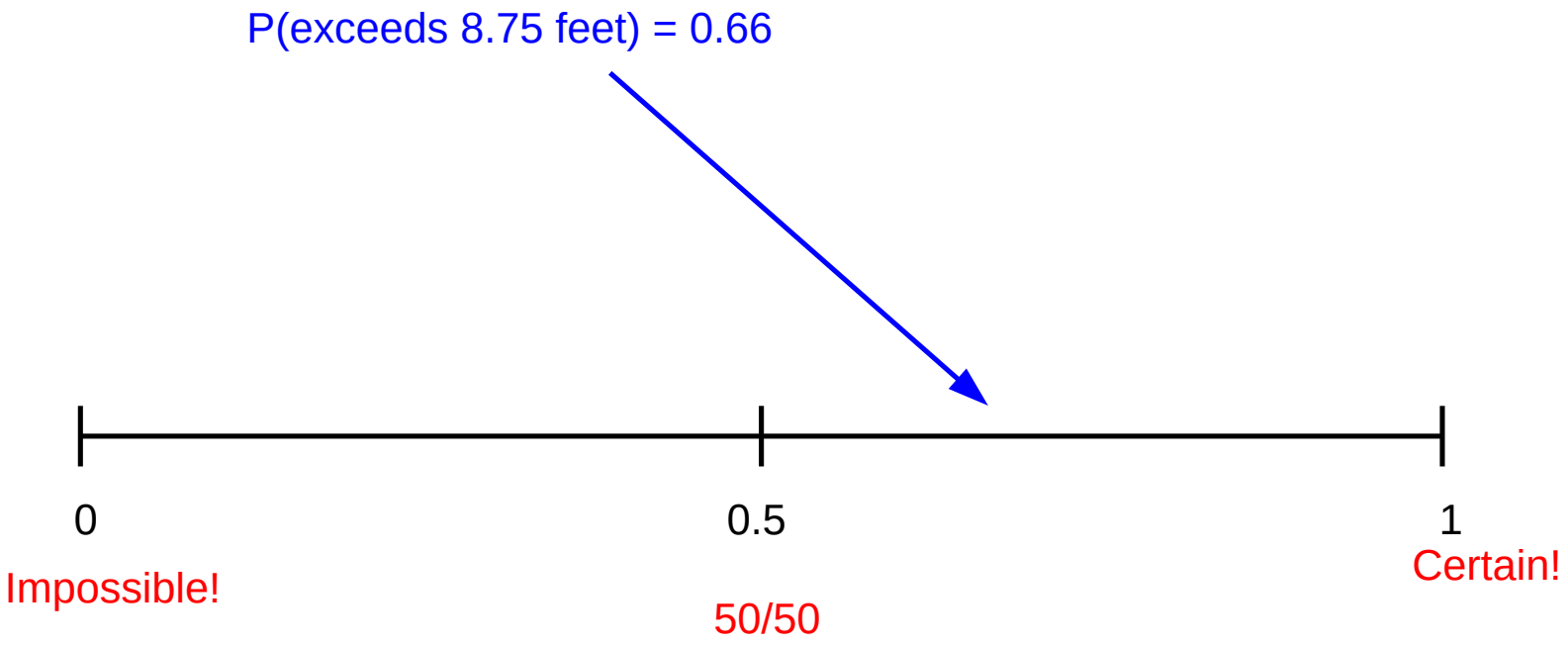
Part A: The relative frequency approach

Think about the **probability scale**:



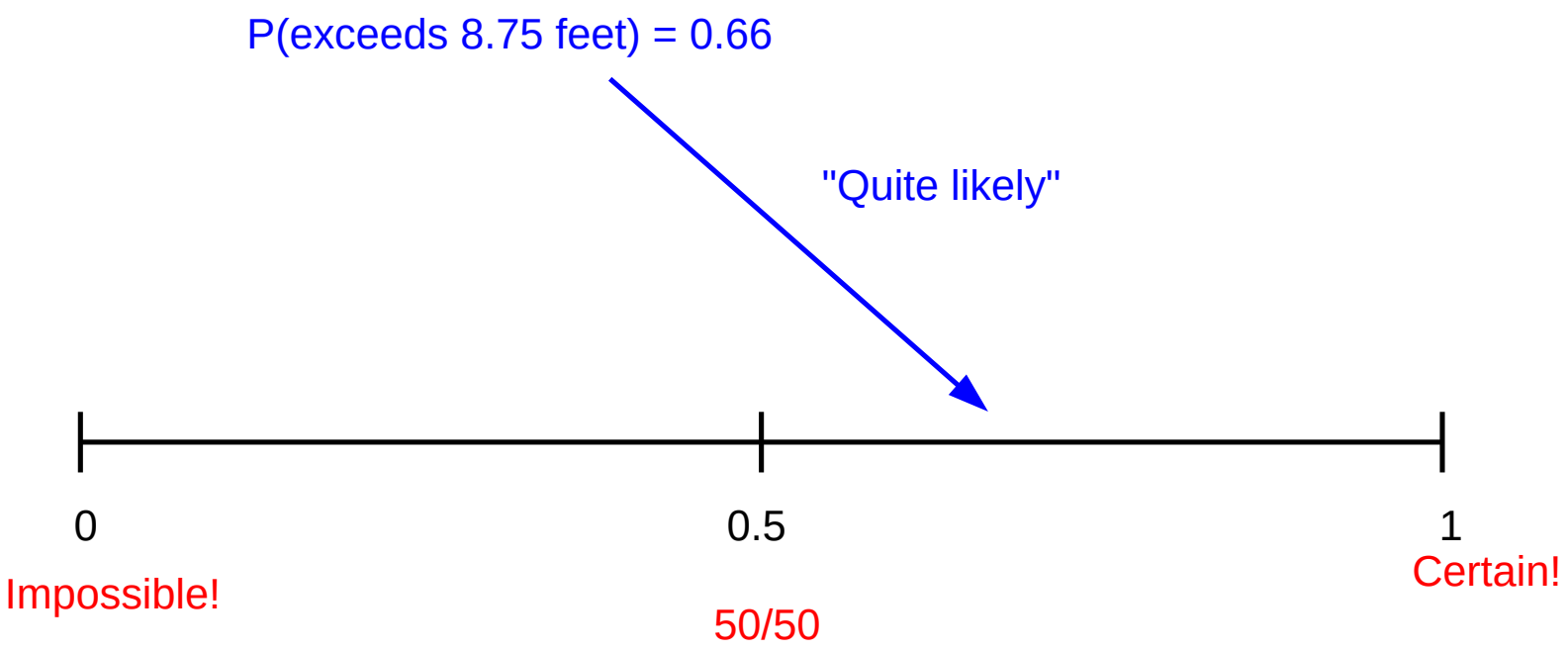
Part A: The relative frequency approach

Think about the **probability scale**:



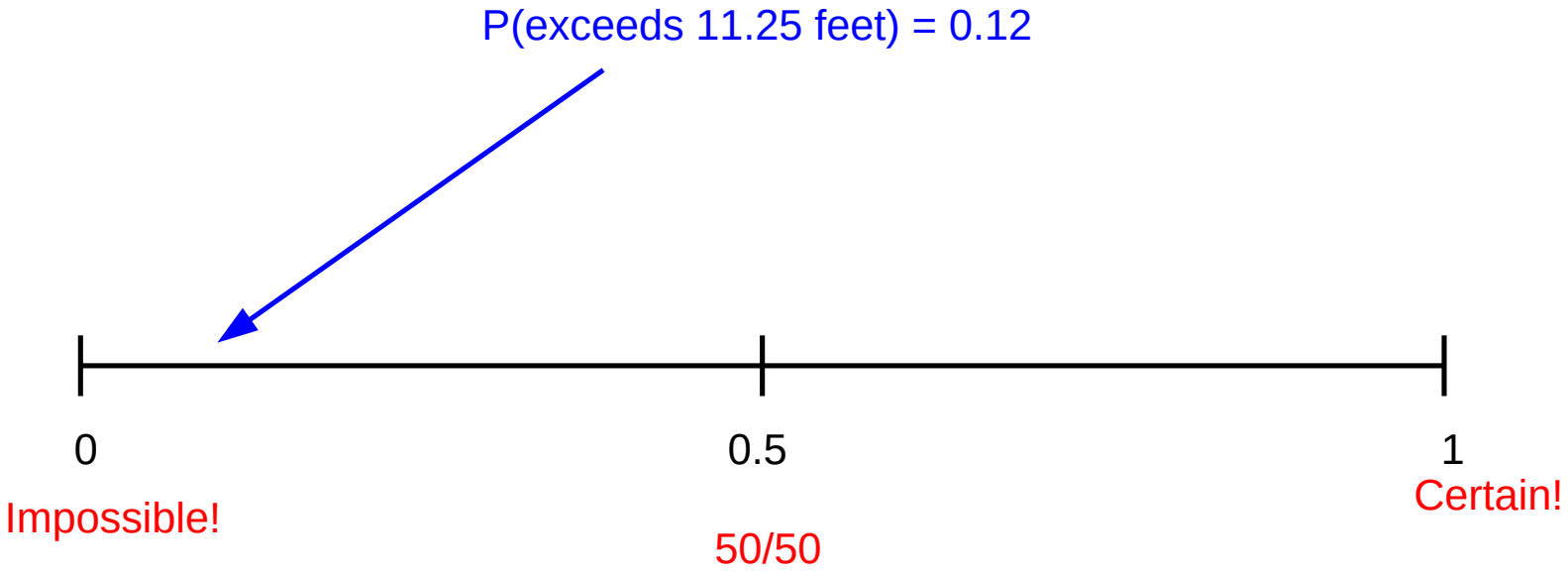
Part A: The relative frequency approach

Think about the **probability scale**:



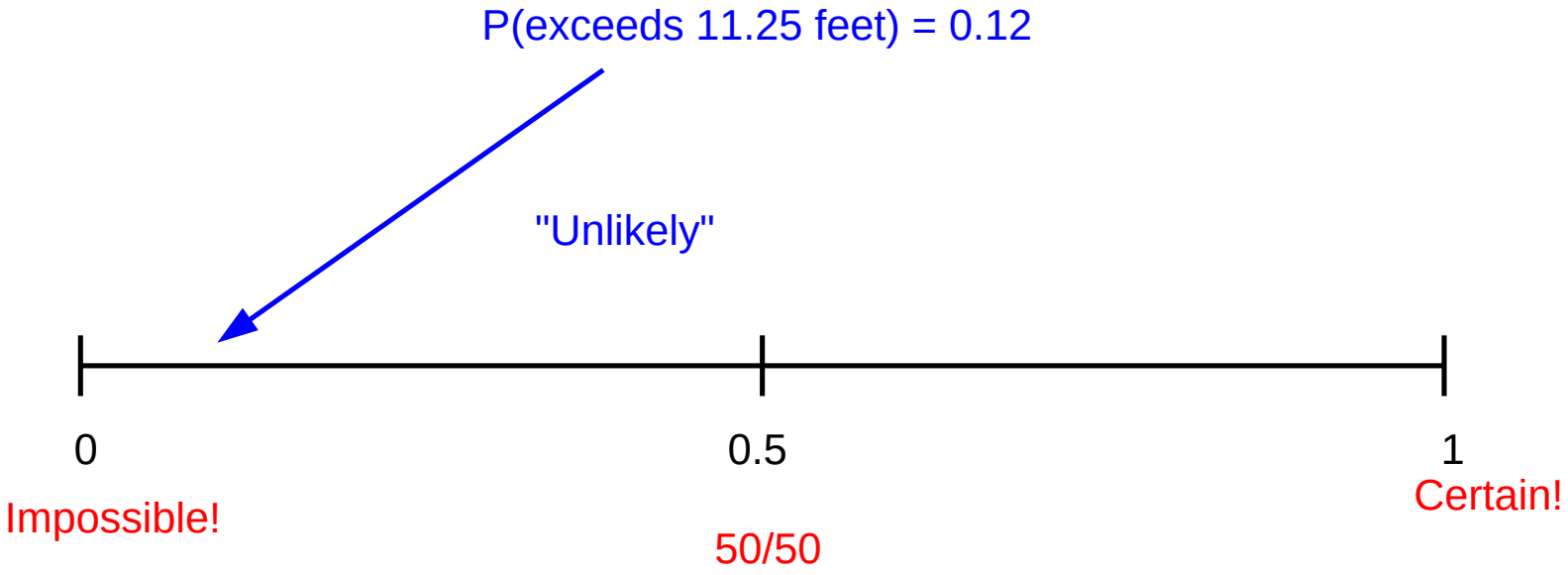
Part A: The relative frequency approach

Think about the **probability scale**:



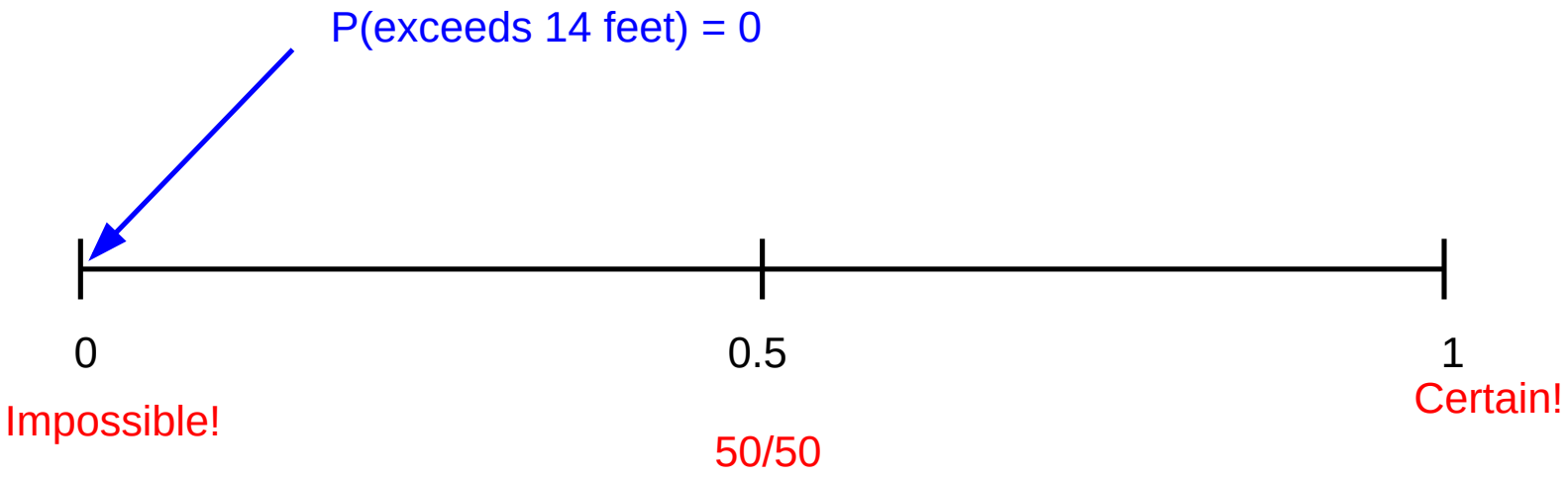
Part A: The relative frequency approach

Think about the **probability scale**:



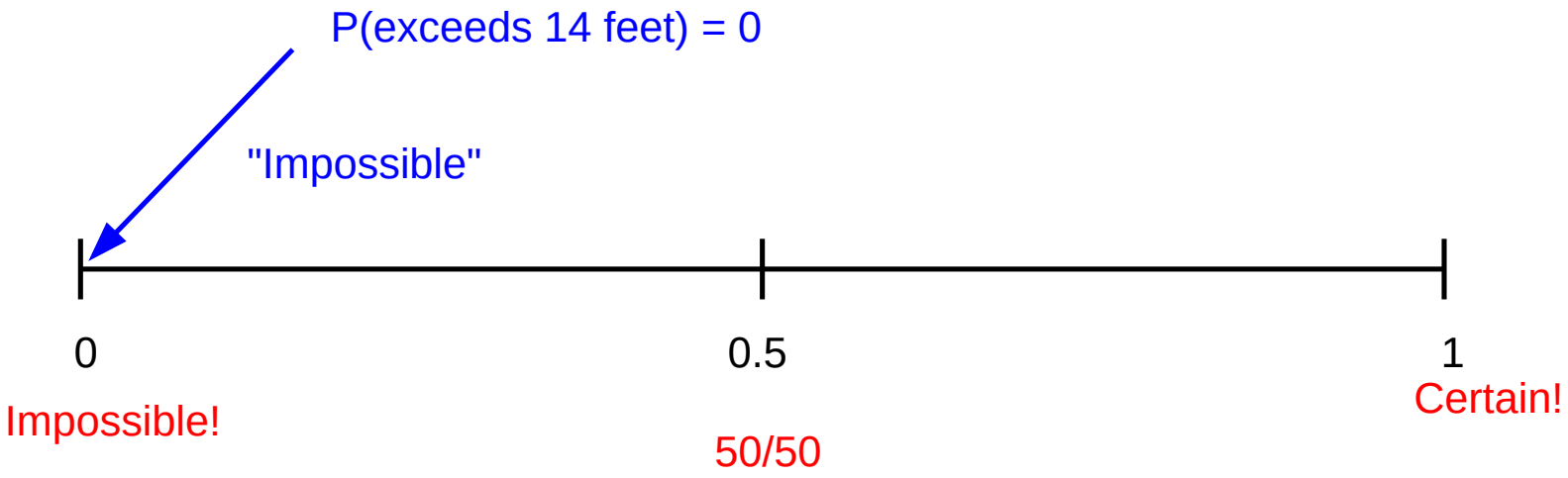
Part A: The relative frequency approach

Think about the **probability scale**:



Part A: The relative frequency approach

Think about the **probability scale**:



Part A: The relative frequency approach

What is wrong with the exceedance probability associated with a sea–surge of 14 feet?

Our answer suggests a sea–surge of more than 14 feet is impossible – however, this *did* happen during Hurricane Katrina (sea–surges reached 14.4 feet!).

Probability models provide a way forward here.

Part B: A probability model for extremes

Before we thing about **probability models**, first of all let's consider exceedance probabilities for a range of values:

Sea–surge	No. of exceedances	Exceedance probability
6.5	50	$50/50 = 1.00$
7.0	48	$48/50 = 0.96$
7.5	44	$44/50 = 0.88$
8.0	41	$41/50 = 0.82$
8.5	36	$36/50 = 0.72$
9.0		
9.5		
10.0		
10.5		
11.0	7	$7/50 = 0.14$
11.5	5	$5/50 = 0.10$
12.0	4	$4/50 = 0.08$
12.5	3	$3/50 = 0.06$

Part B: A probability model for extremes

Before we thing about **probability models**, first of all let's consider exceedance probabilities for a range of values:

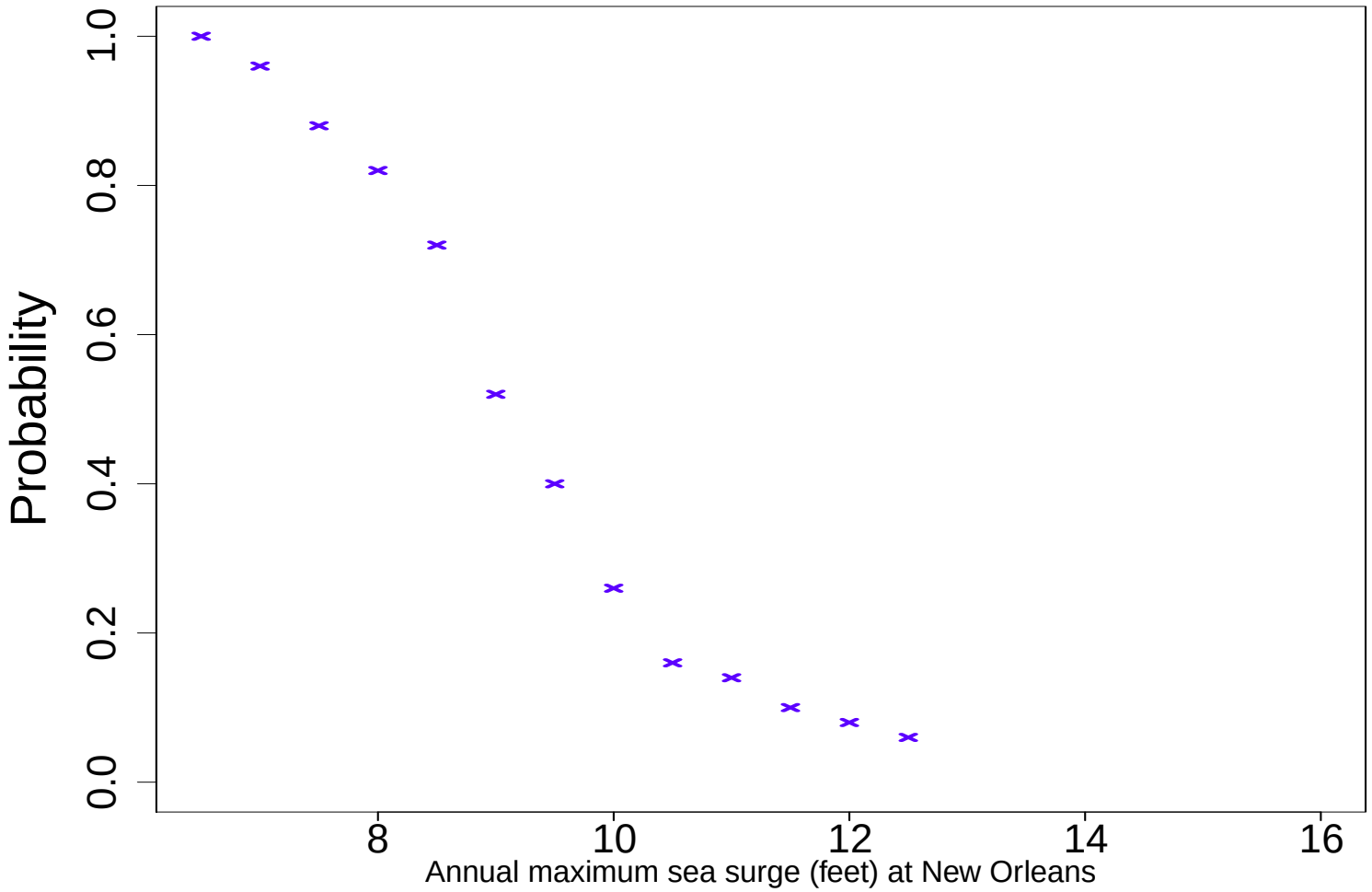
Sea–surge	No. of exceedances	Exceedance probability
6.5	50	$50/50 = 1.00$
7.0	48	$48/50 = 0.96$
7.5	44	$44/50 = 0.88$
8.0	41	$41/50 = 0.82$
8.5	36	$36/50 = 0.72$
9.0	**26**	$\mathbf{26/50 = 0.52}$
9.5	**20**	$\mathbf{20/50 = 0.40}$
10.0	**13**	$\mathbf{13/50 = 0.26}$
10.5	**8**	$\mathbf{8/50 = 0.16}$
11.0	7	$7/50 = 0.14$
11.5	5	$5/50 = 0.10$
12.0	4	$4/50 = 0.08$
12.5	3	$3/50 = 0.06$

Part B: A probability model for extremes

Before we thing about **probability models**, first of all let's consider exceedance probabilities for a range of values:

Sea–surge	No. of exceedances	Exceedance probability
6.5	50	$50/50 =$ **1.00**
7.0	48	$48/50 =$ **0.96**
7.5	44	$44/50 =$ **0.88**
8.0	41	$41/50 =$ **0.82**
8.5	36	$36/50 =$ **0.72**
9.0	26	$26/50 =$ **0.52**
9.5	20	$20/50 =$ **0.40**
10.0	13	$13/50 =$ **0.26**
10.5	8	$8/50 =$ **0.16**
11.0	7	$7/50 =$ **0.14**
11.5	5	$5/50 =$ **0.10**
12.0	4	$4/50 =$ **0.08**
12.5	3	$3/50 =$ **0.06**

Part B: A probability model for extremes

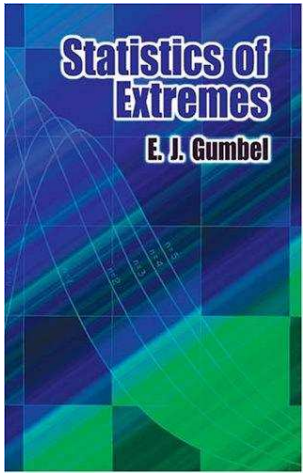


Part B: A probability model for extremes

Between 1920 and the mid–1950s, some eminent mathematicians developed probability models for extremes.

The aim was to provide a **mathematical formula** that could "predict" the exceedance probabilities of real–life extremes, as in the table we've just drawn up for the New Orleans data.

One of these mathematicians was **Emil Julius Gumbel**.



Part B: A probability model for extremes

The so–called **Gumbel model** for the exceedance probabilities of extremes is given by the following formula:

$$\boldsymbol{P}(X > x) = 1 - \exp\left[-\exp\left\{-\left(\frac{x-\mu}{\sigma}\right)\right\}\right],$$

where

- μ is the "**location**" parameter
- σ is the "**scale**" parameter
- X is our "**variable**"
- x is the **value** of our variable
- "exp" is the **exponential function** (button on calculator)

Part B: A probability model for extremes

Using techniques students learn about at University, we can use the data to estimate the values of the location and scale parameters.

Doing so, gives

$$\mu = 8.536 \qquad \text{and} \qquad \sigma = 1.2$$

for our dataset. Plugging these into Gumbel's formula gives

$$\boldsymbol{P}(X > x) = 1 - \exp\left[-\exp\left\{-\left(\frac{x - 8.536}{1.2}\right)\right\}\right].$$

Then we can work out the exceedance probabilities for particular values x. For example, for $x = 7.5$, we get:

$$\boldsymbol{P}(X > 7.5) = 1 - \exp\left[-\exp\left\{-\left(\frac{7.5 - 8.536}{1.2}\right)\right\}\right] = \mathbf{0.907}.$$

Part B: A probability model for extremes

We can compare the values from Gumbel's formula to those we obtained from the data itself. For example:

$$\boldsymbol{P}(X > \mathbf{10}) = 1 - \exp\left[-\exp\left\{-\left(\frac{\mathbf{10} - 8.536}{1.2}\right)\right\}\right] = \mathbf{0.26}$$

(= 0.26 from the data)

$$\boldsymbol{P}(X > \mathbf{11.5}) = 1 - \exp\left[-\exp\left\{-\left(\frac{\mathbf{11.5} - 8.536}{1.2}\right)\right\}\right] = \mathbf{0.08}$$

(= 0.1 from the data)

$$\boldsymbol{P}(X > \mathbf{14}) = 1 - \exp\left[-\exp\left\{-\left(\frac{\mathbf{14} - 8.536}{1.2}\right)\right\}\right] = \mathbf{0.01}$$

(= 0 from the data – impossible!)

Part B: A probability model for extremes

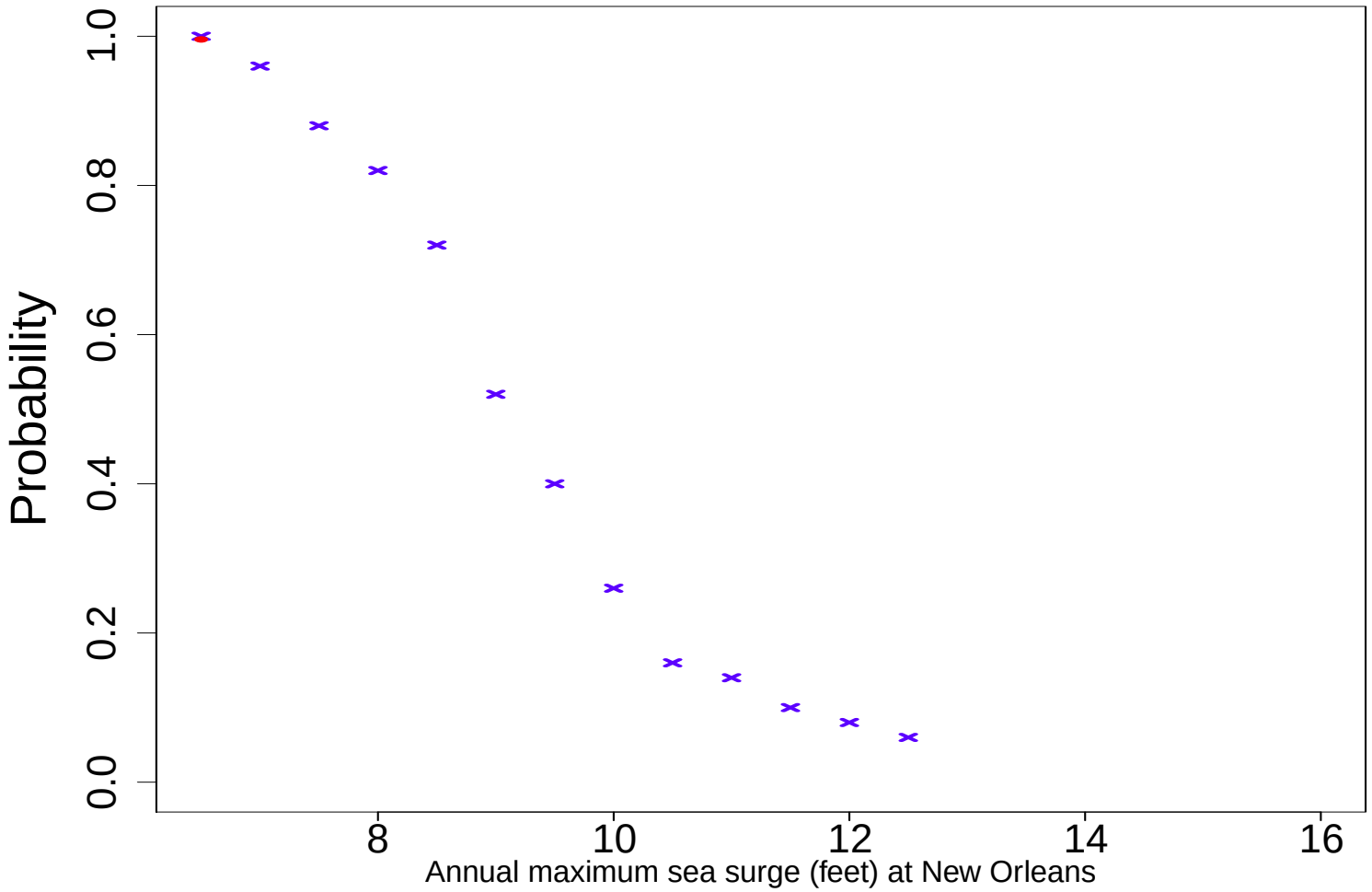
We could use Gumbel's formula for lots of different values of x and then plot them on the same graph as the real data.

What do you notice when we do this?

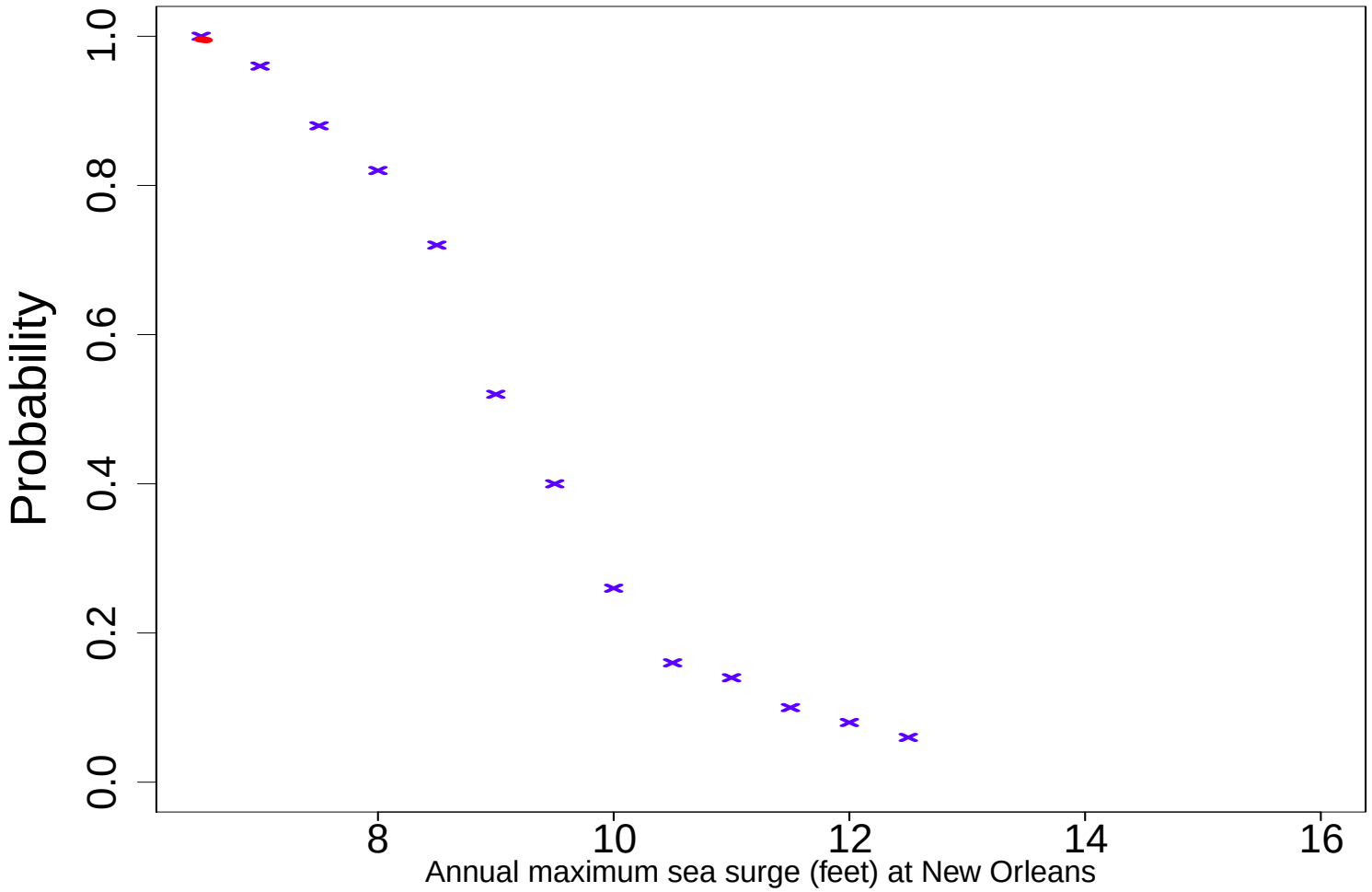
The **Gumbel** model is not the only model to choose from:

- The **Fréchet** model
- The **Weibull** model

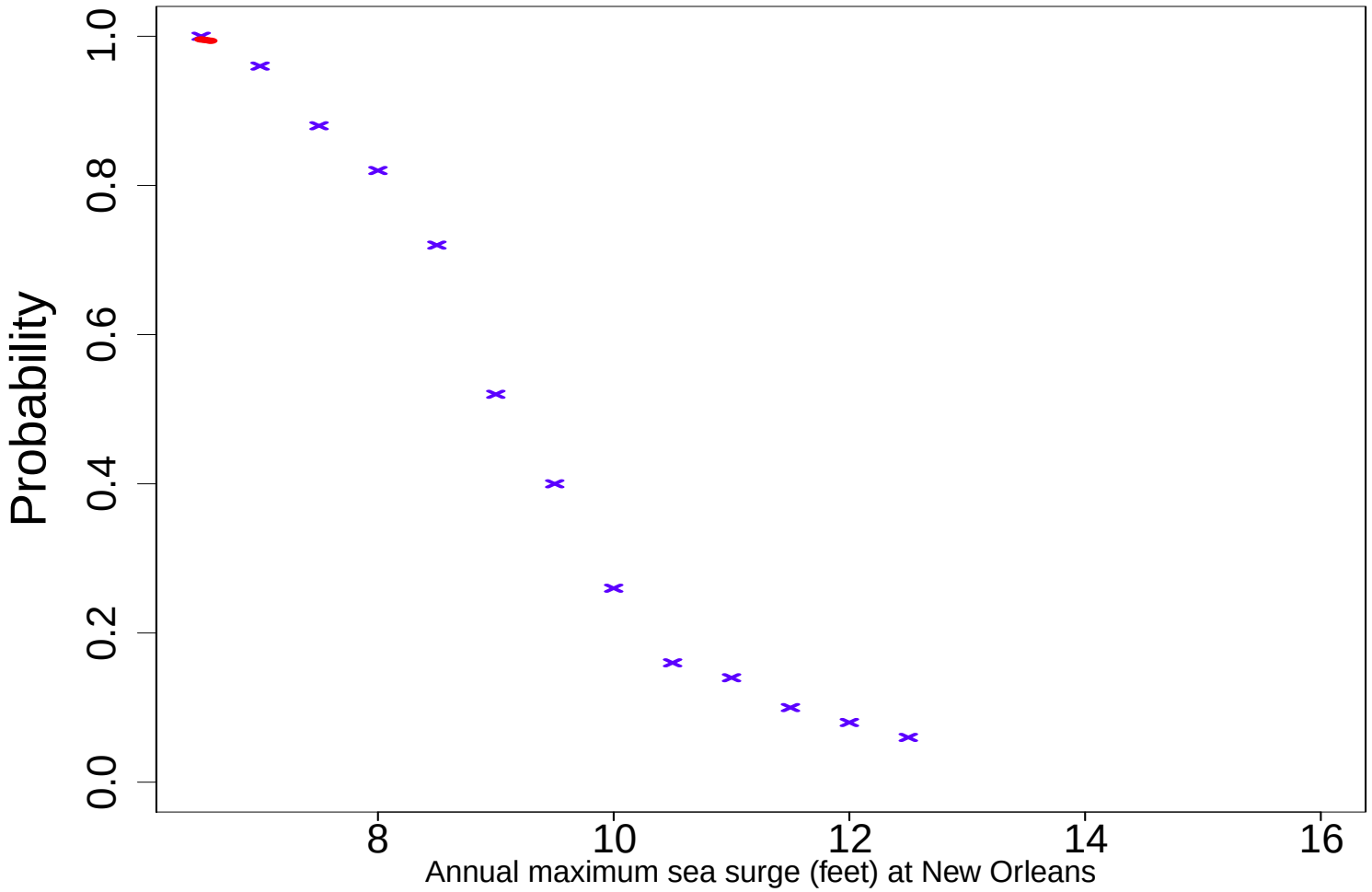
Part B: A probability model for extremes



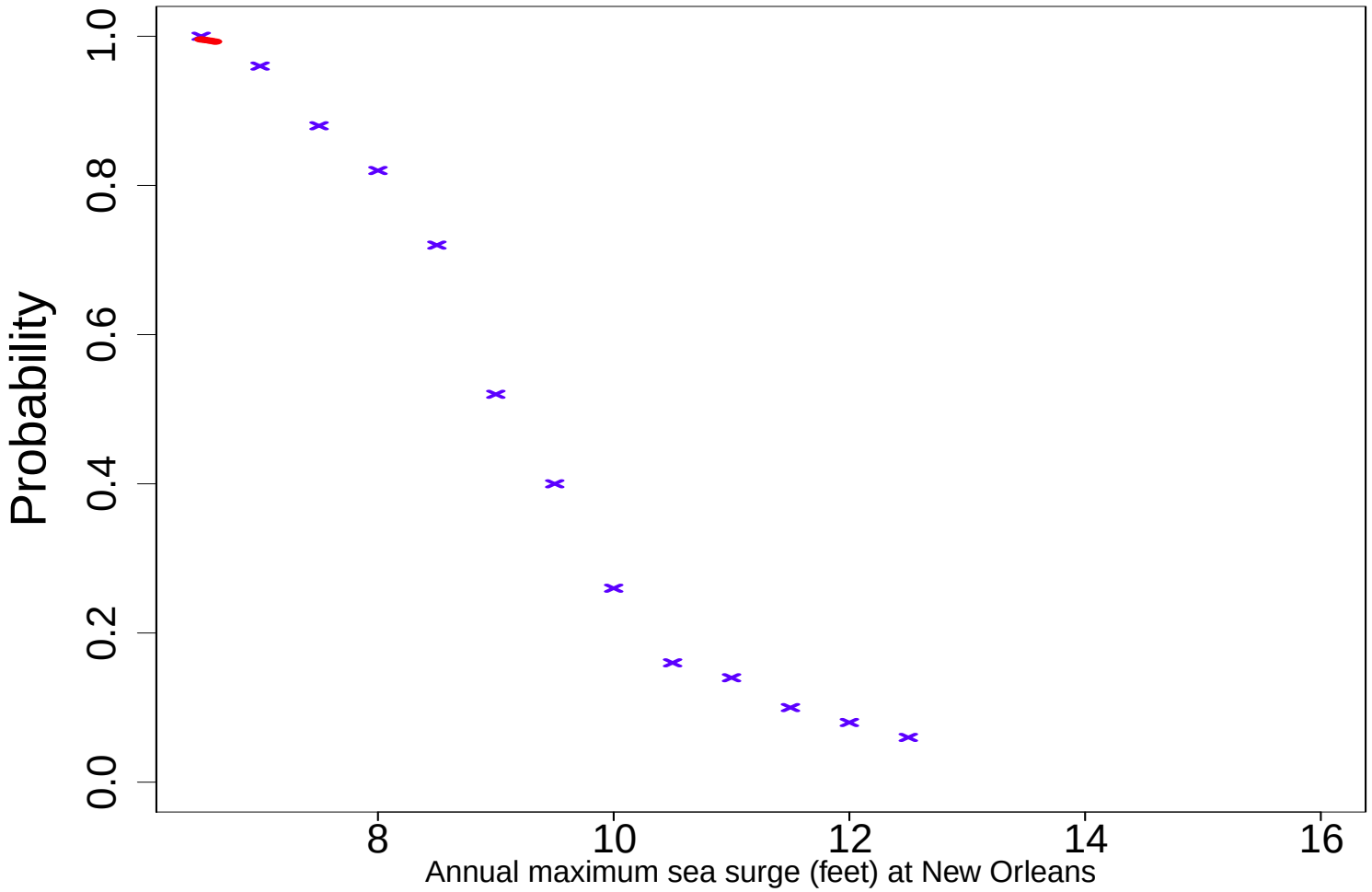
Part B: A probability model for extremes



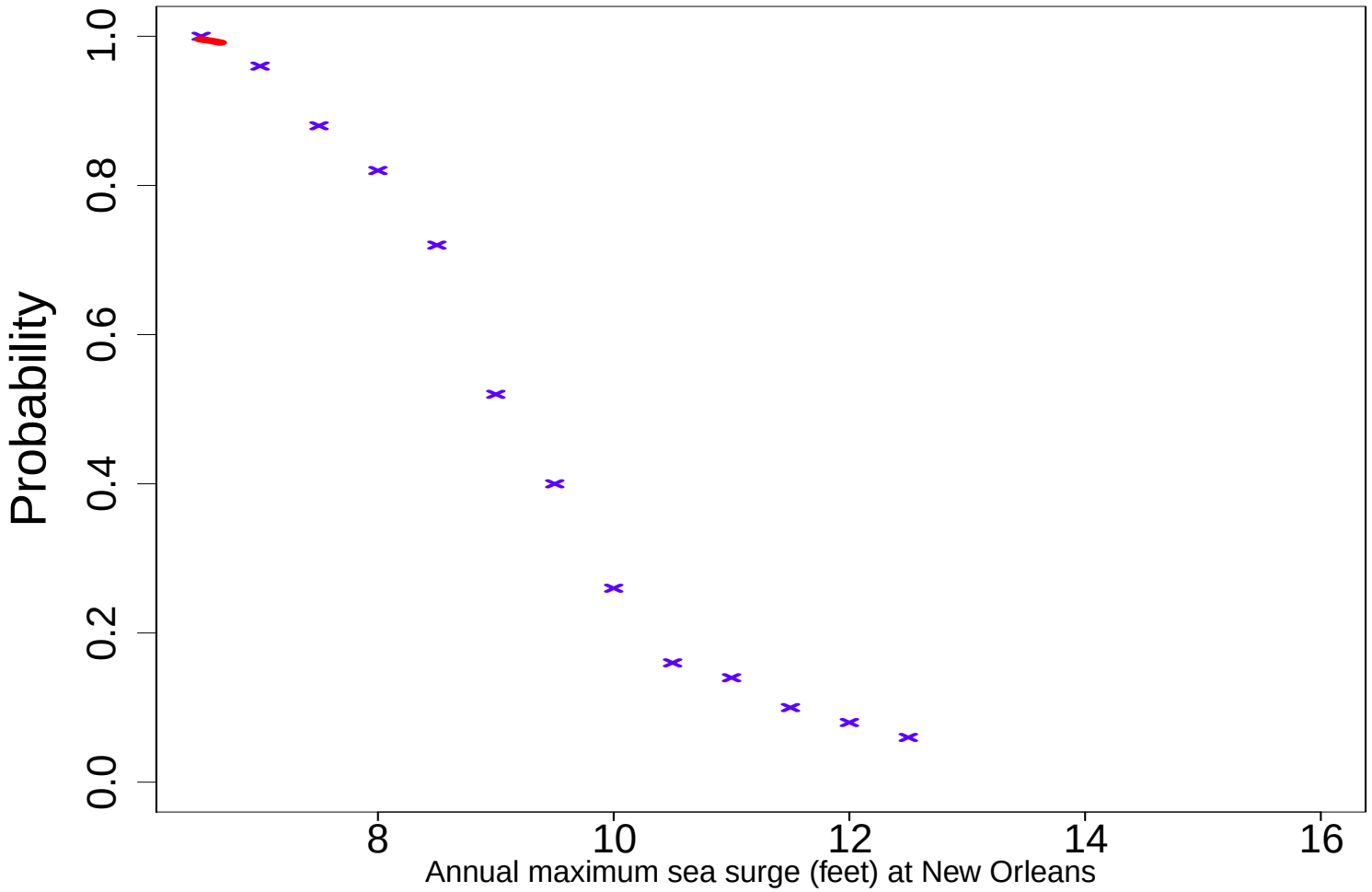
Part B: A probability model for extremes



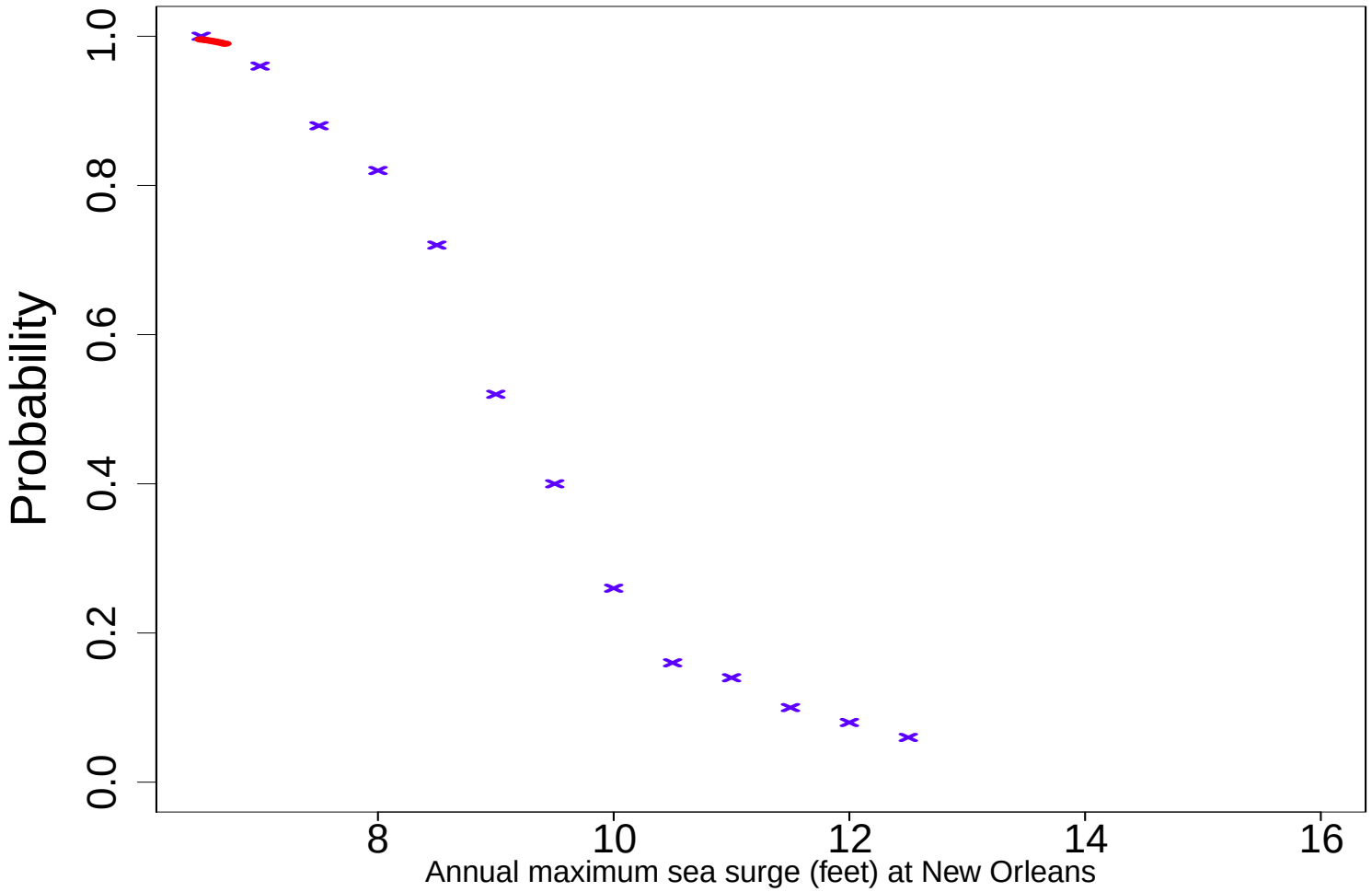
Part B: A probability model for extremes



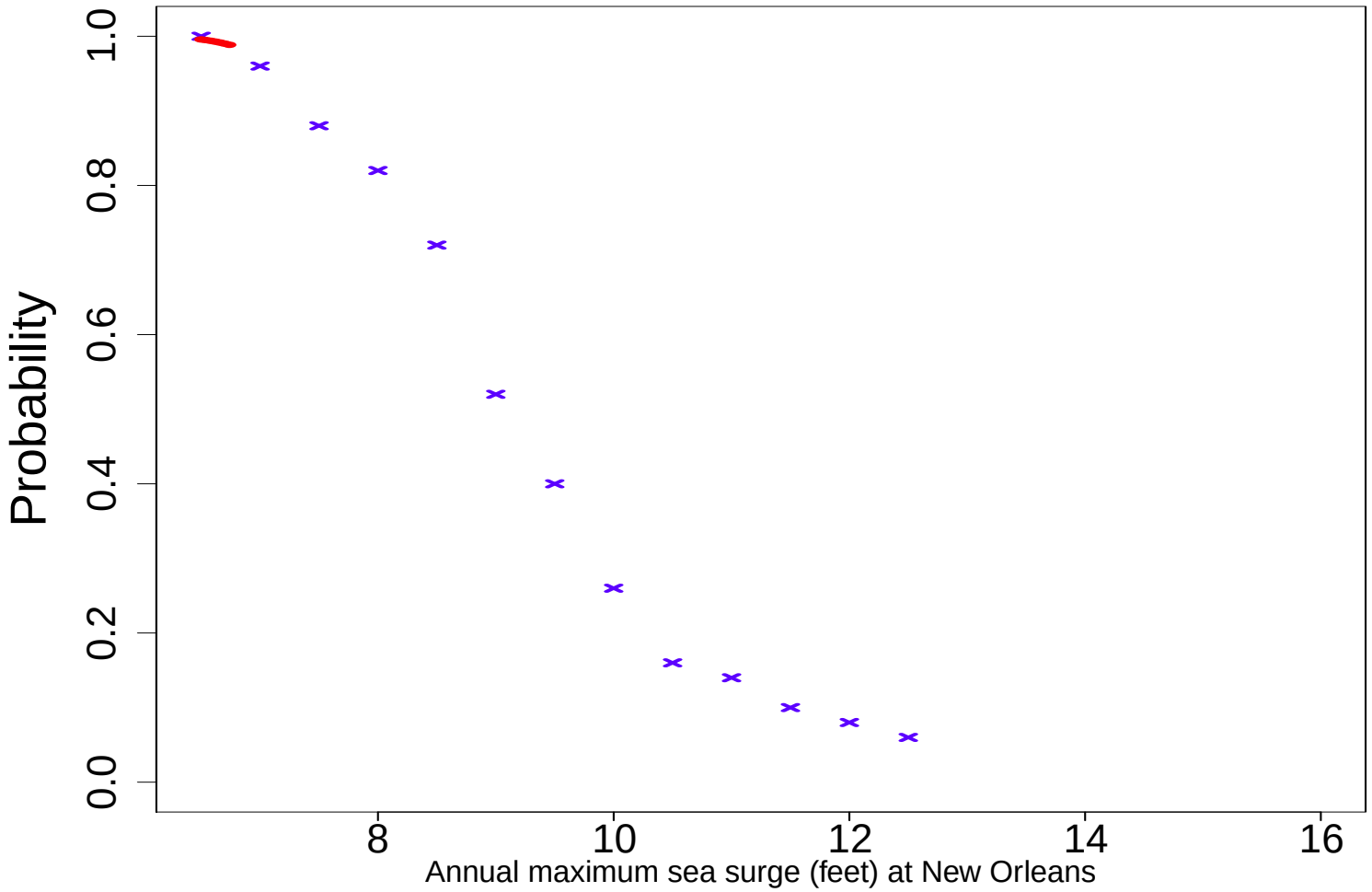
Part B: A probability model for extremes



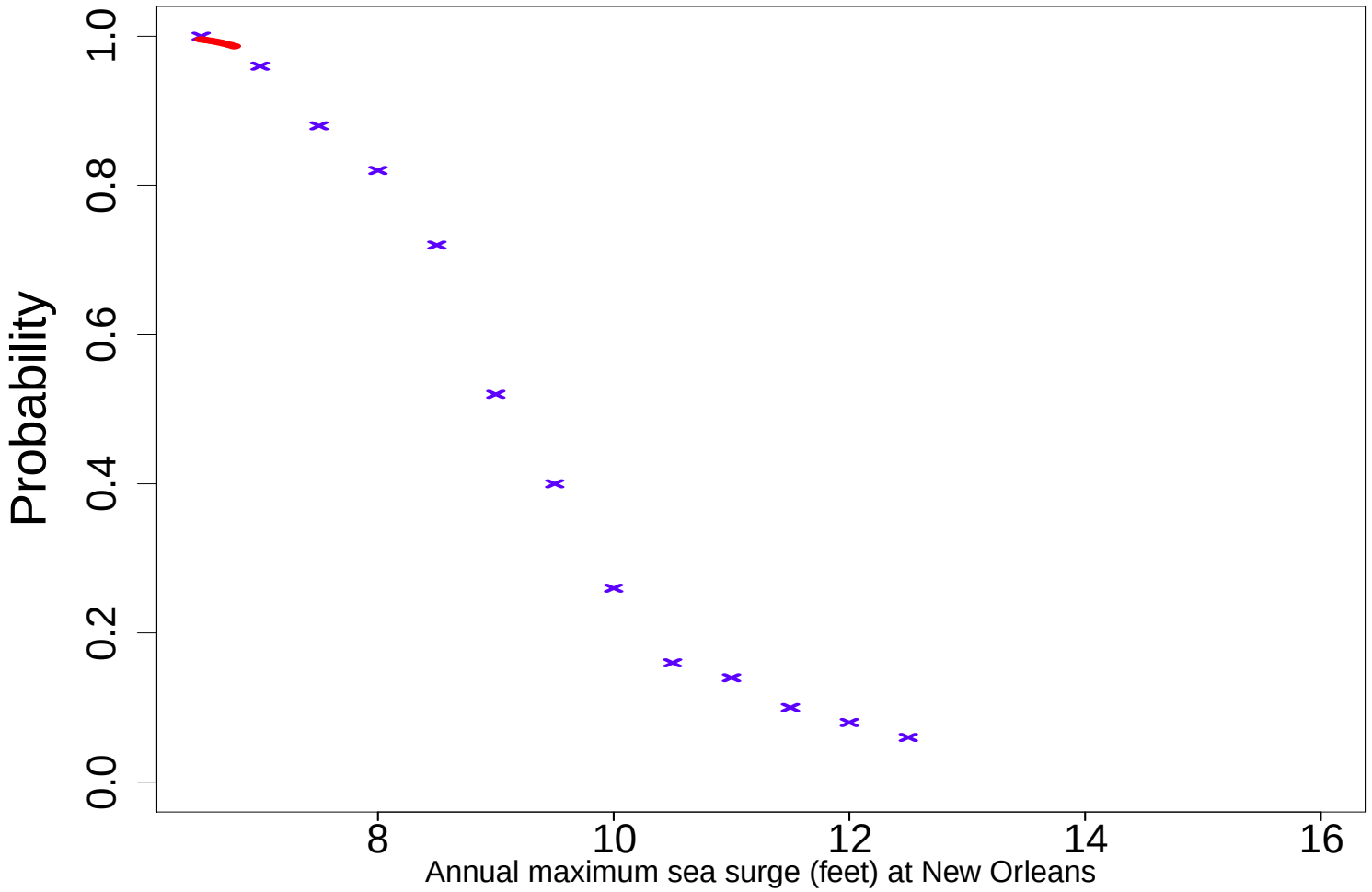
Part B: A probability model for extremes



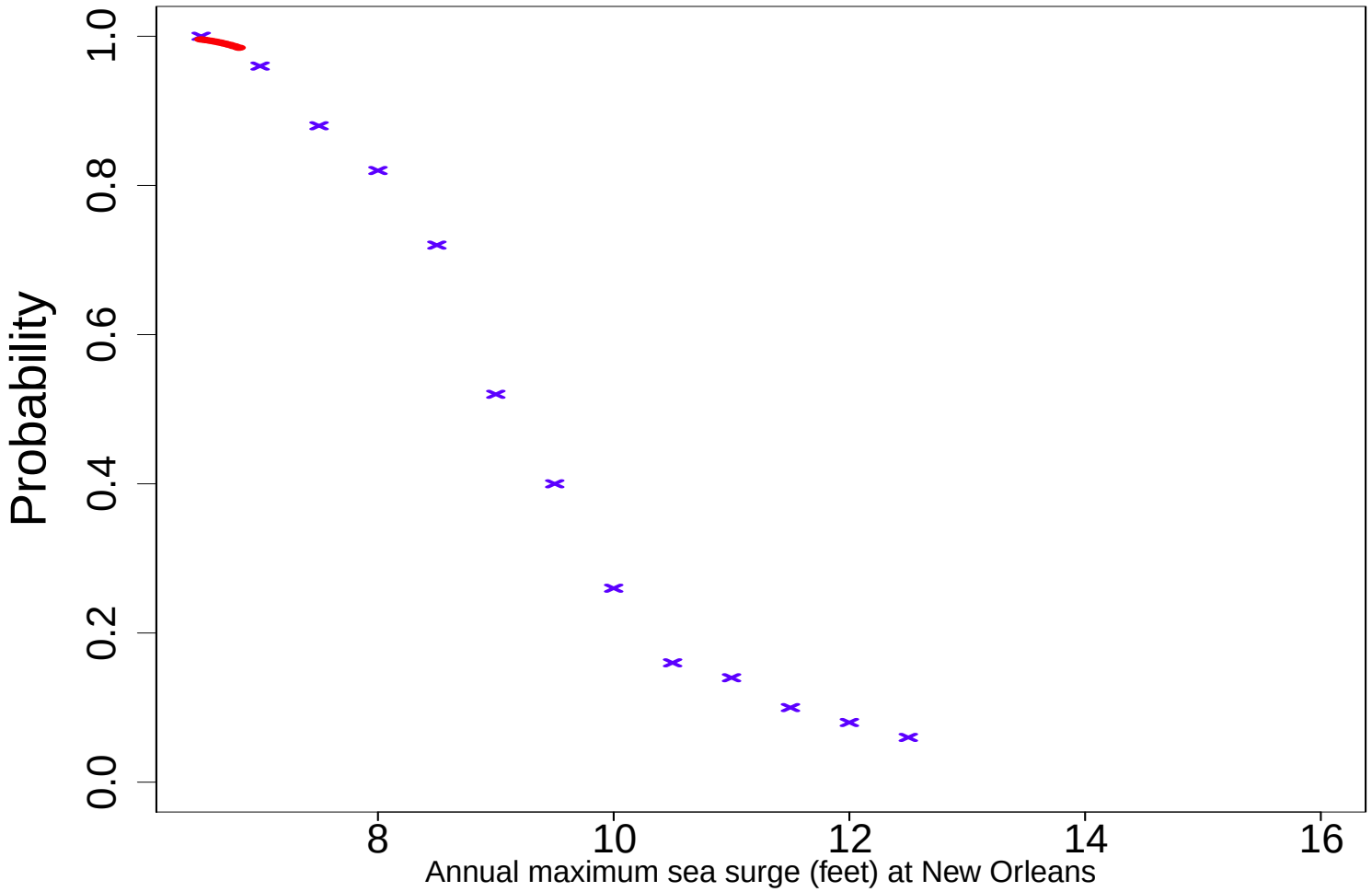
Part B: A probability model for extremes



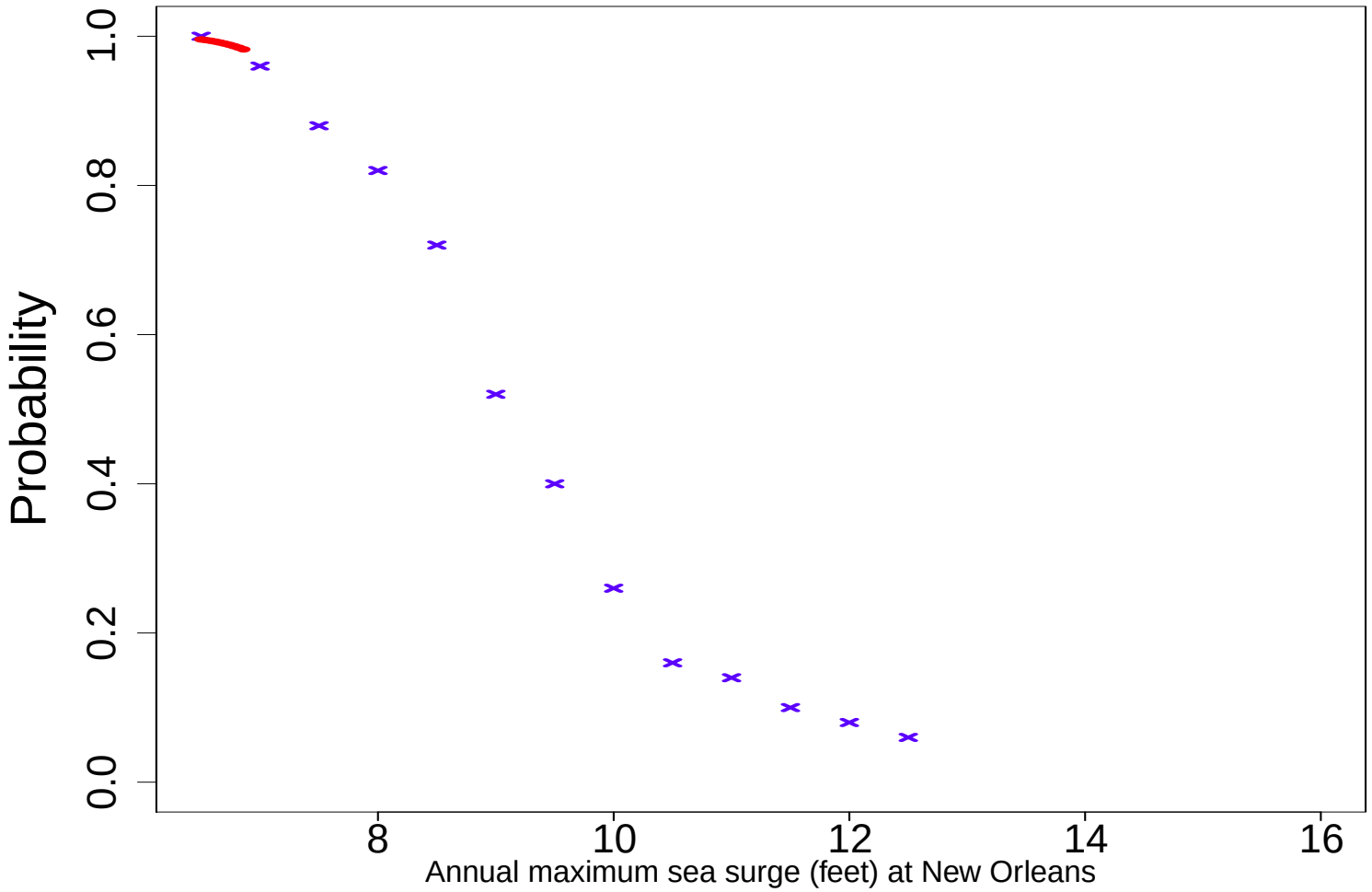
Part B: A probability model for extremes



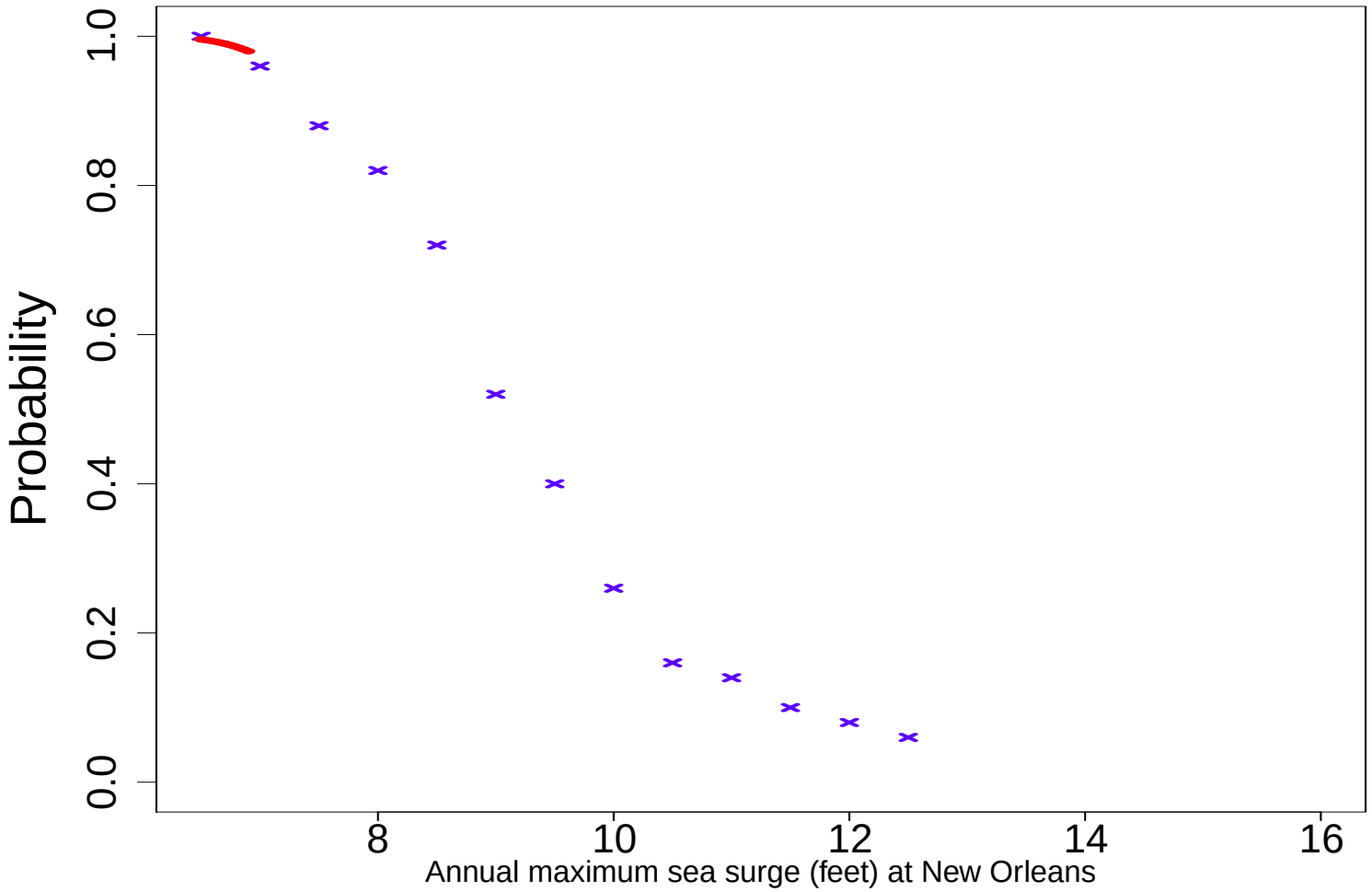
Part B: A probability model for extremes



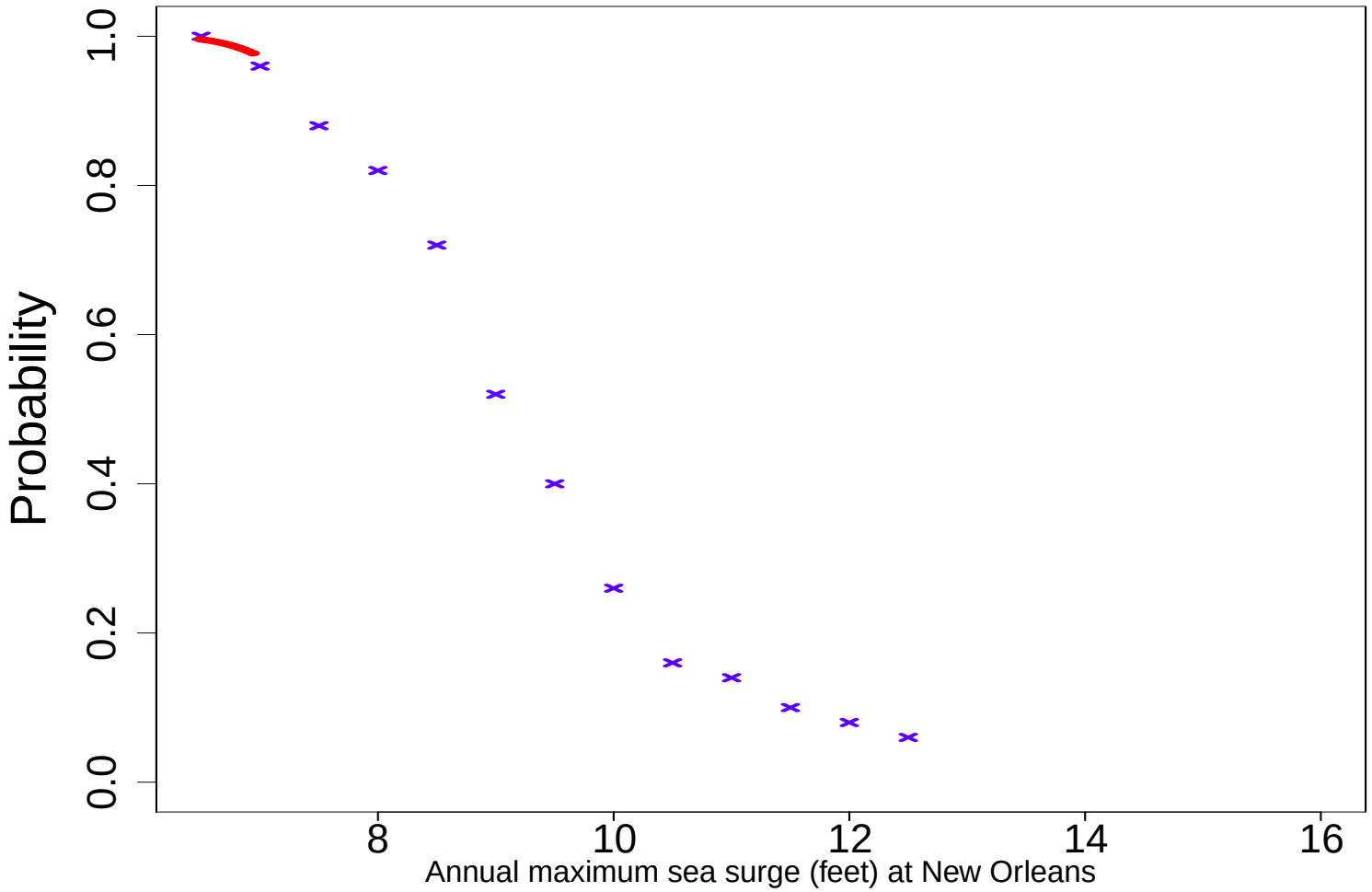
Part B: A probability model for extremes



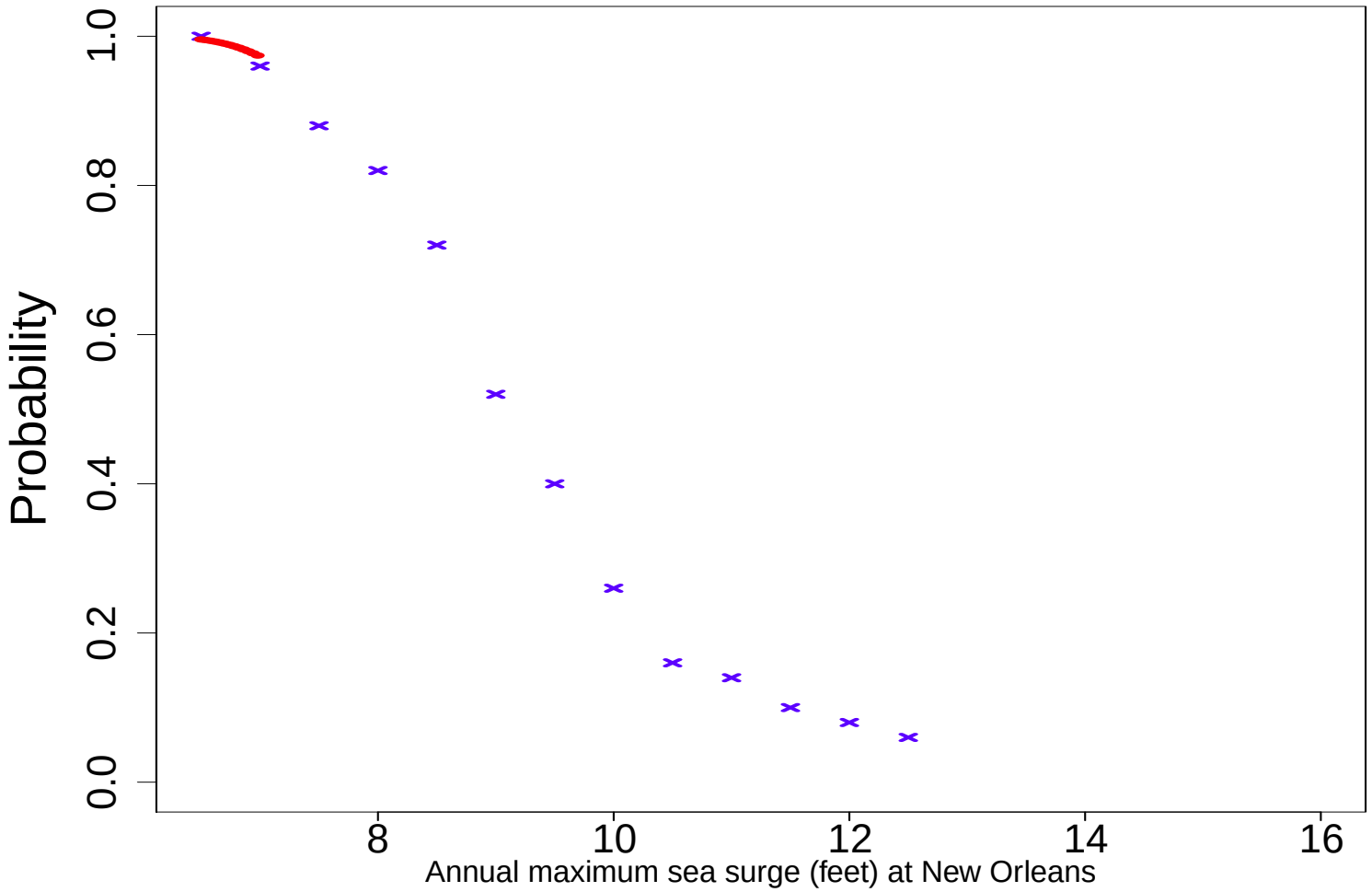
Part B: A probability model for extremes



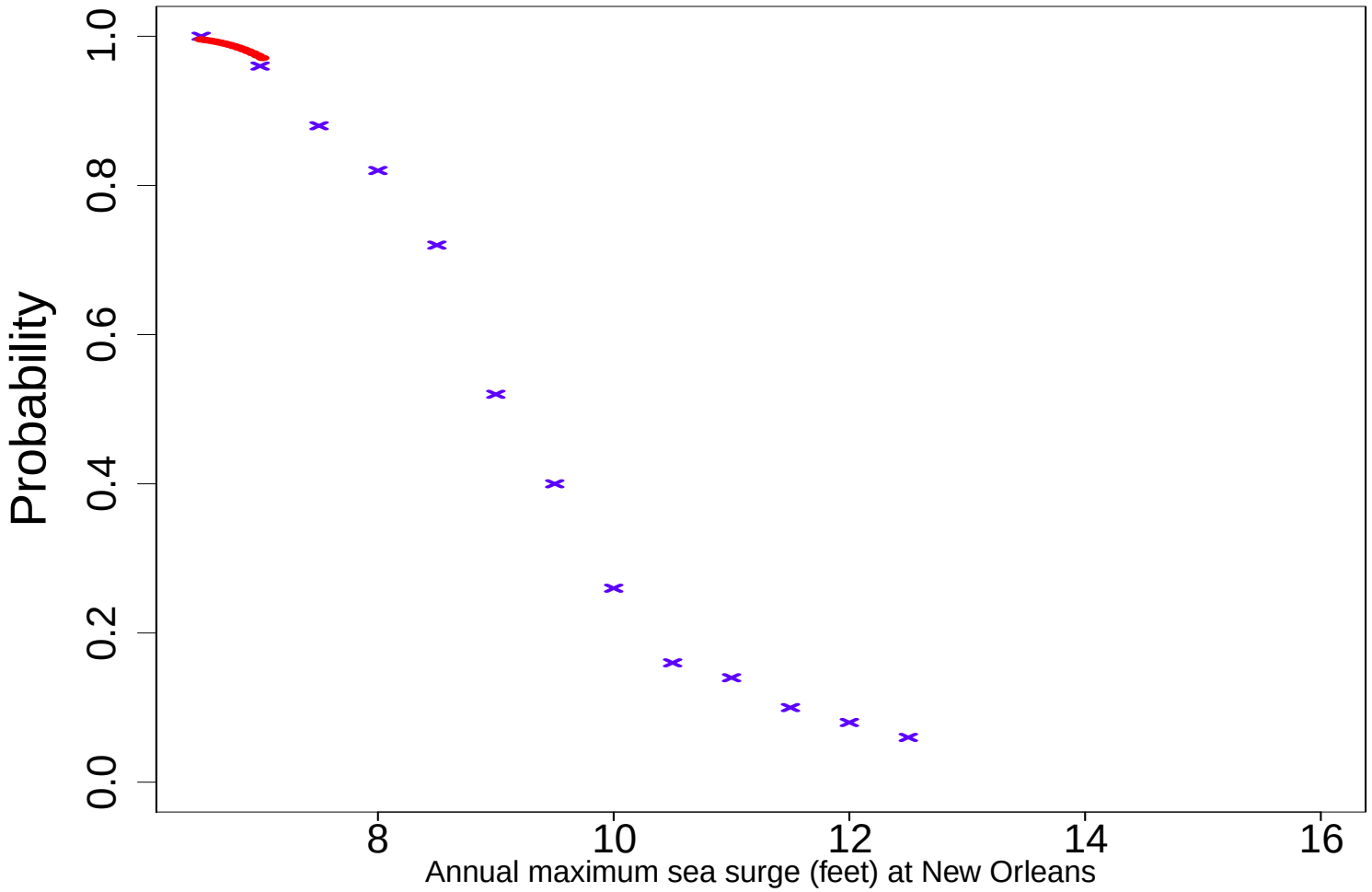
Part B: A probability model for extremes



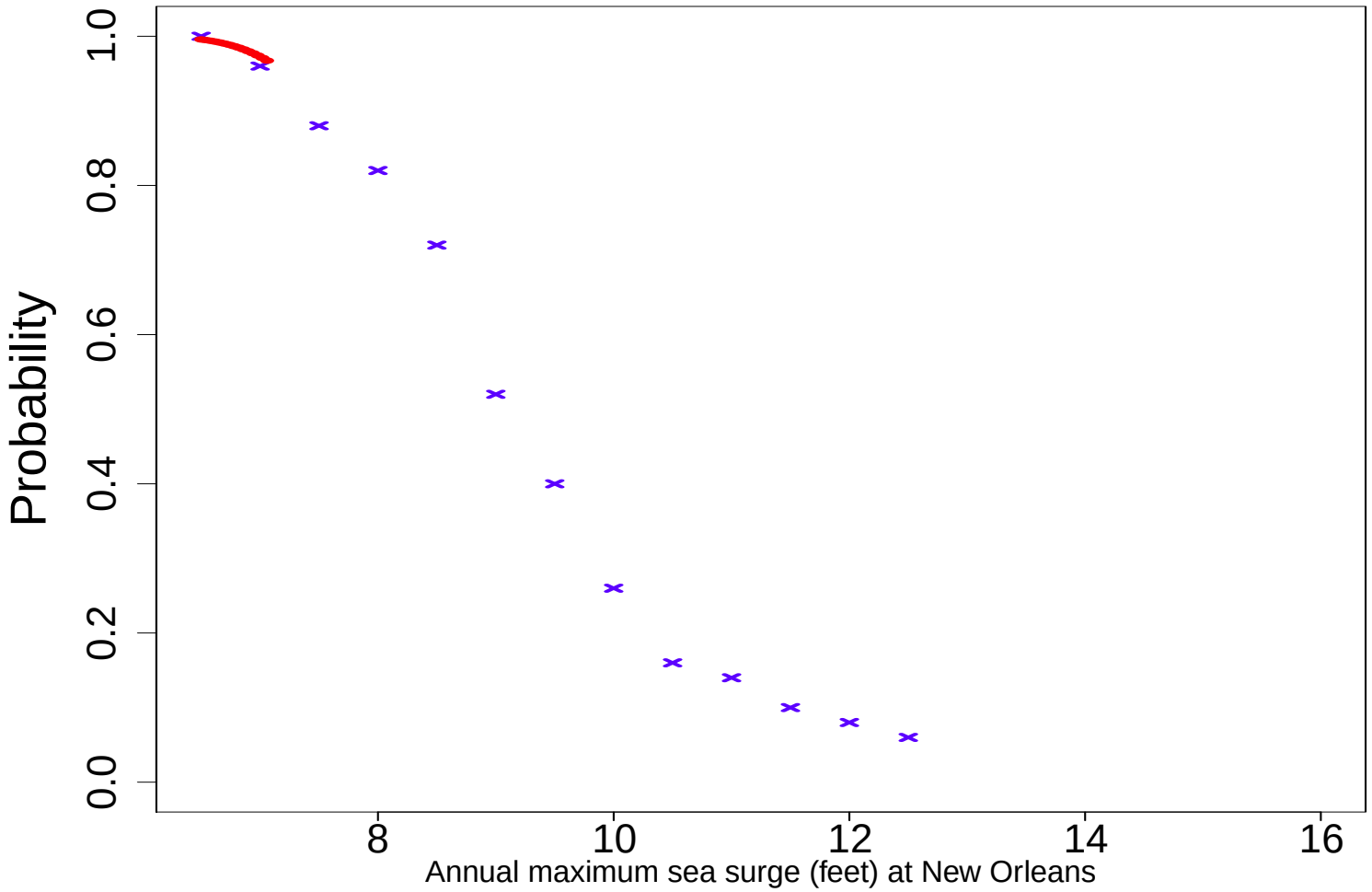
Part B: A probability model for extremes



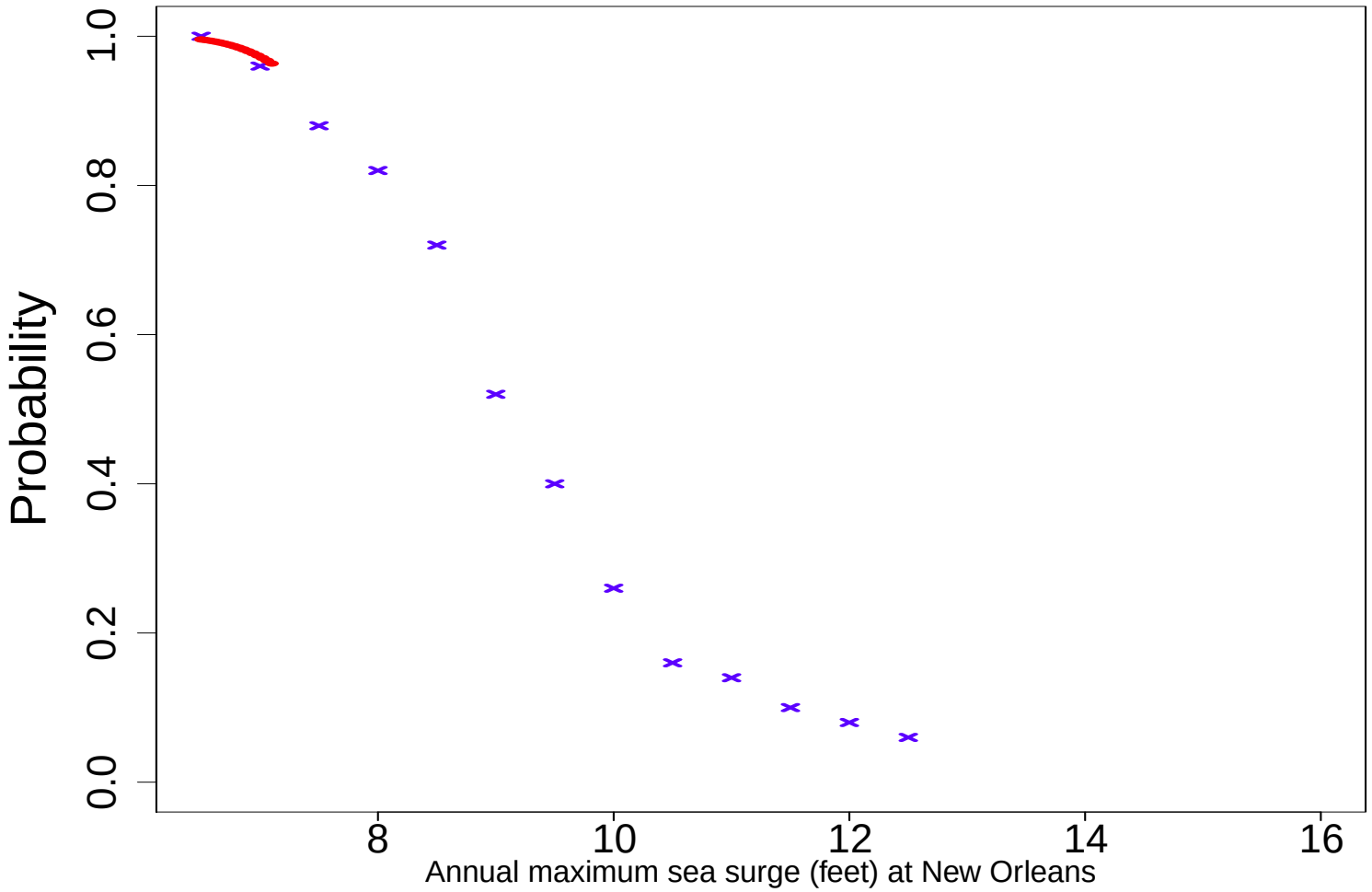
Part B: A probability model for extremes



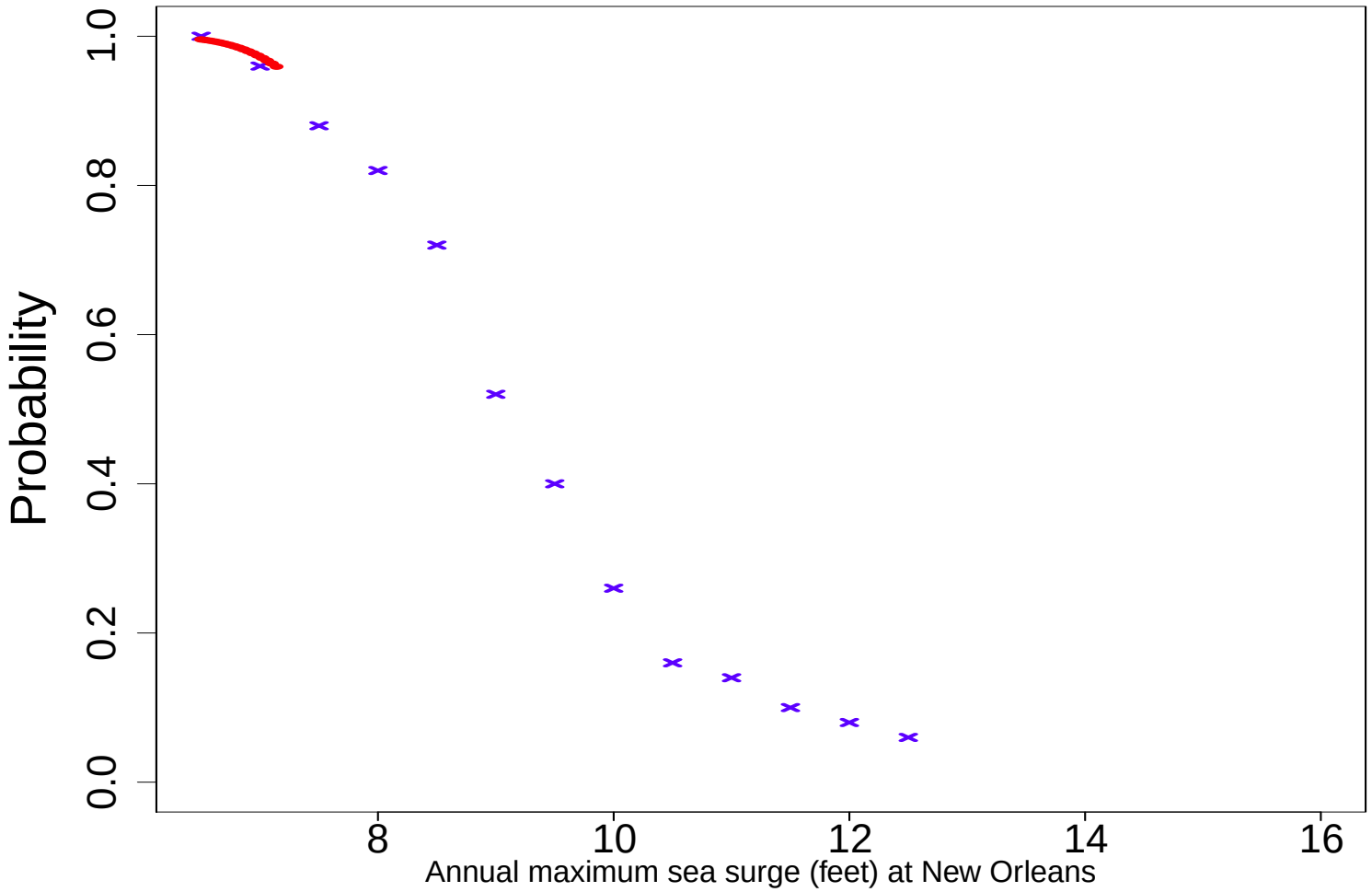
Part B: A probability model for extremes



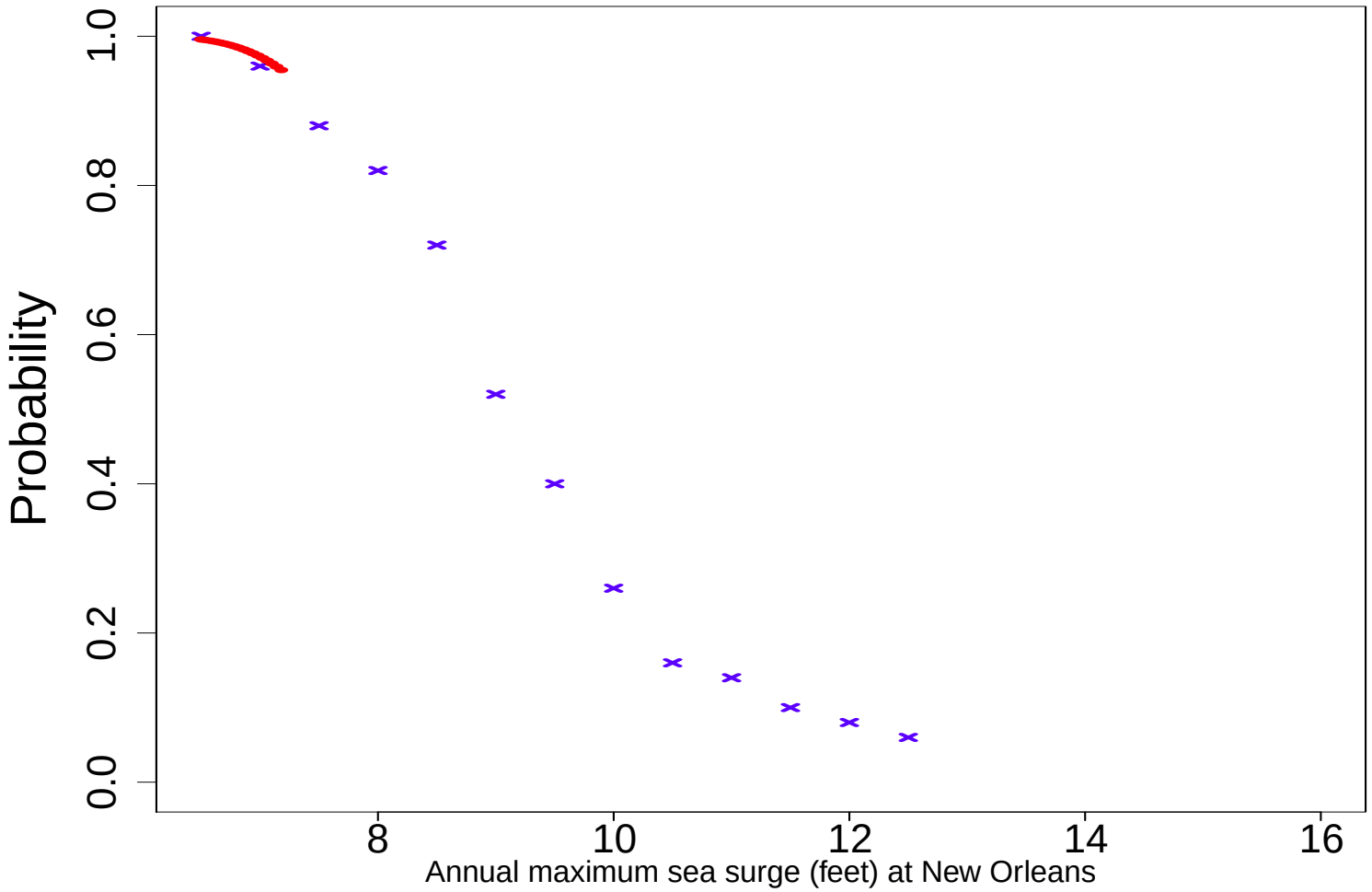
Part B: A probability model for extremes



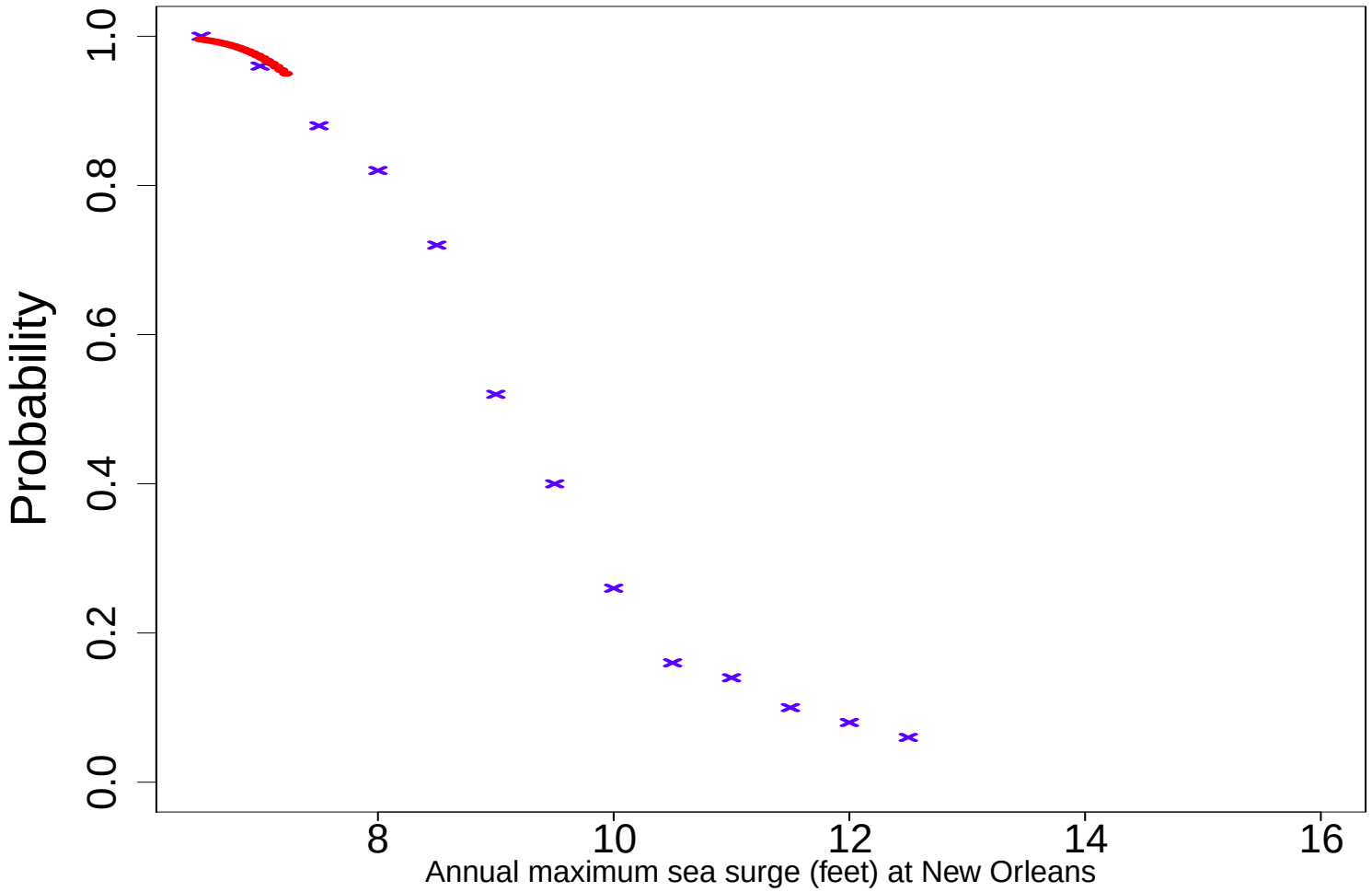
Part B: A probability model for extremes



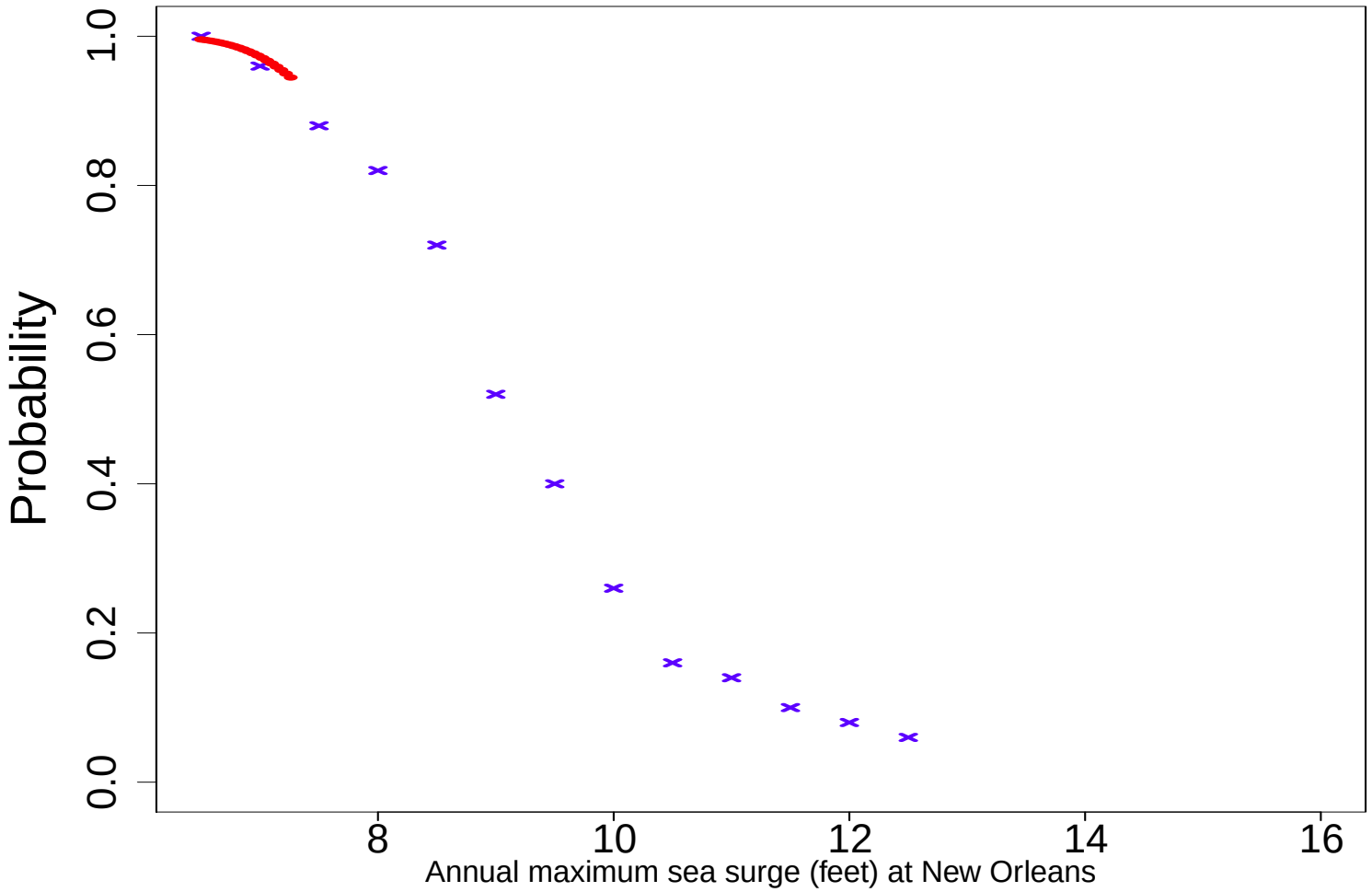
Part B: A probability model for extremes



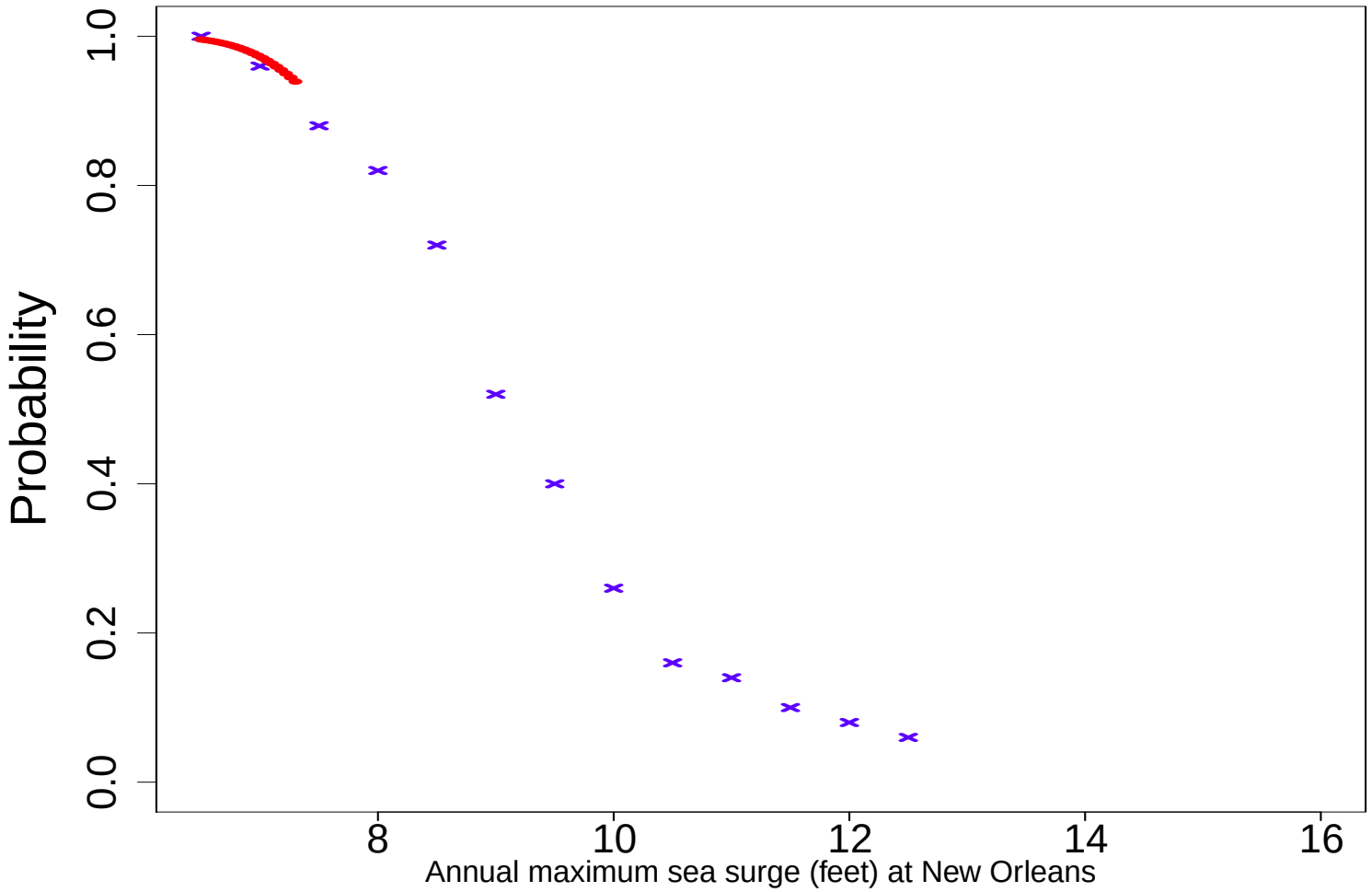
Part B: A probability model for extremes



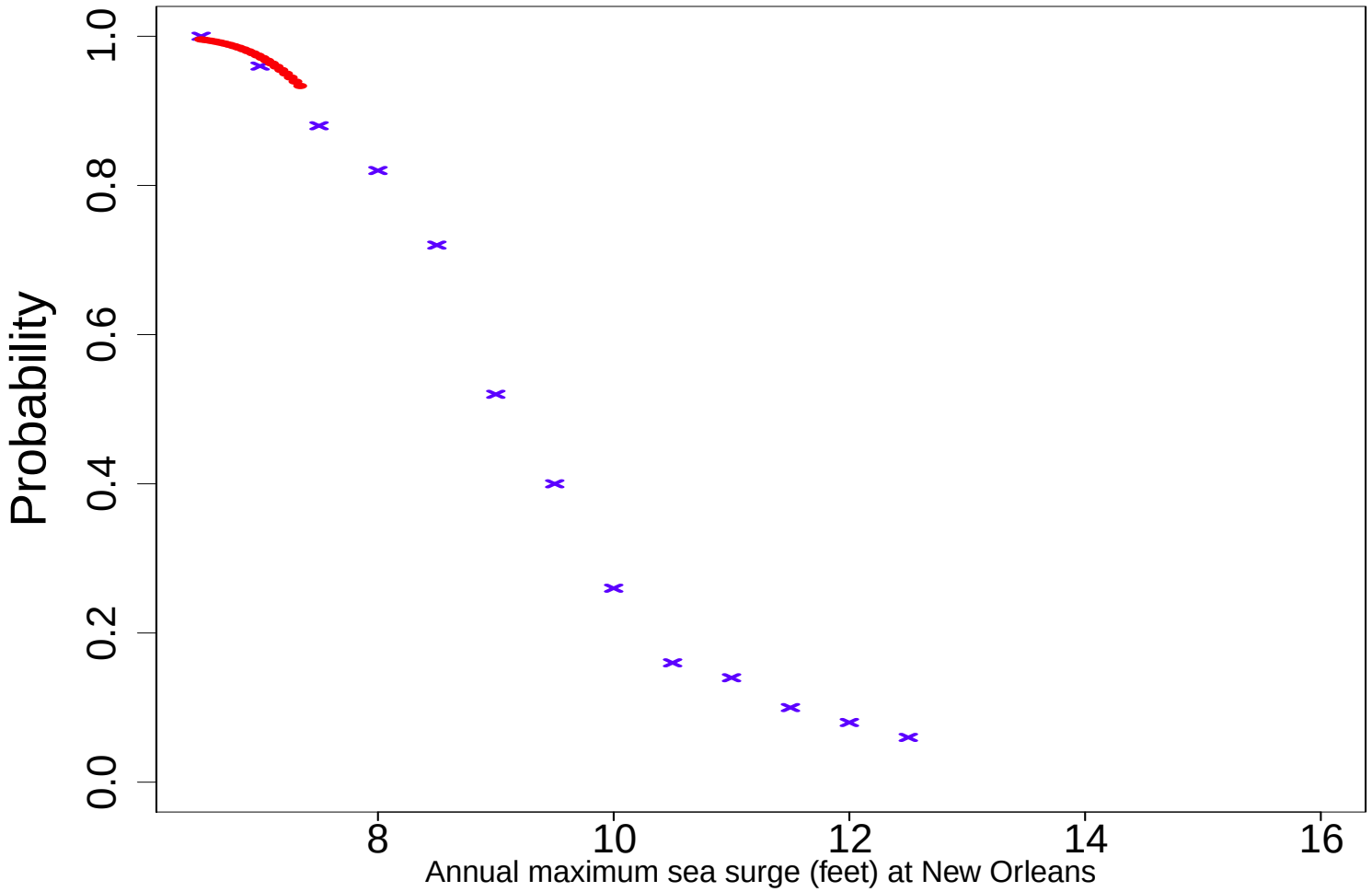
Part B: A probability model for extremes



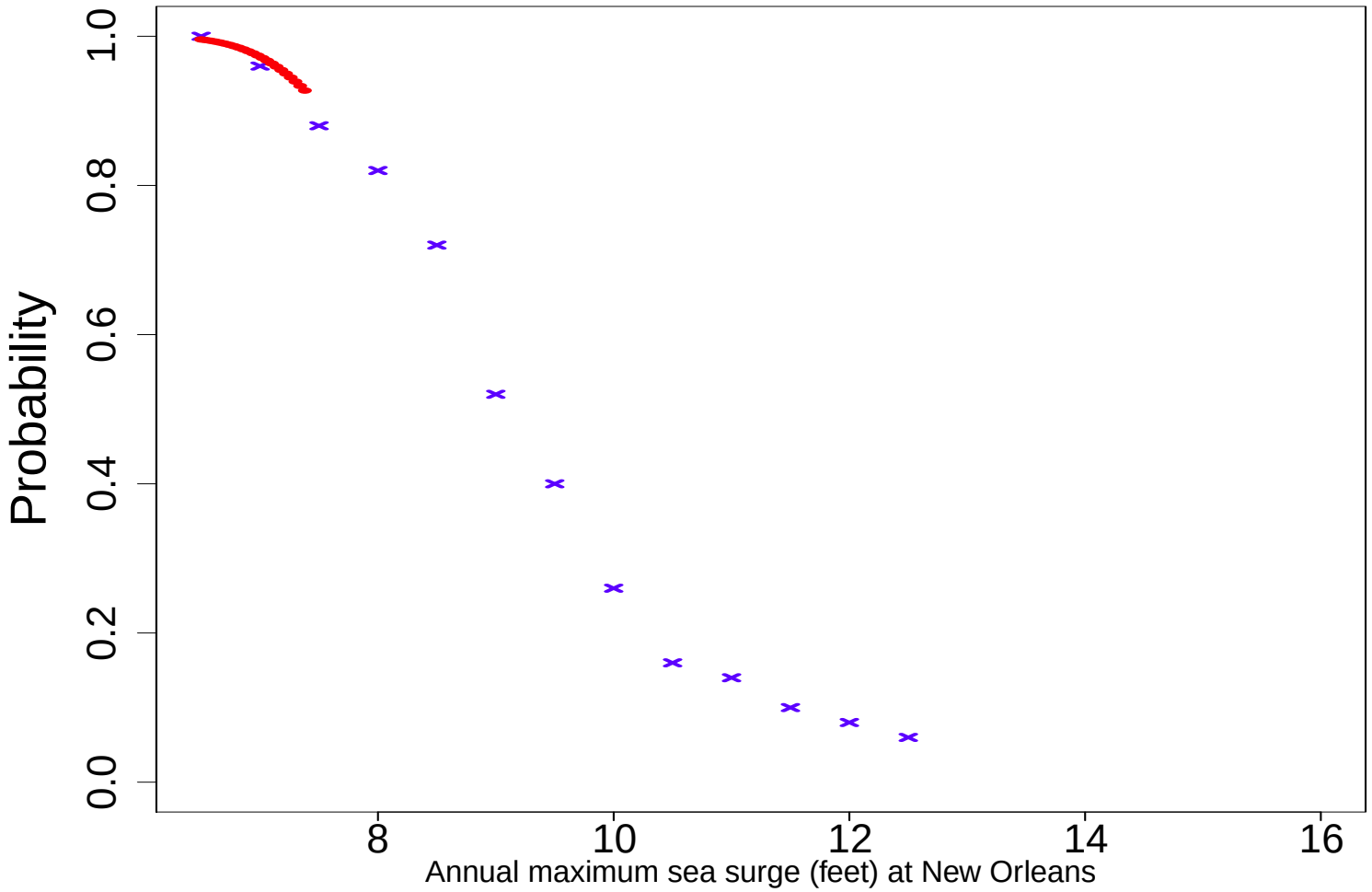
Part B: A probability model for extremes



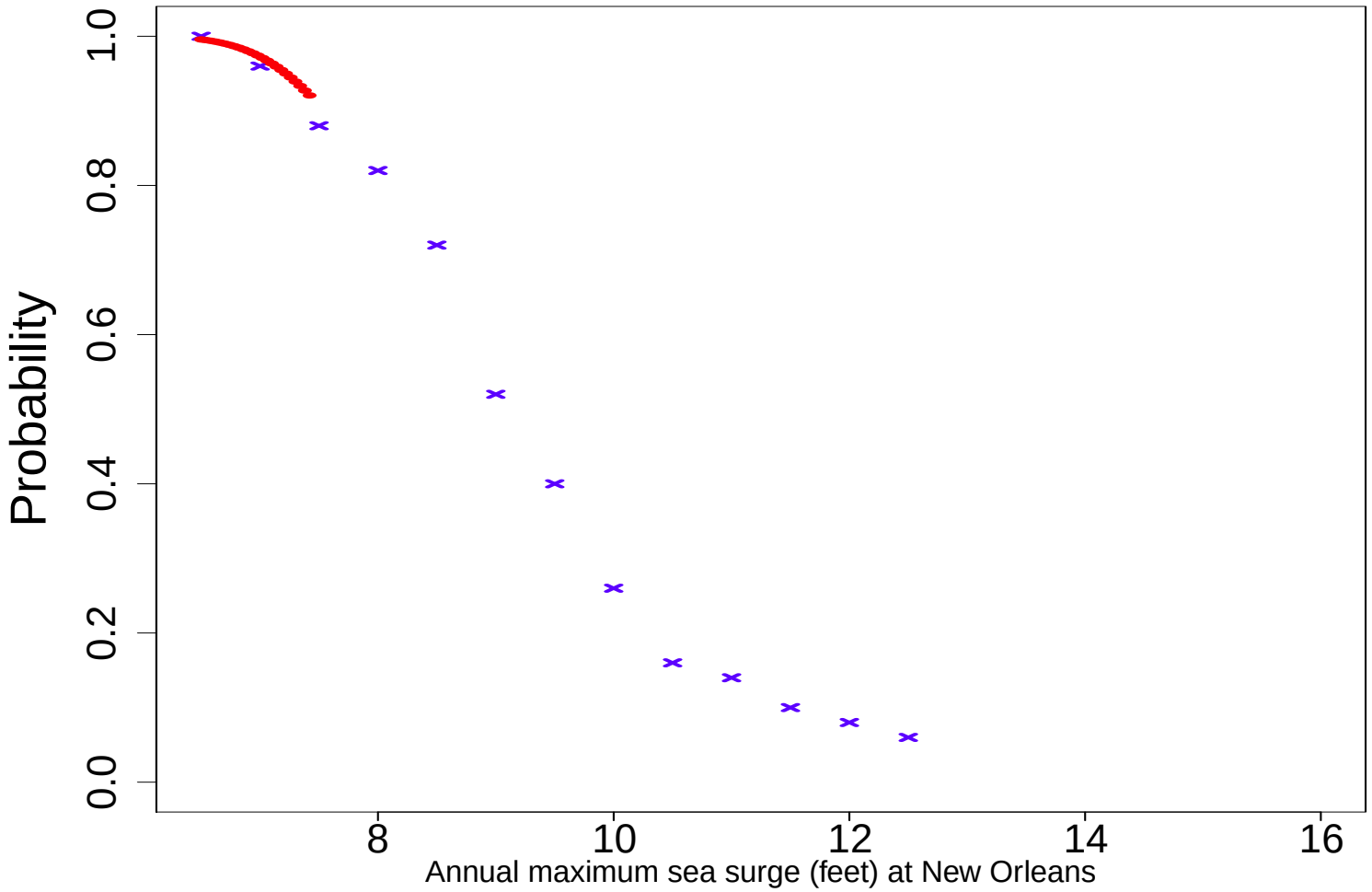
Part B: A probability model for extremes



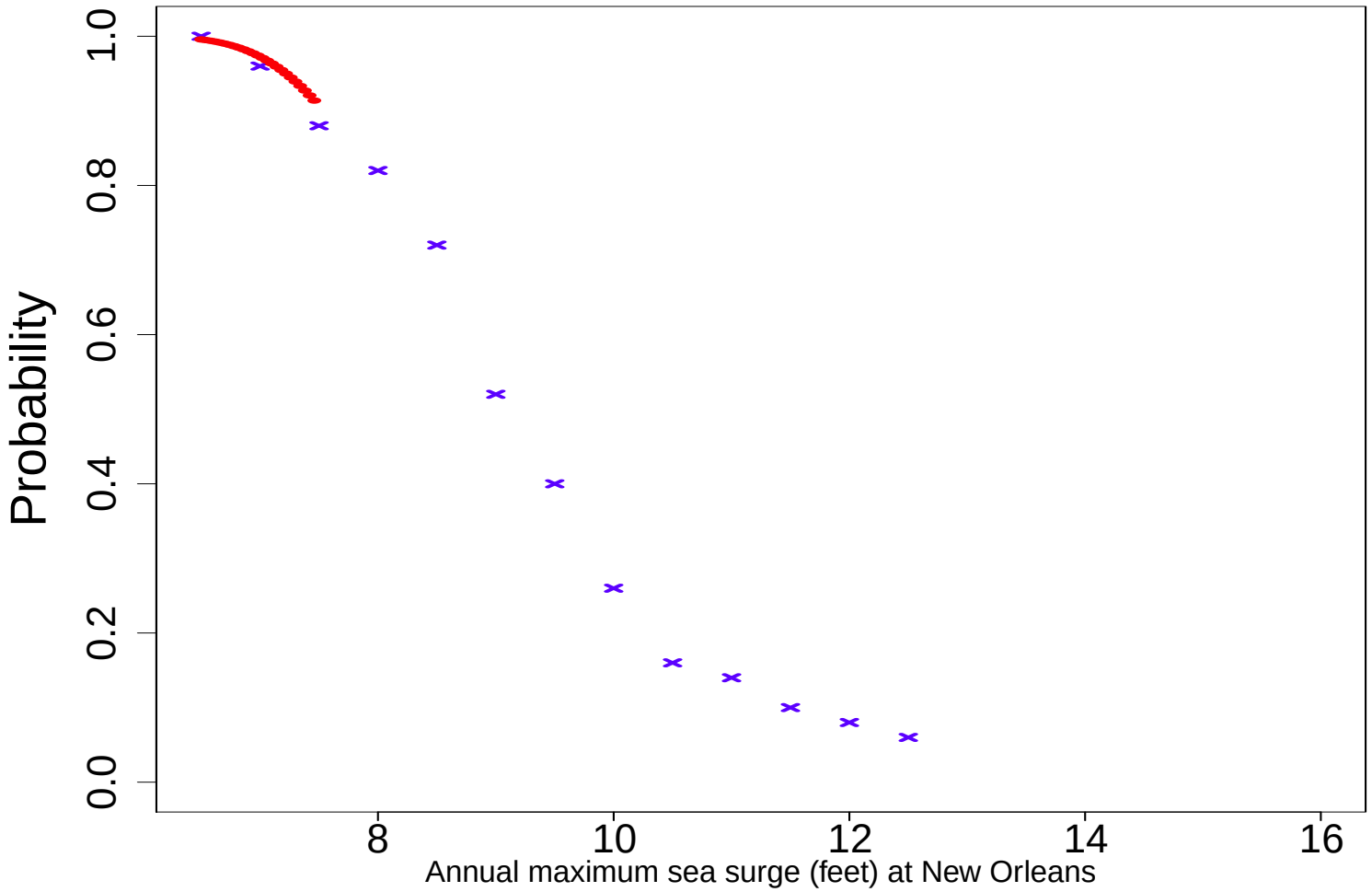
Part B: A probability model for extremes



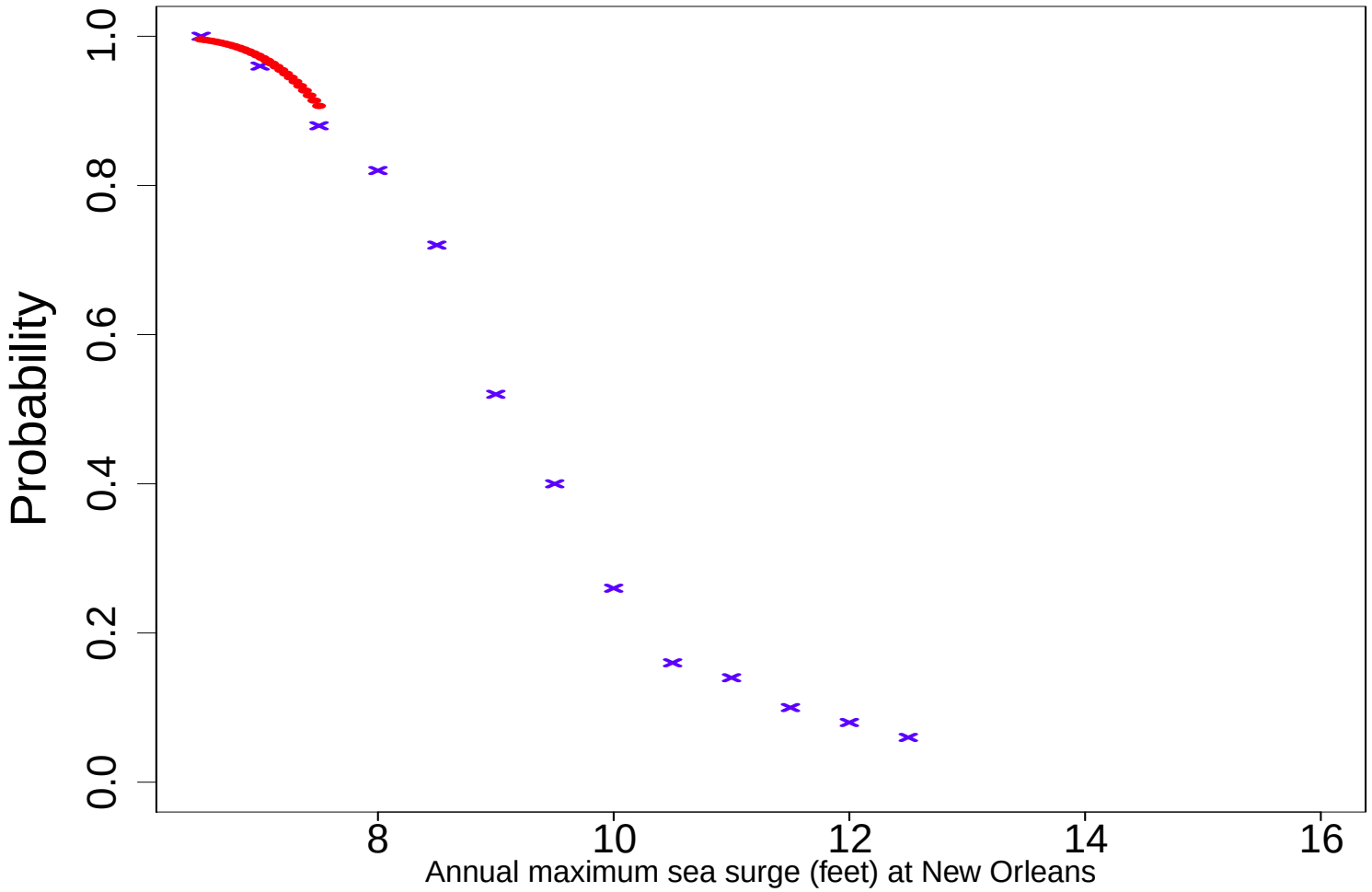
Part B: A probability model for extremes



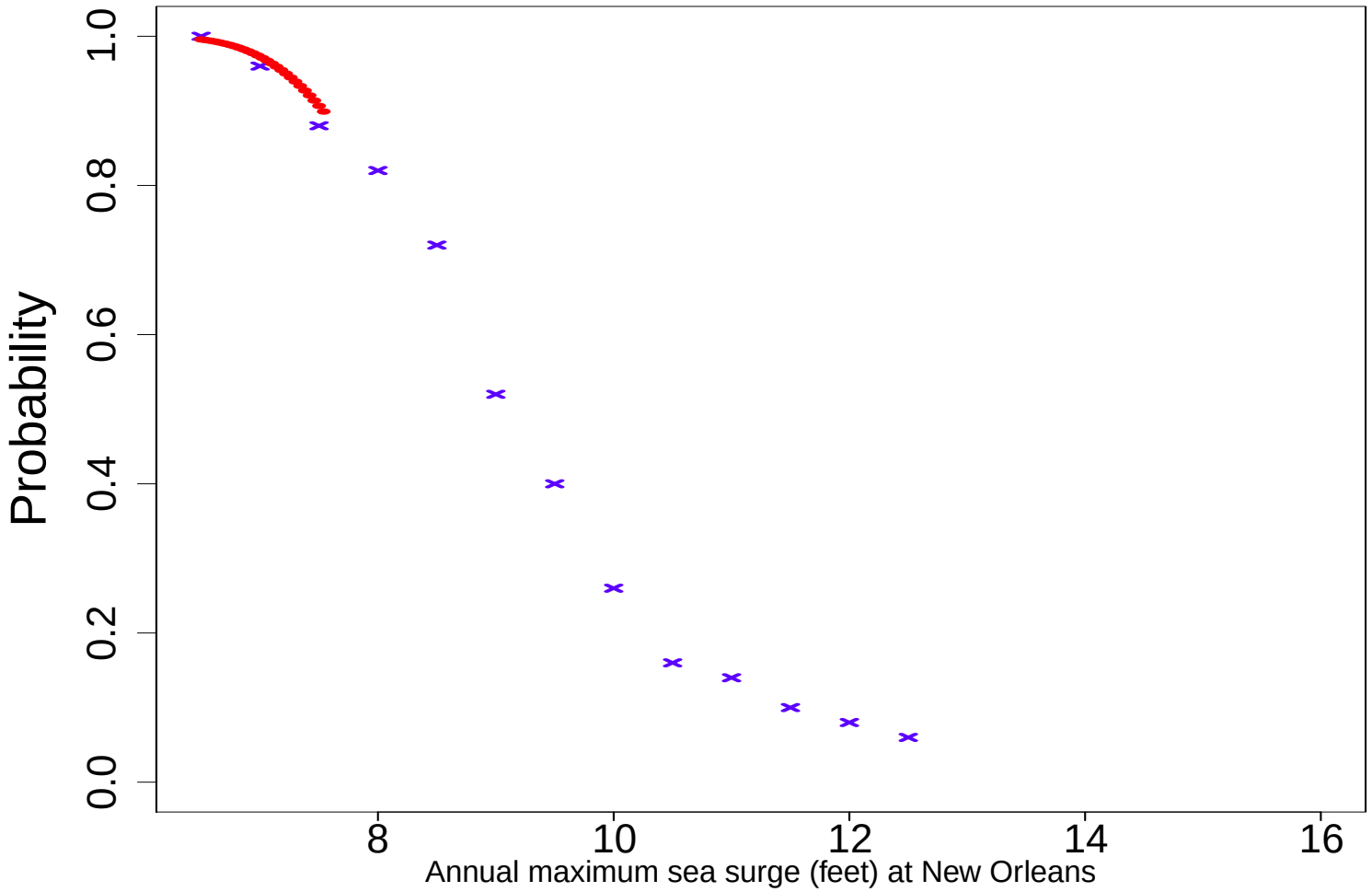
Part B: A probability model for extremes



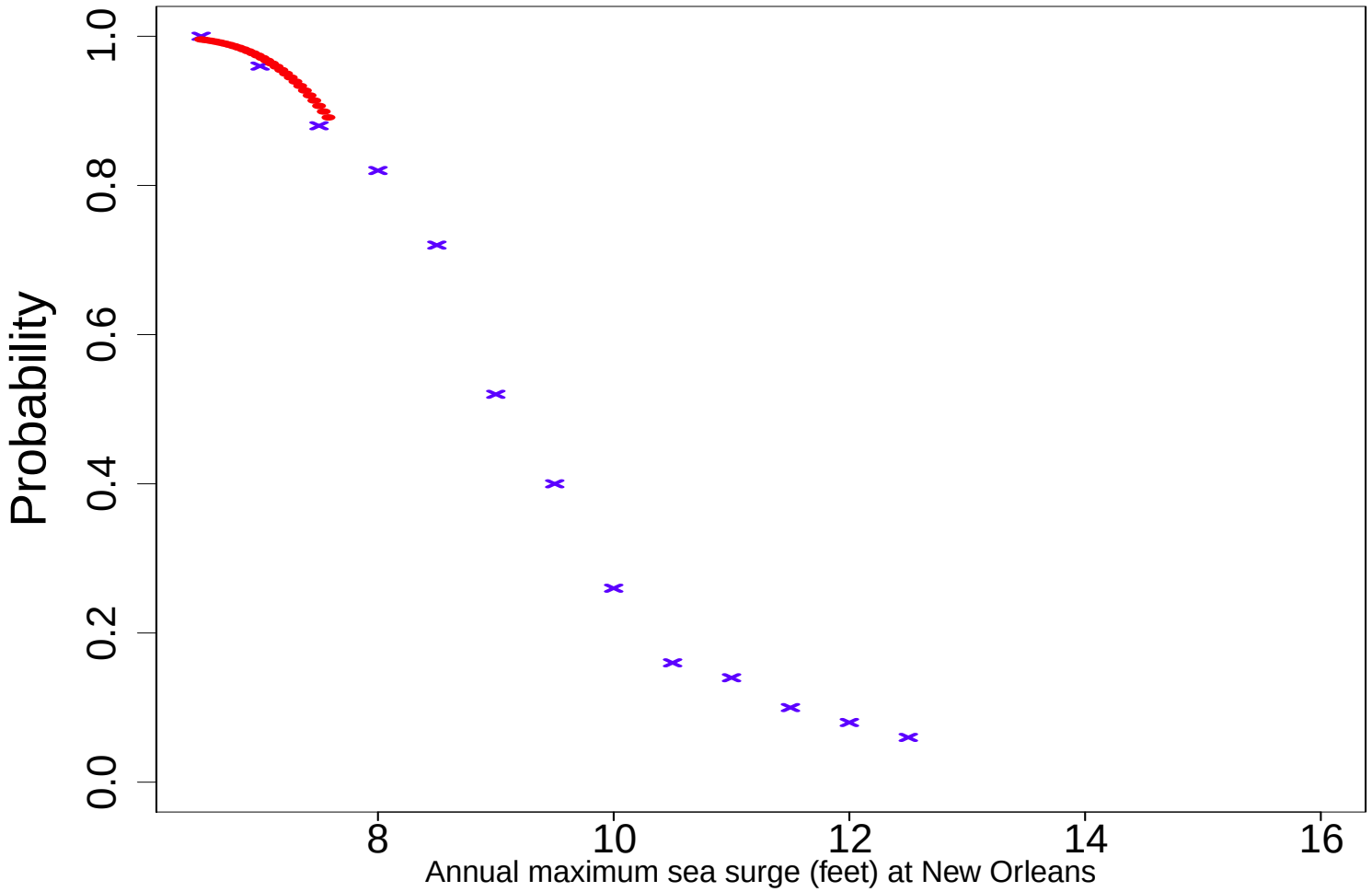
Part B: A probability model for extremes



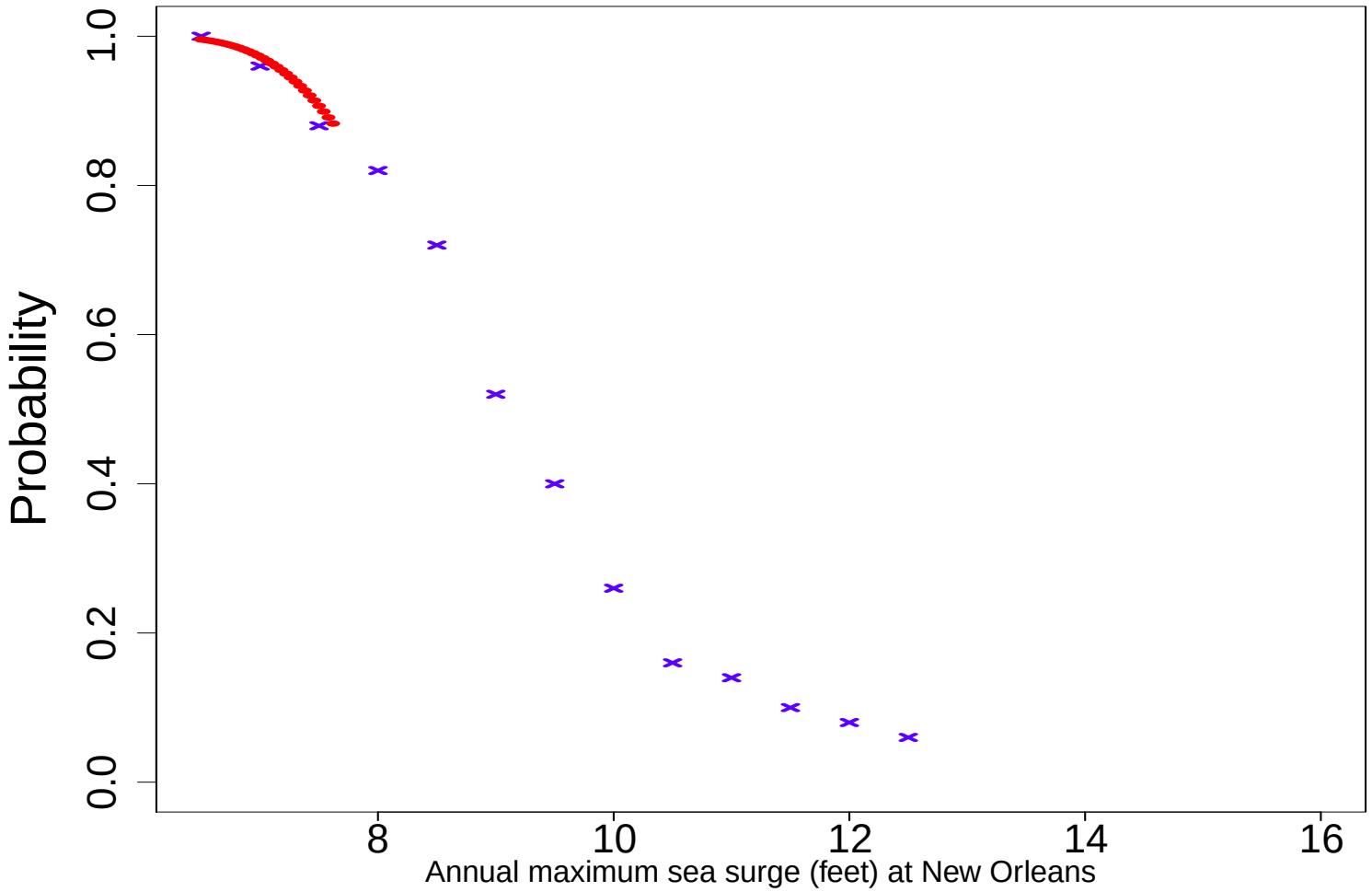
Part B: A probability model for extremes



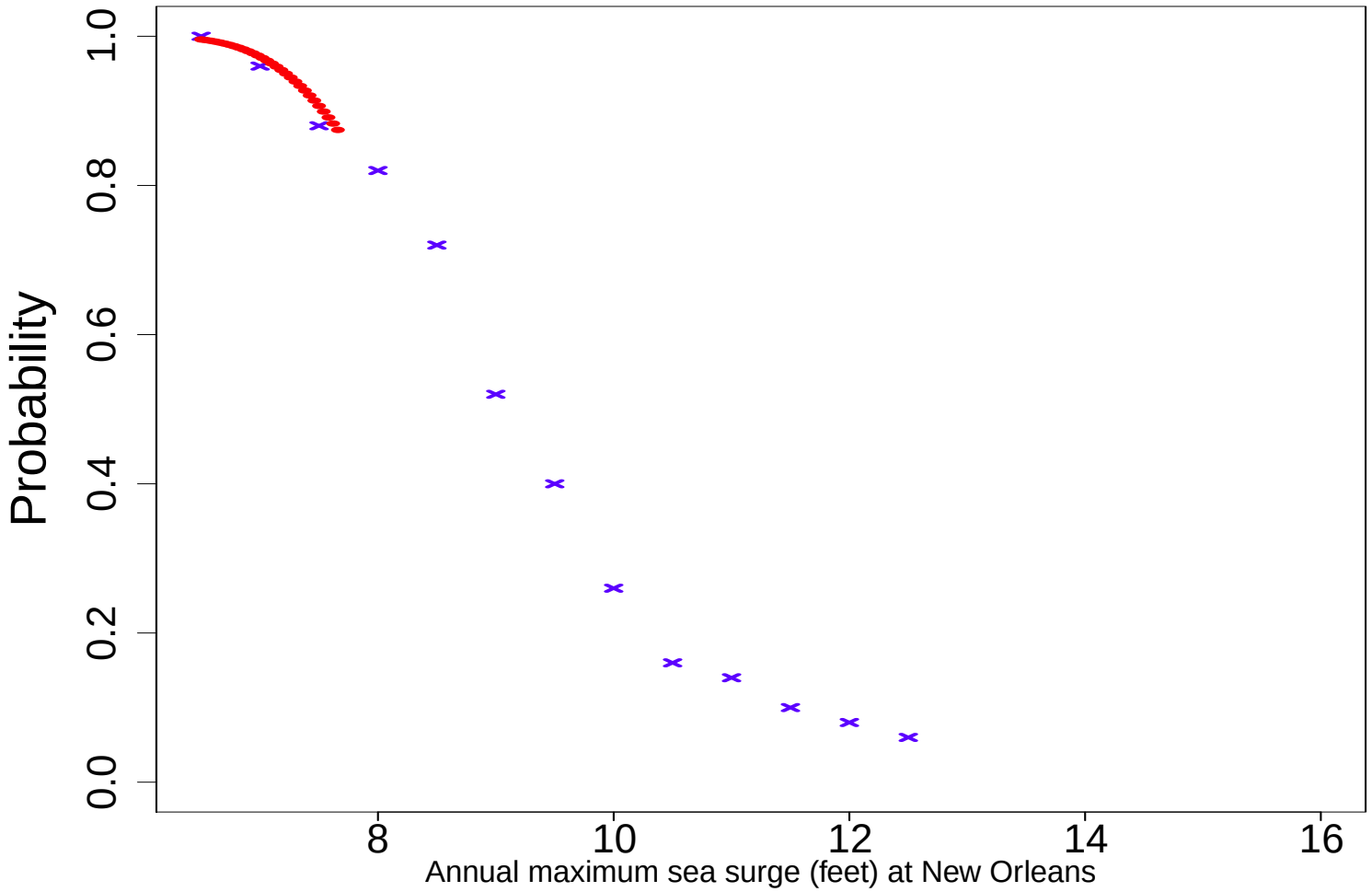
Part B: A probability model for extremes



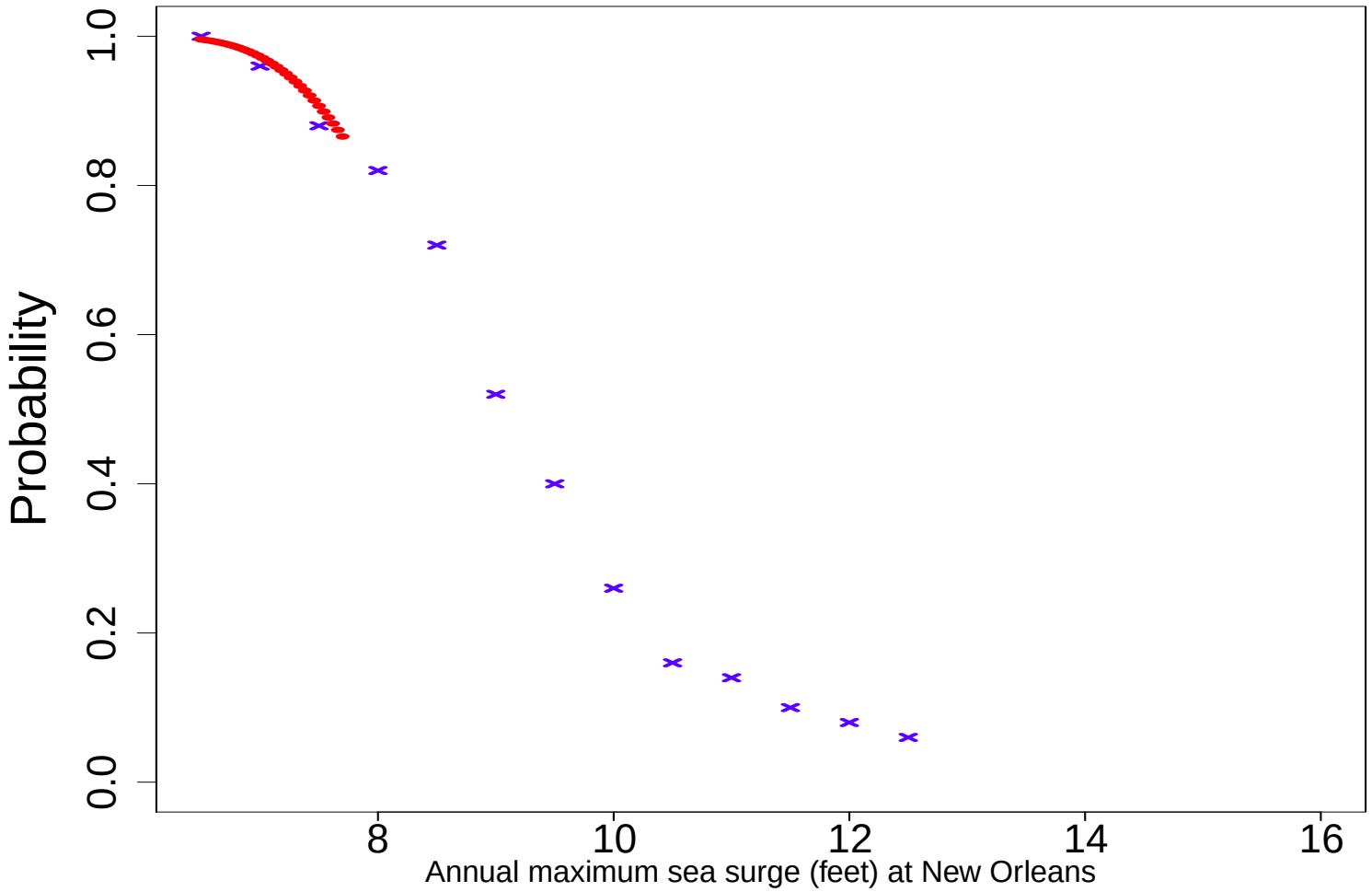
Part B: A probability model for extremes



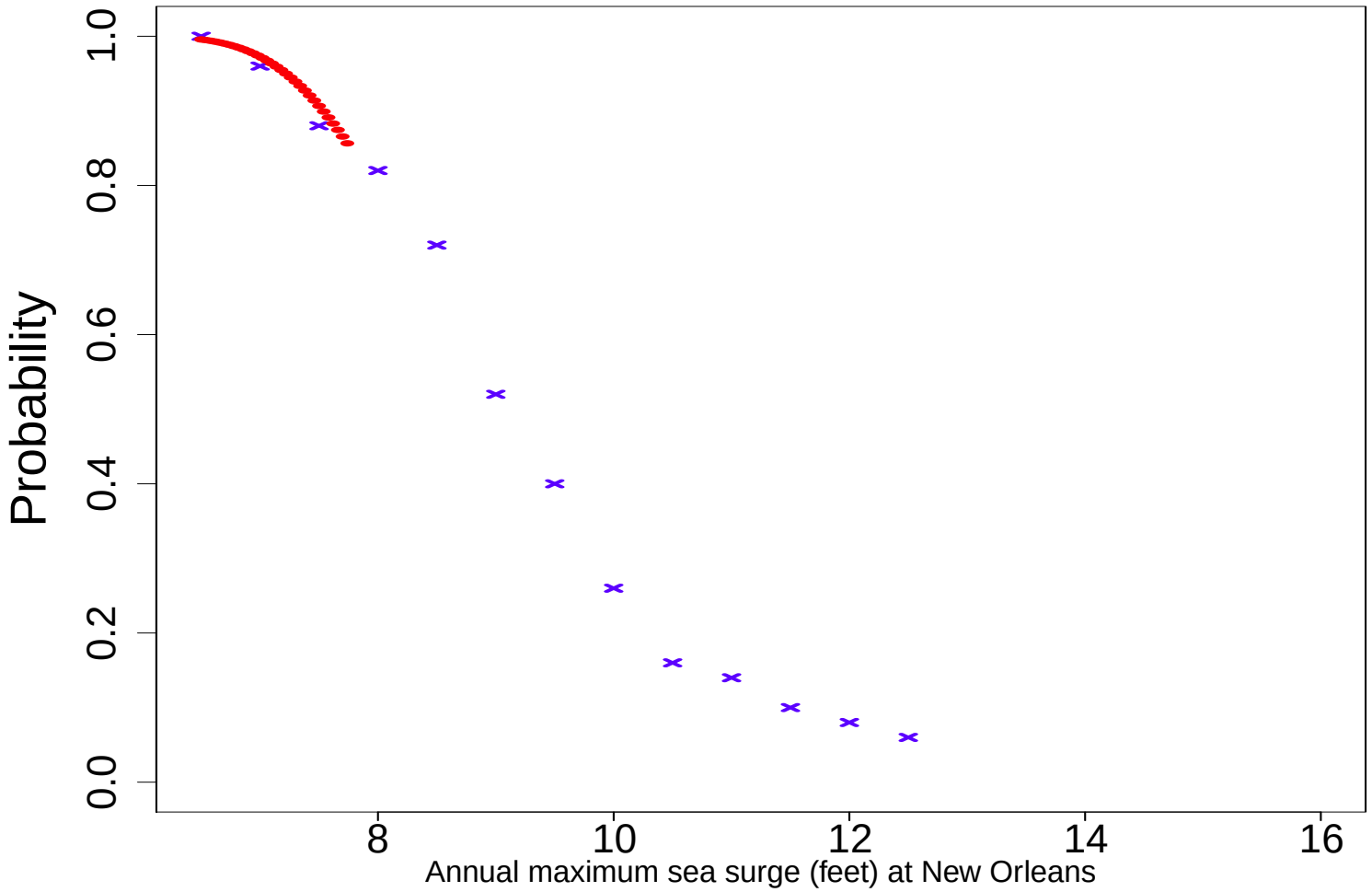
Part B: A probability model for extremes



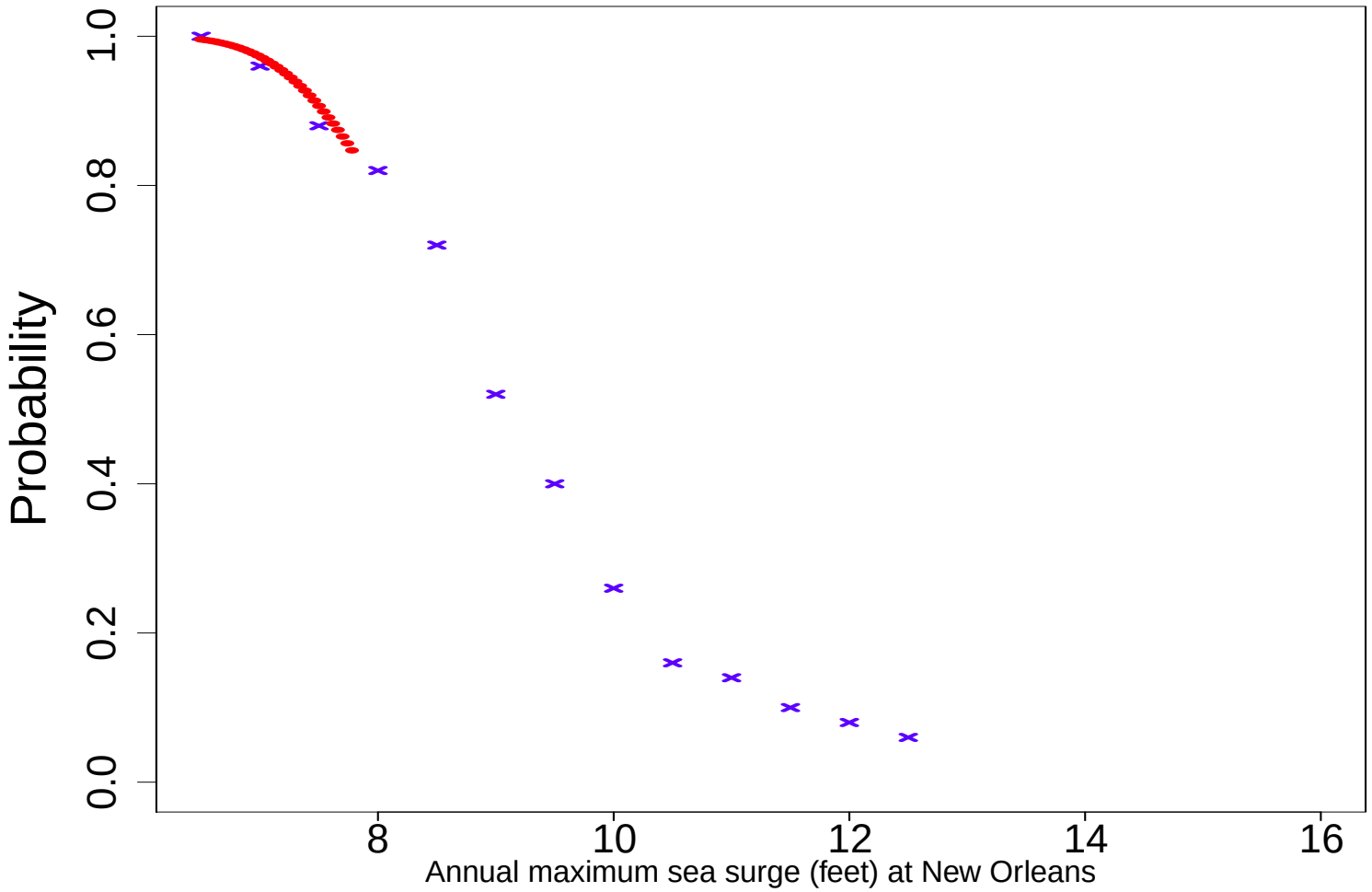
Part B: A probability model for extremes



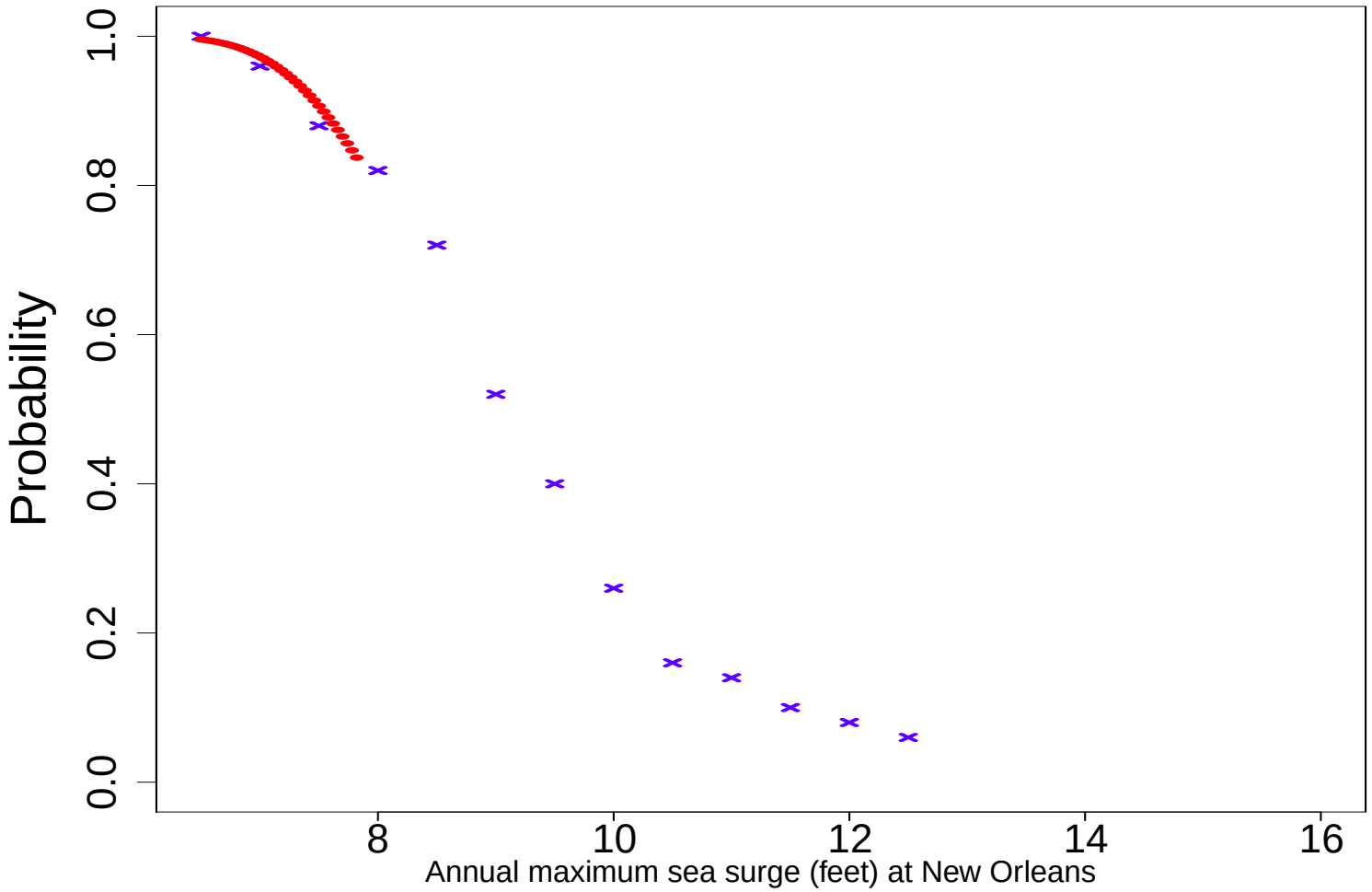
Part B: A probability model for extremes



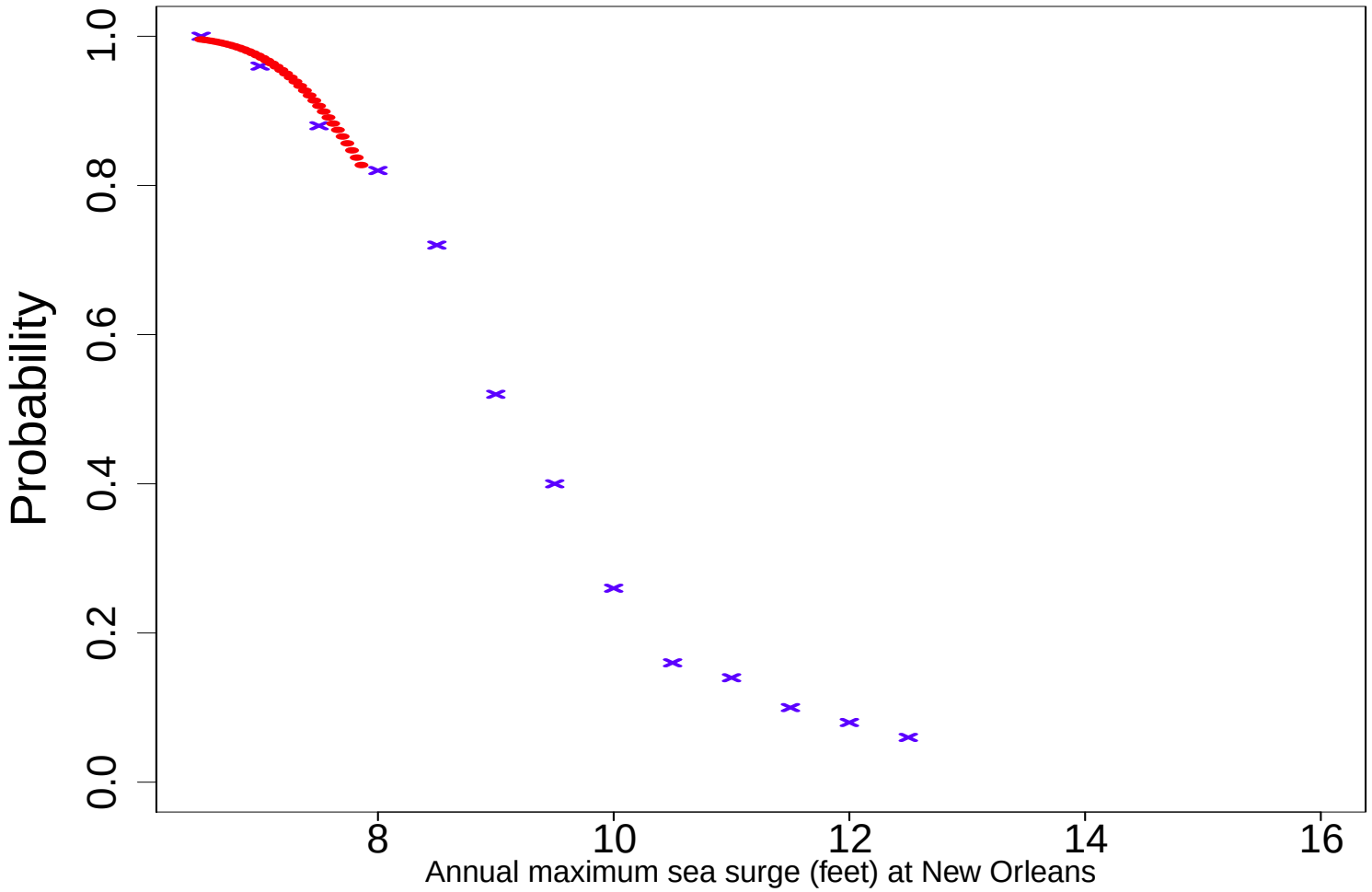
Part B: A probability model for extremes



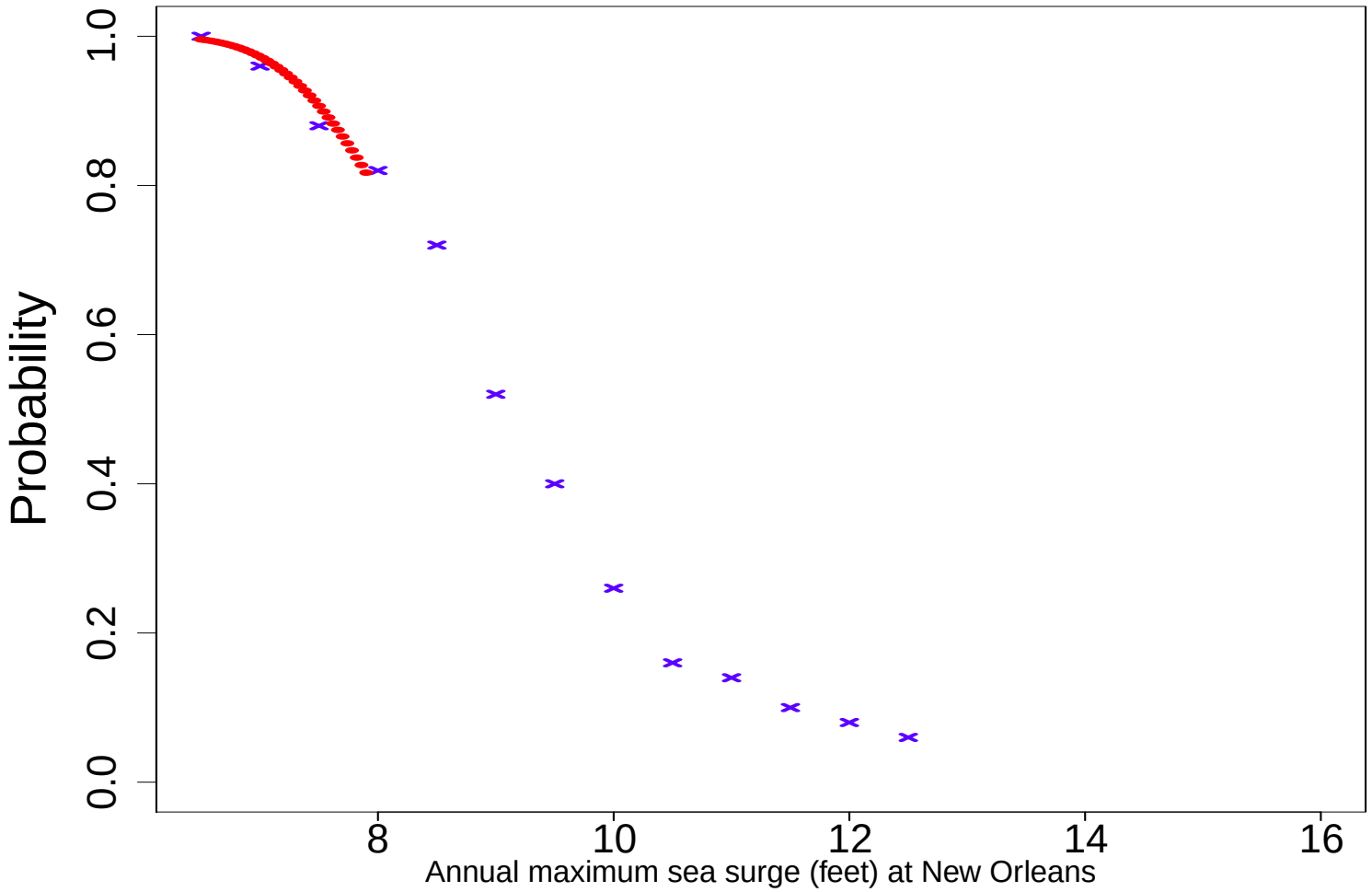
Part B: A probability model for extremes



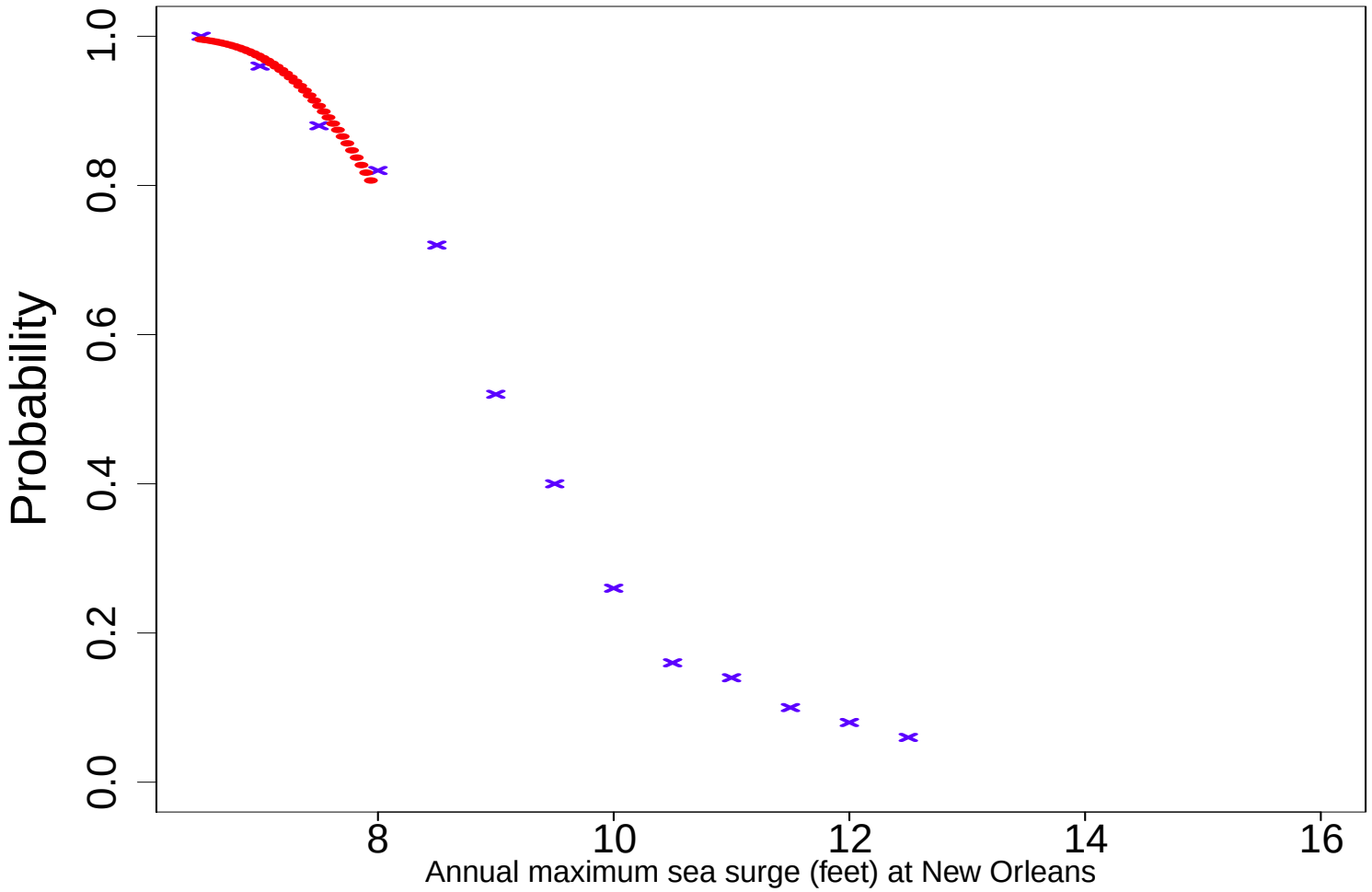
Part B: A probability model for extremes



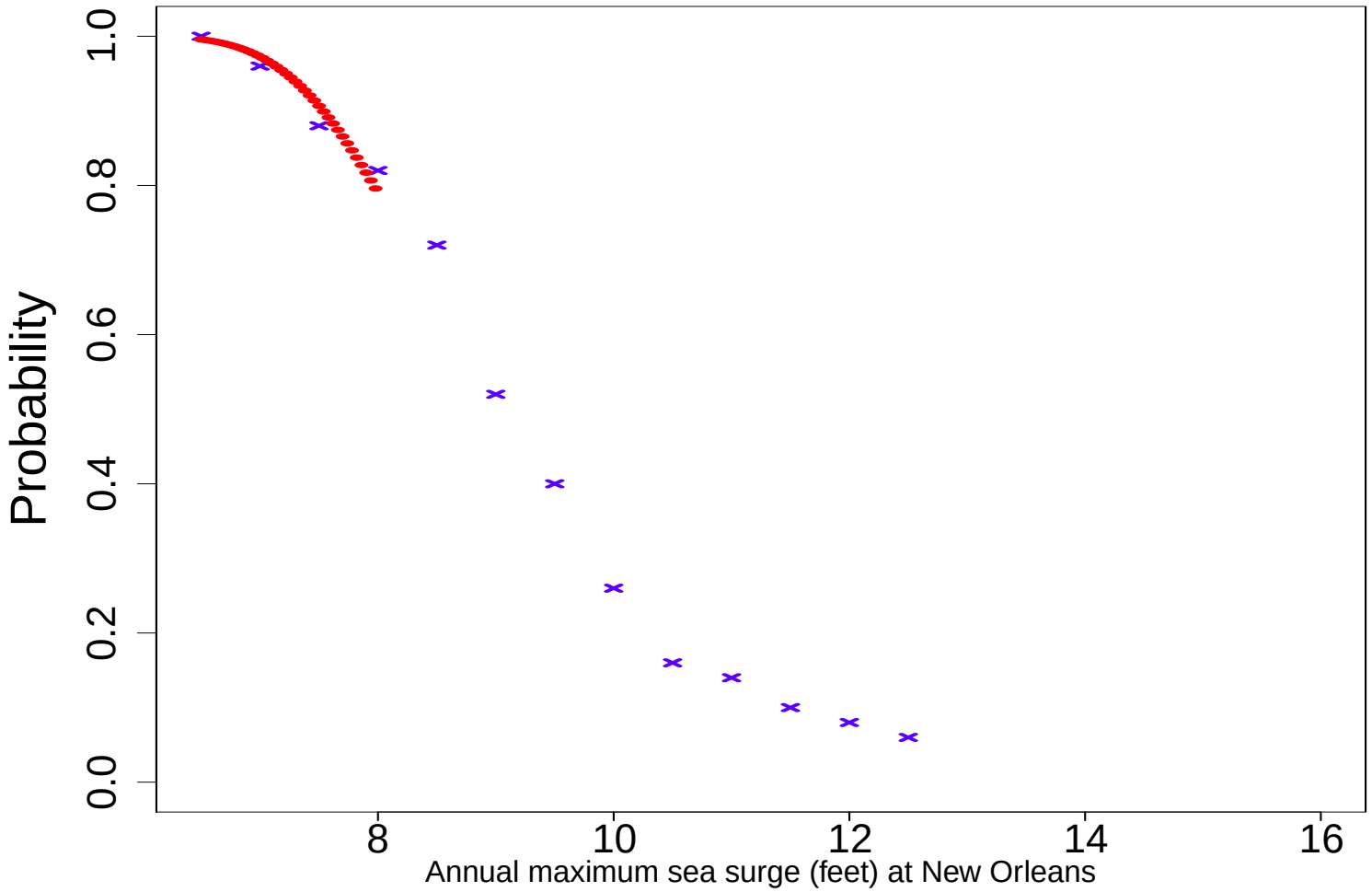
Part B: A probability model for extremes



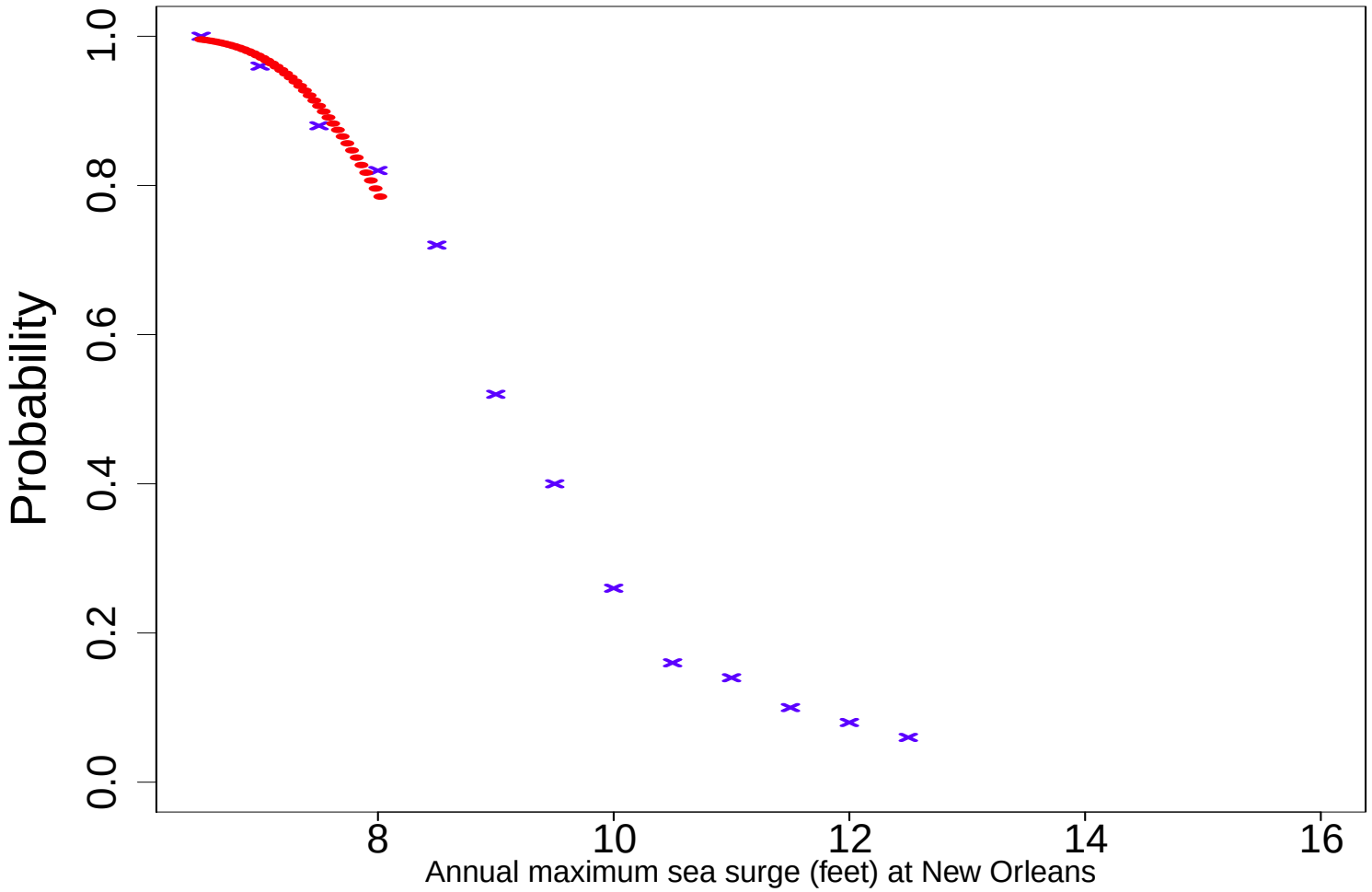
Part B: A probability model for extremes



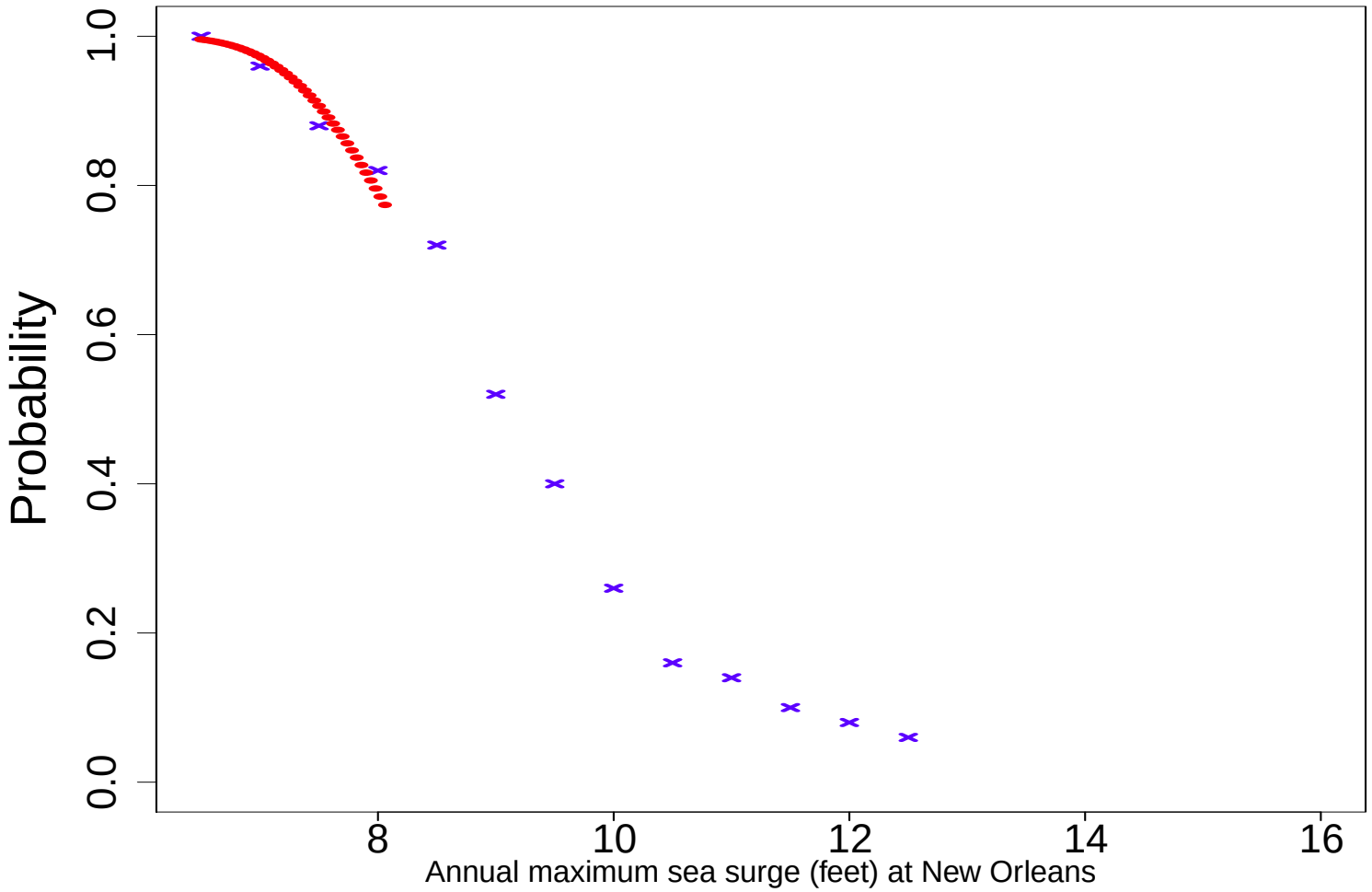
Part B: A probability model for extremes



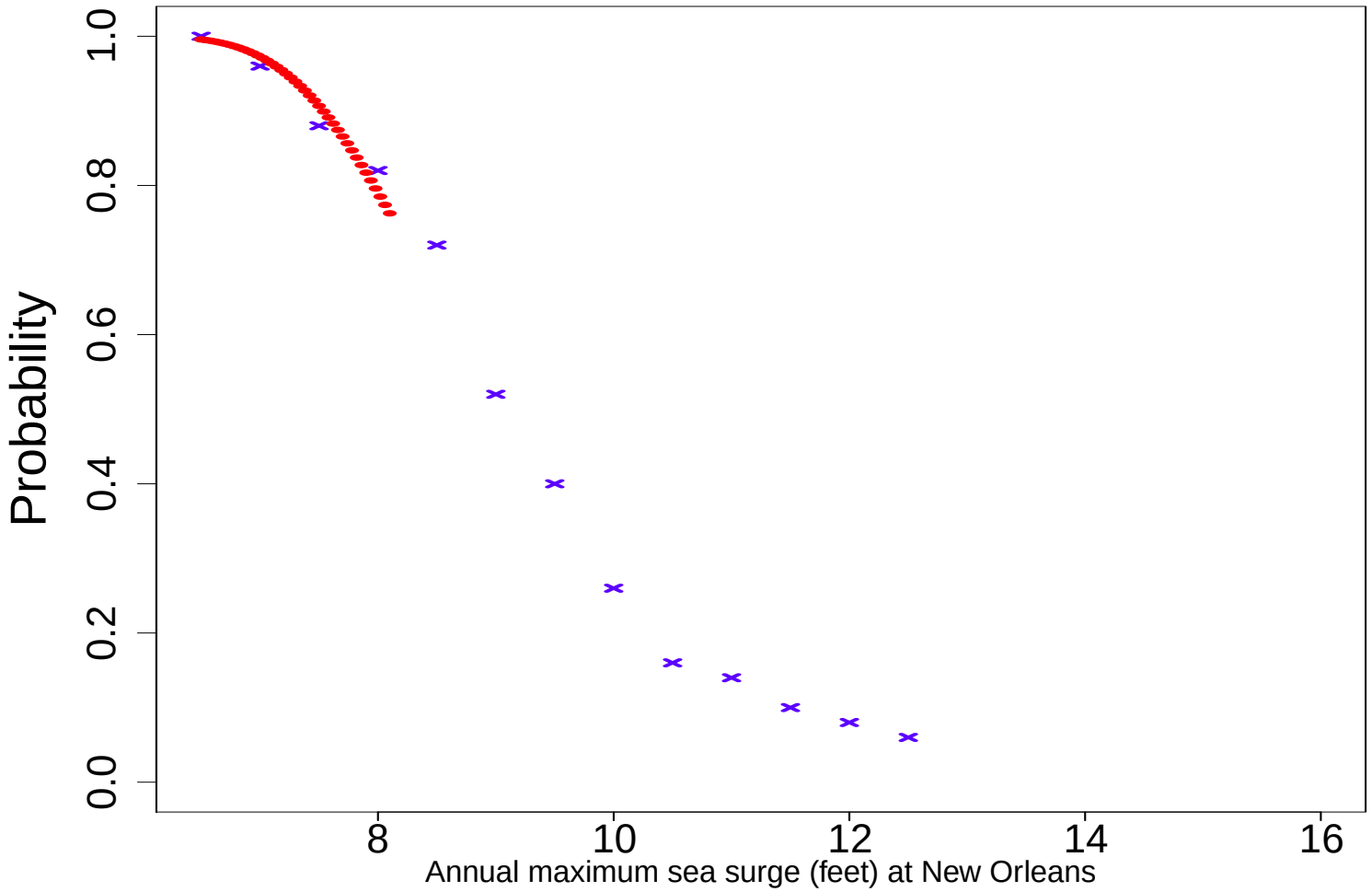
Part B: A probability model for extremes



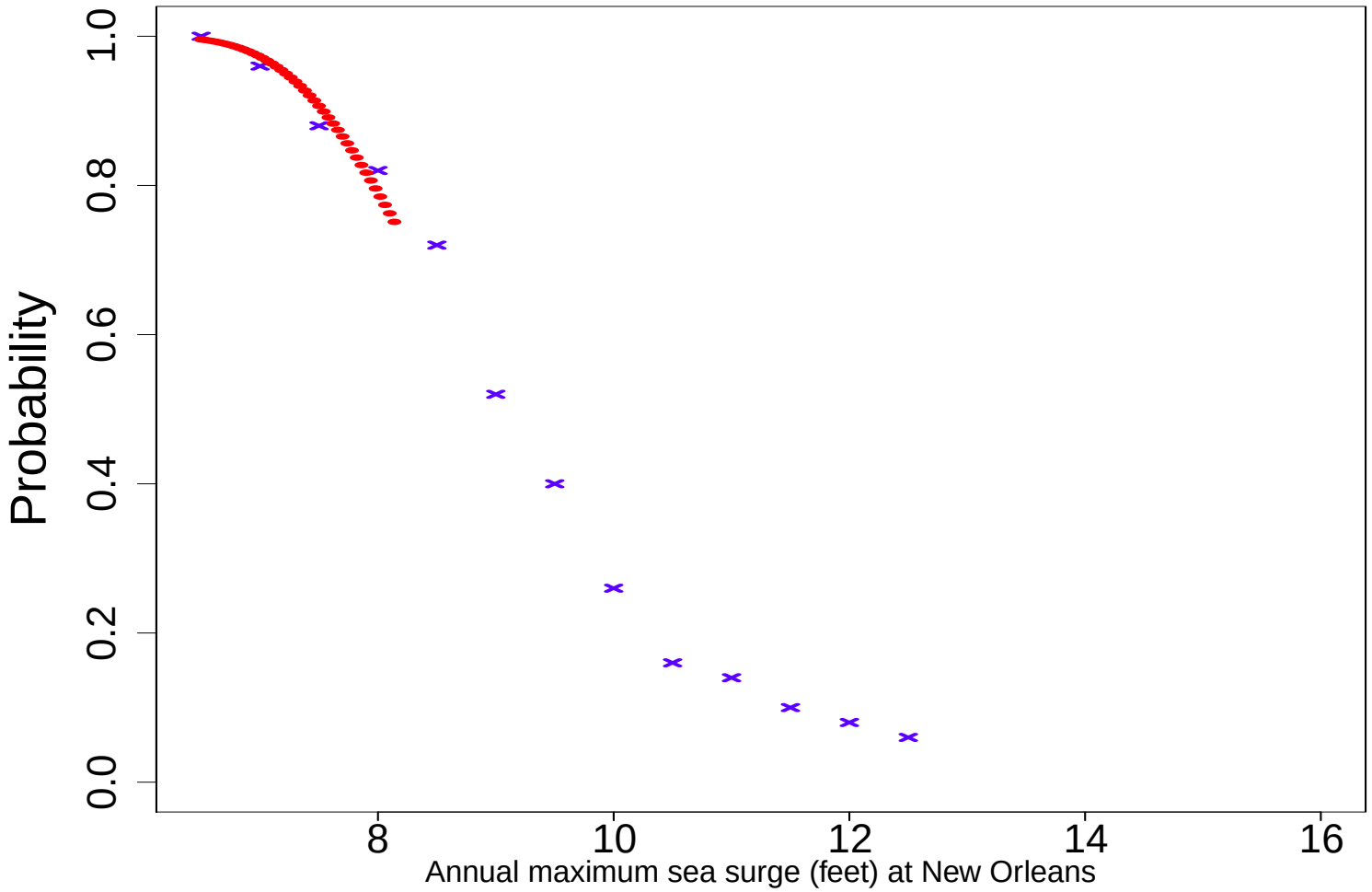
Part B: A probability model for extremes



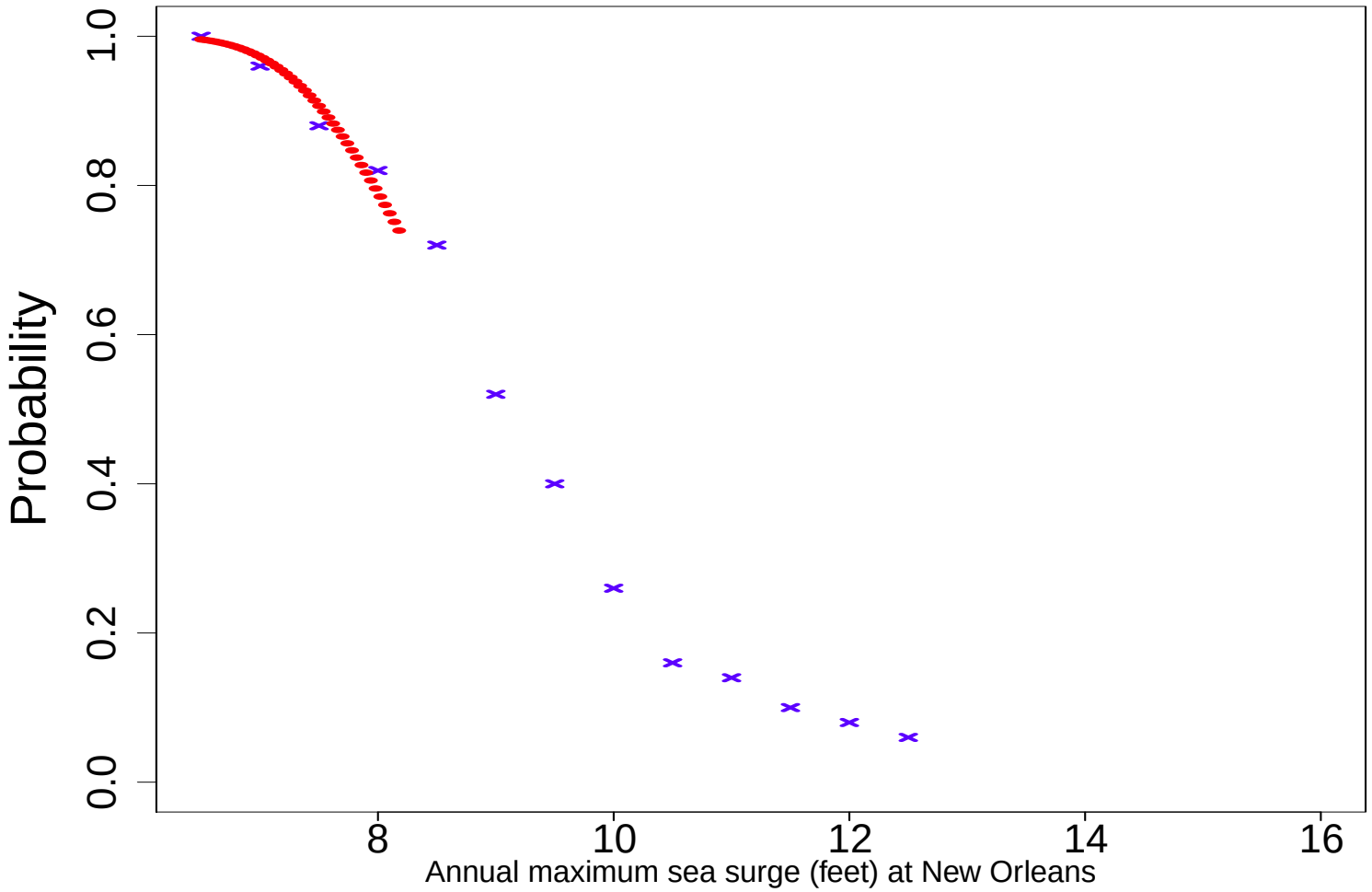
Part B: A probability model for extremes



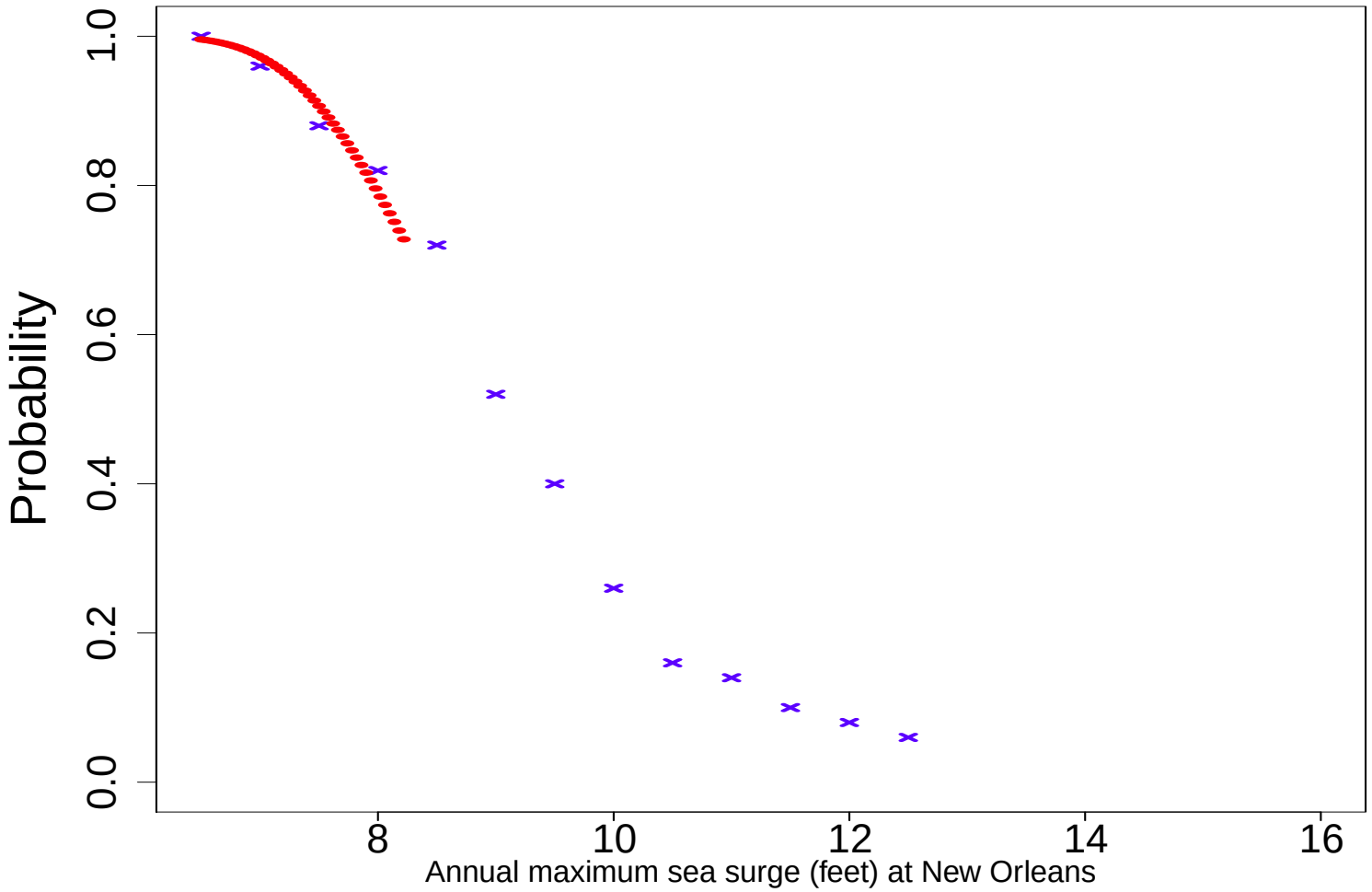
Part B: A probability model for extremes



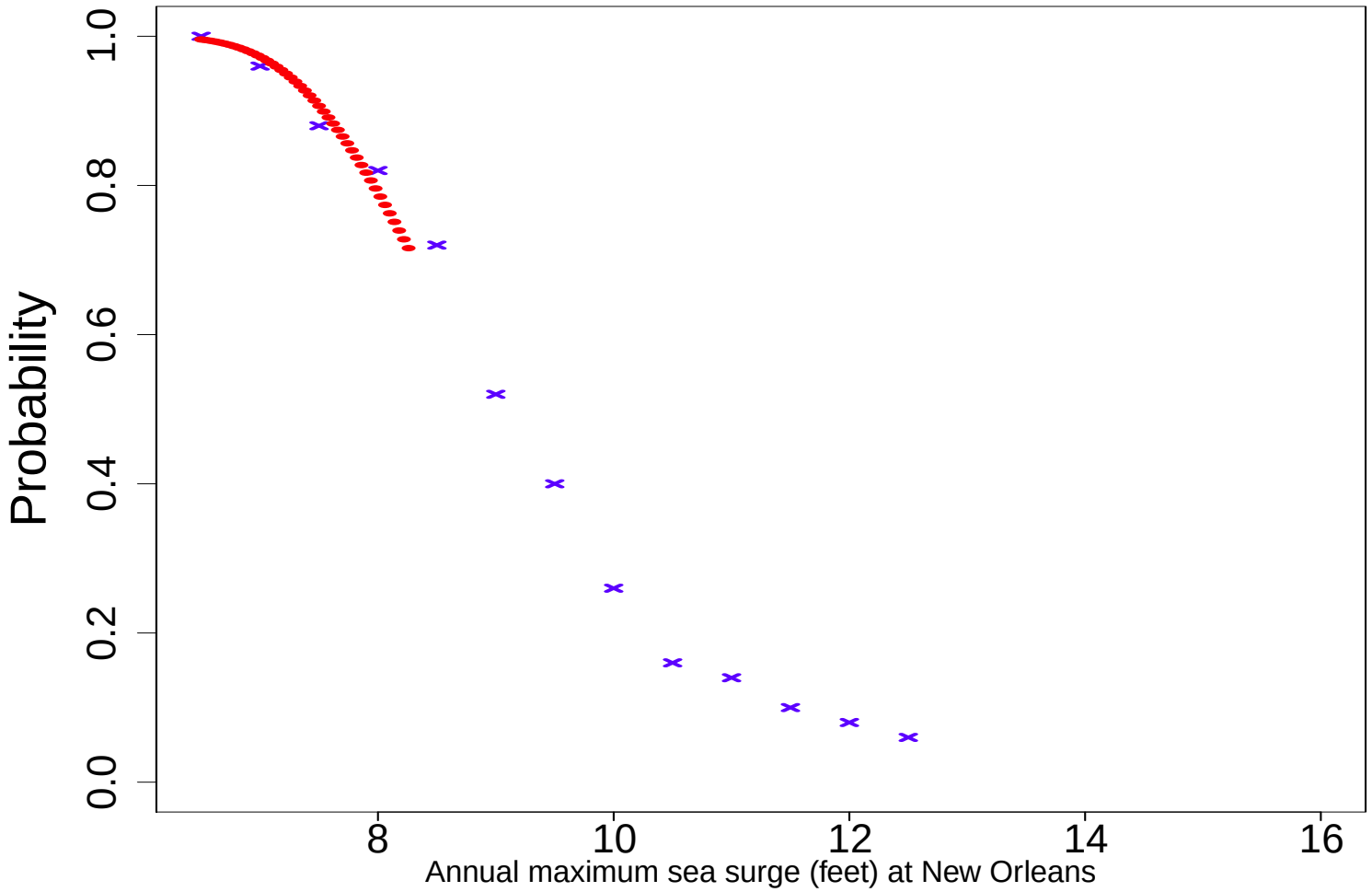
Part B: A probability model for extremes



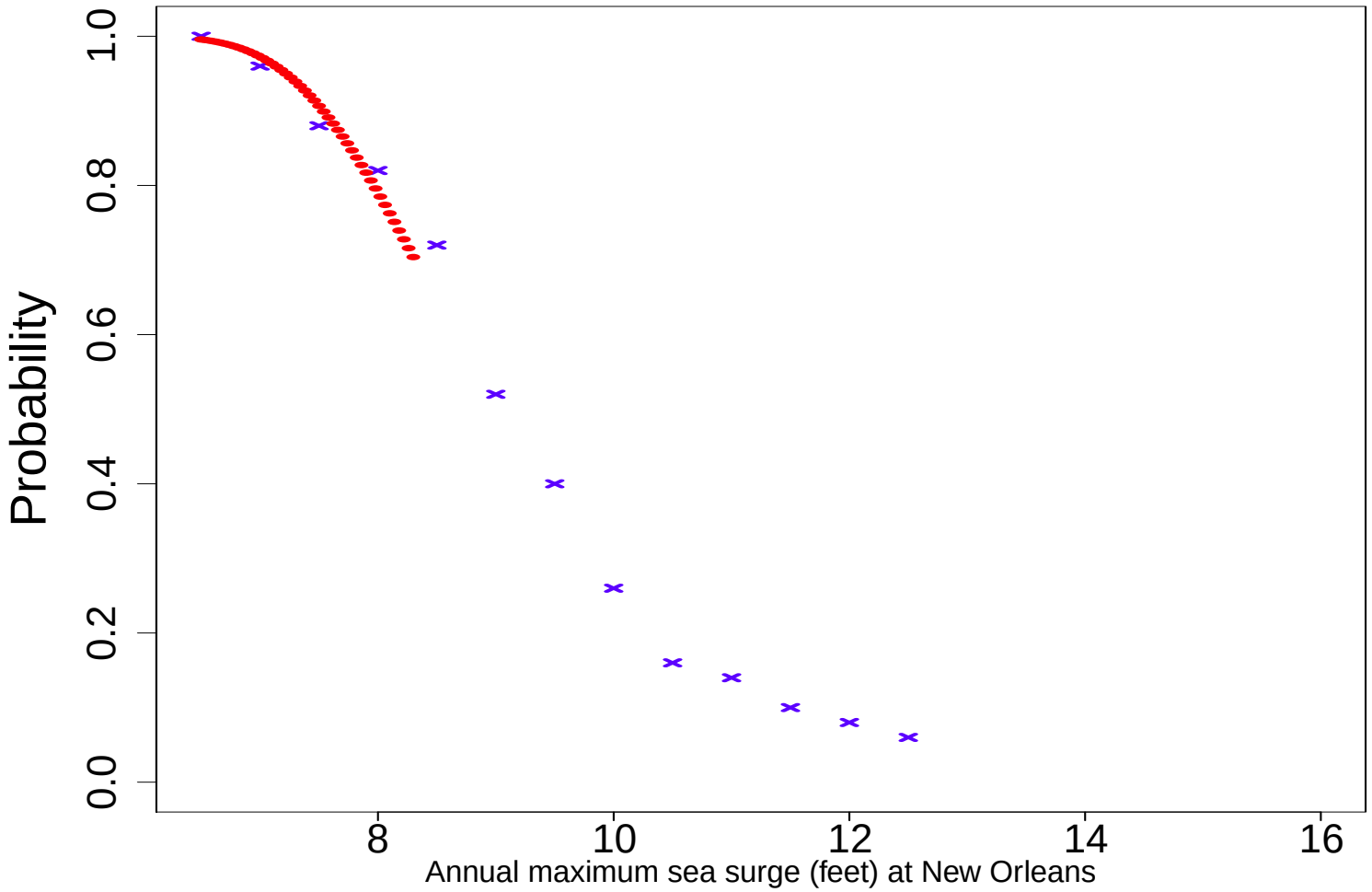
Part B: A probability model for extremes



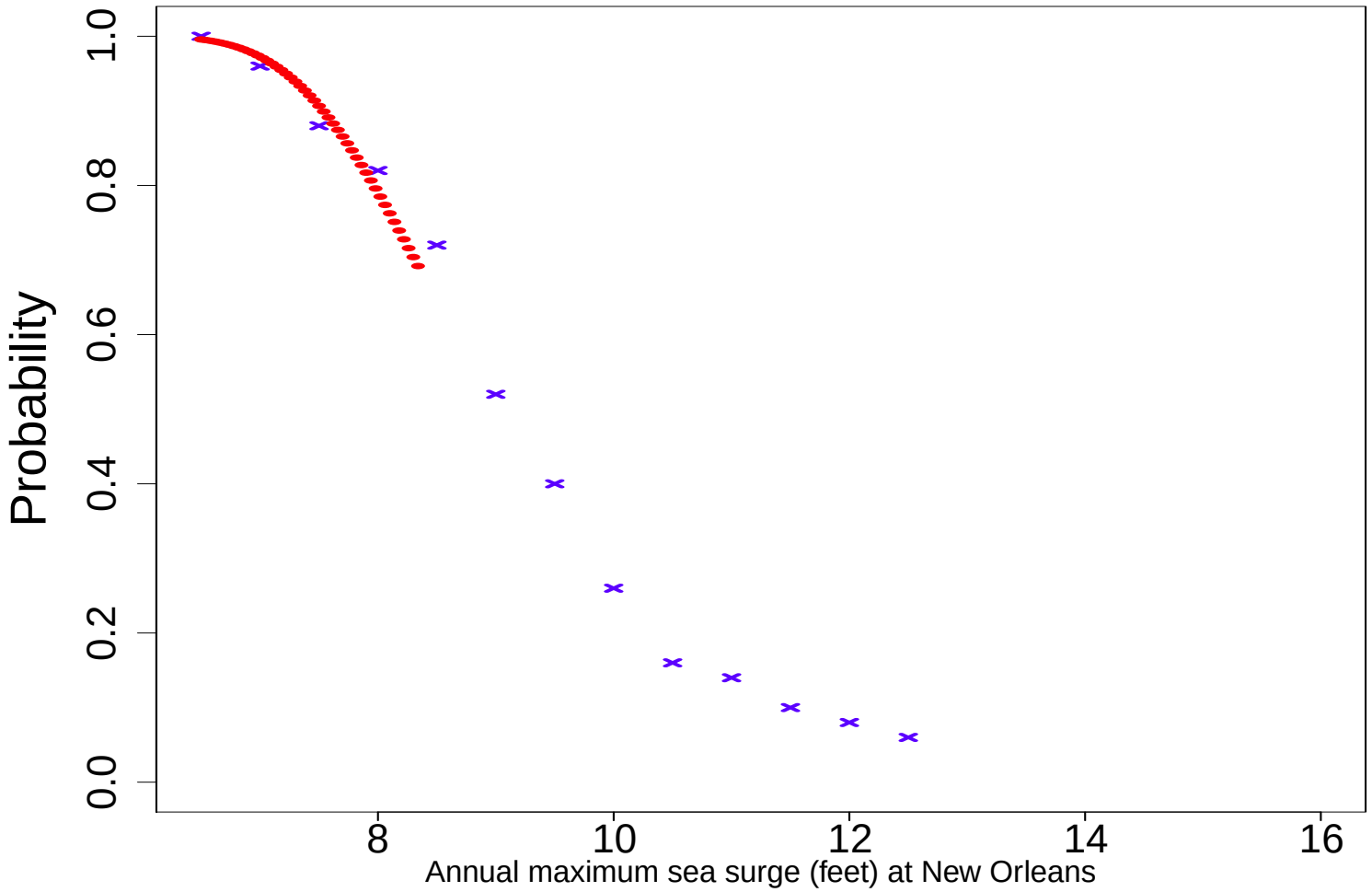
Part B: A probability model for extremes



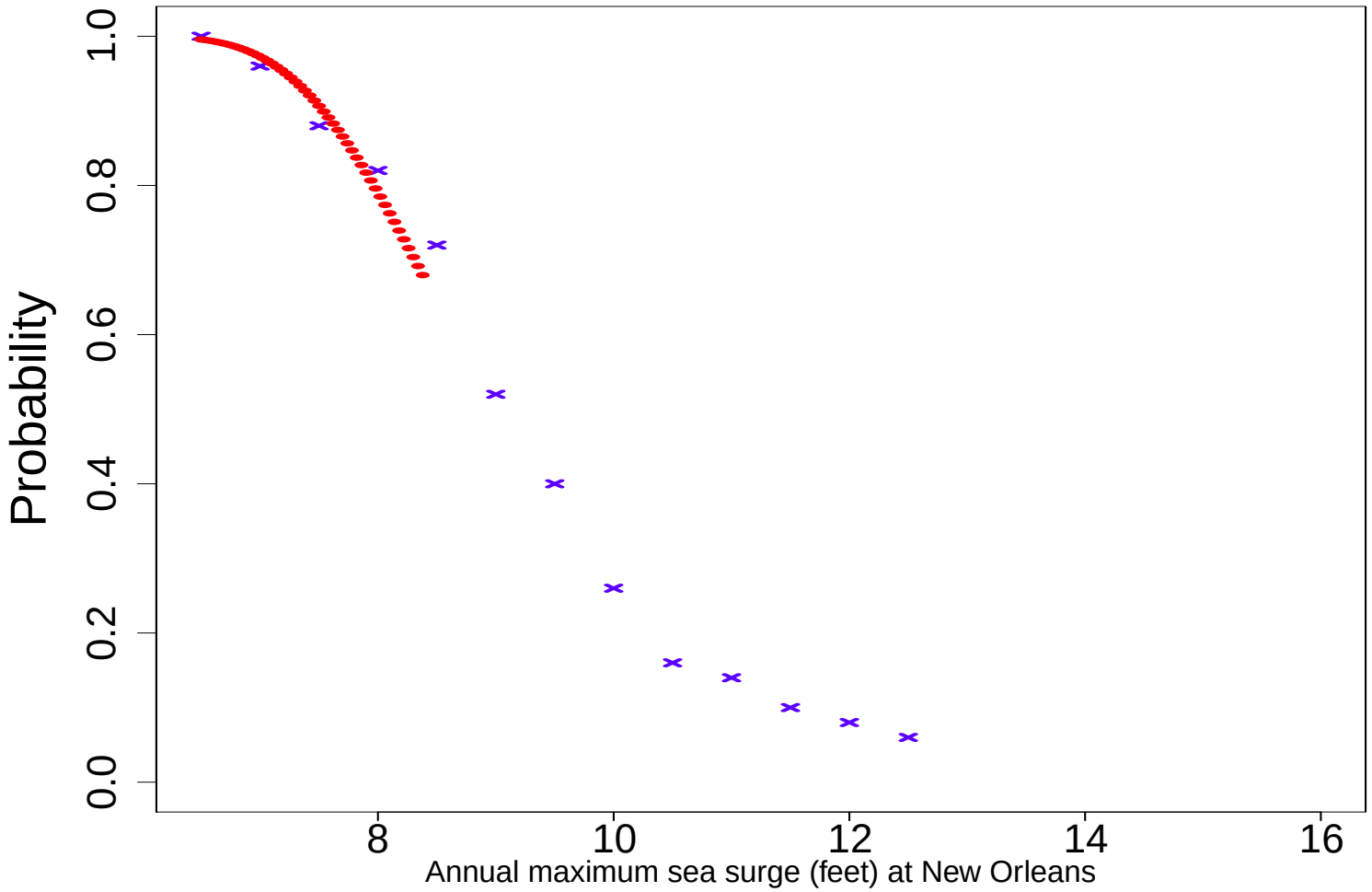
Part B: A probability model for extremes



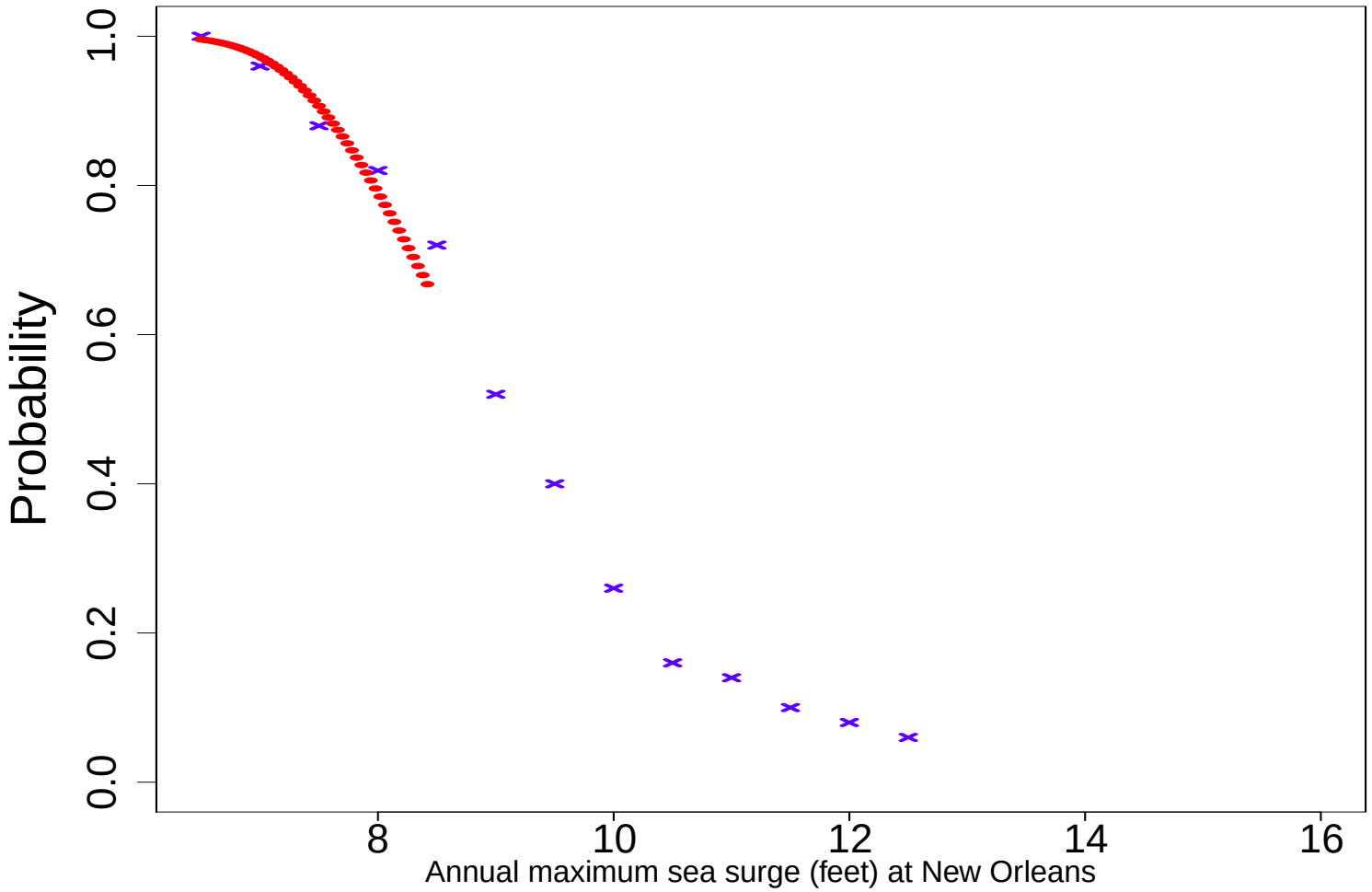
Part B: A probability model for extremes



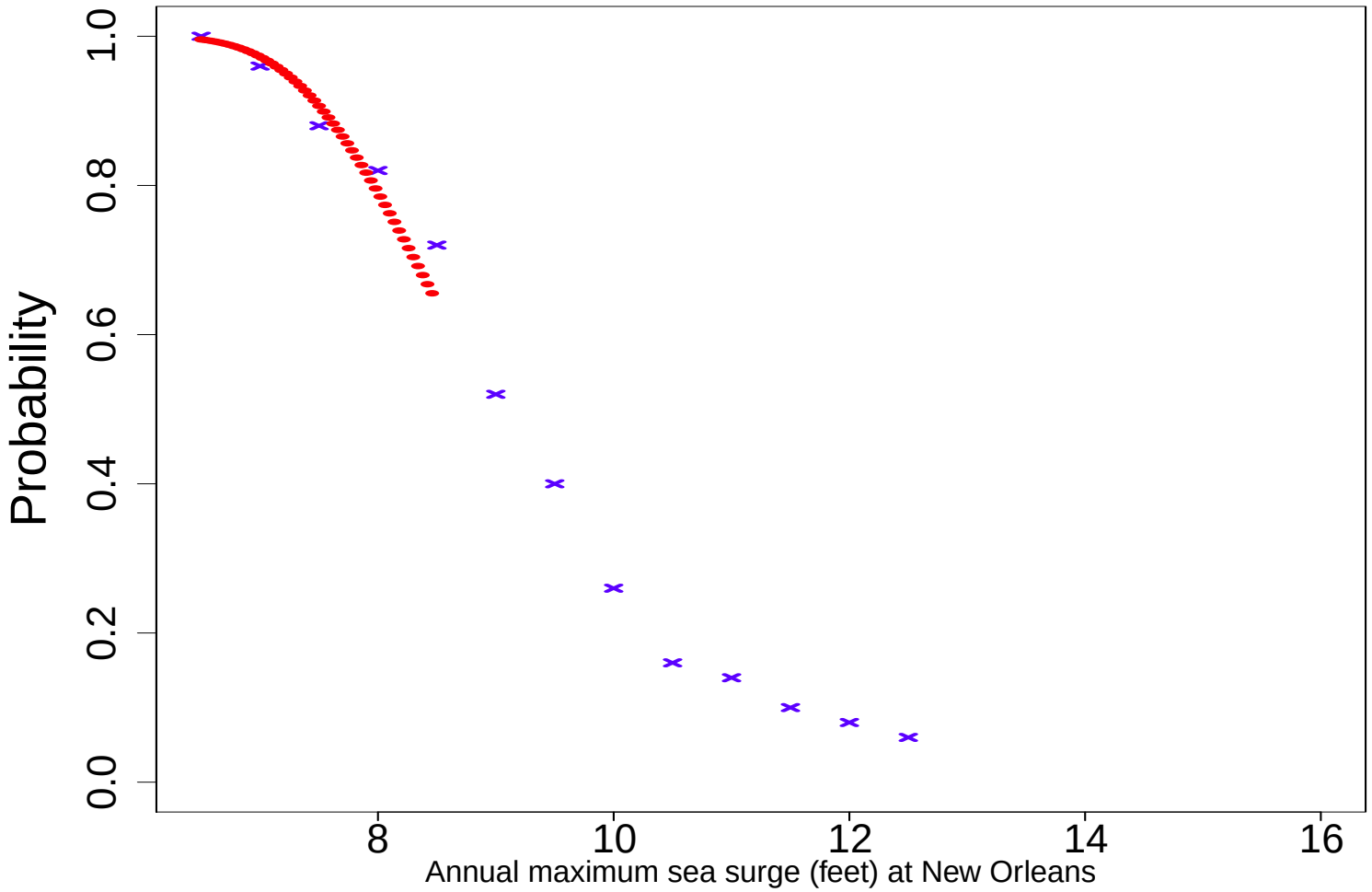
Part B: A probability model for extremes



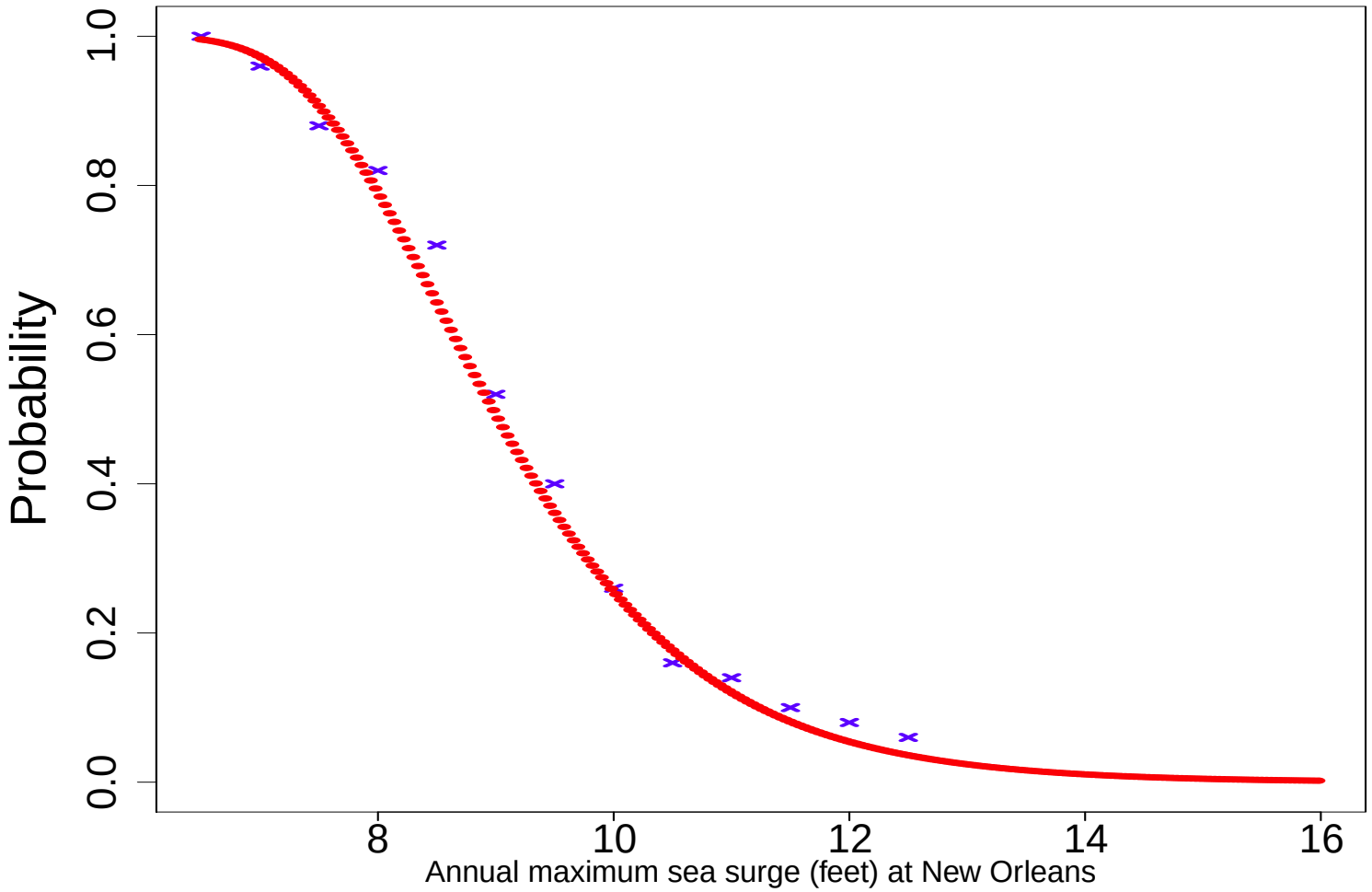
Part B: A probability model for extremes



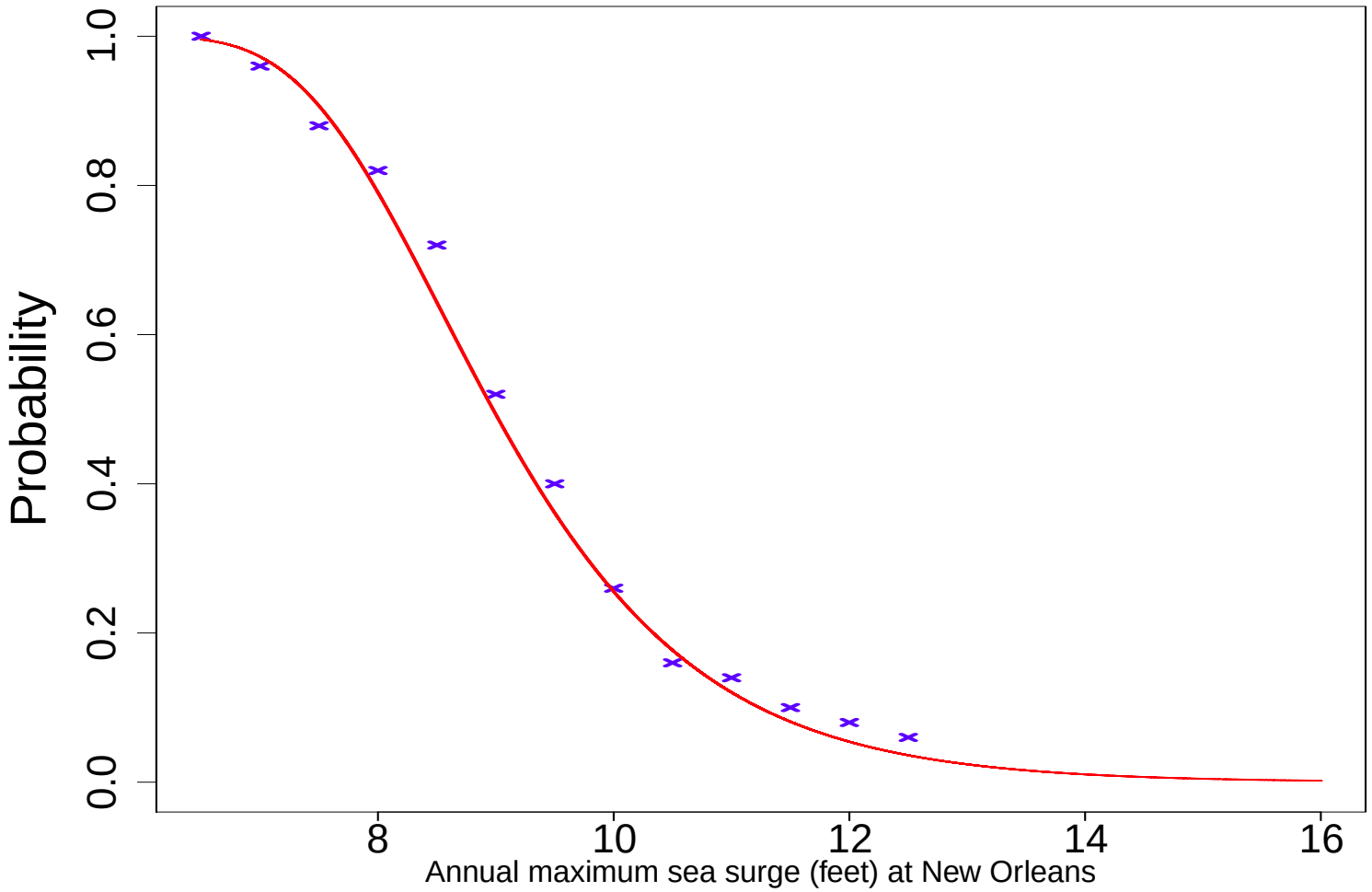
Part B: A probability model for extremes



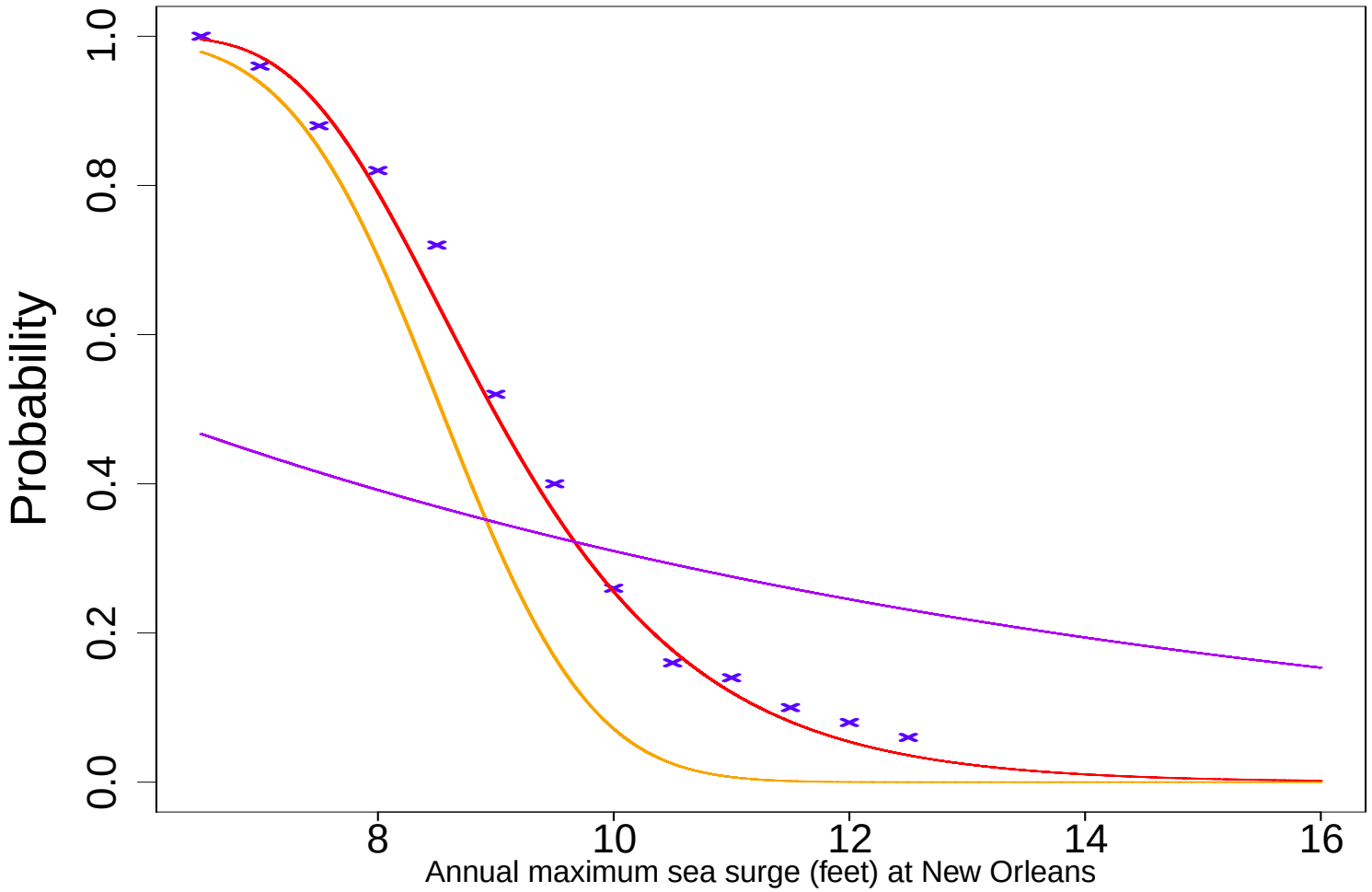
Part B: A probability model for extremes



Part B: A probability model for extremes



Part B: A probability model for extremes



Part B: A probability model for extremes

The Gumbel model does really well at predicting the sea–surge exceedance probabilities at New Orleans!

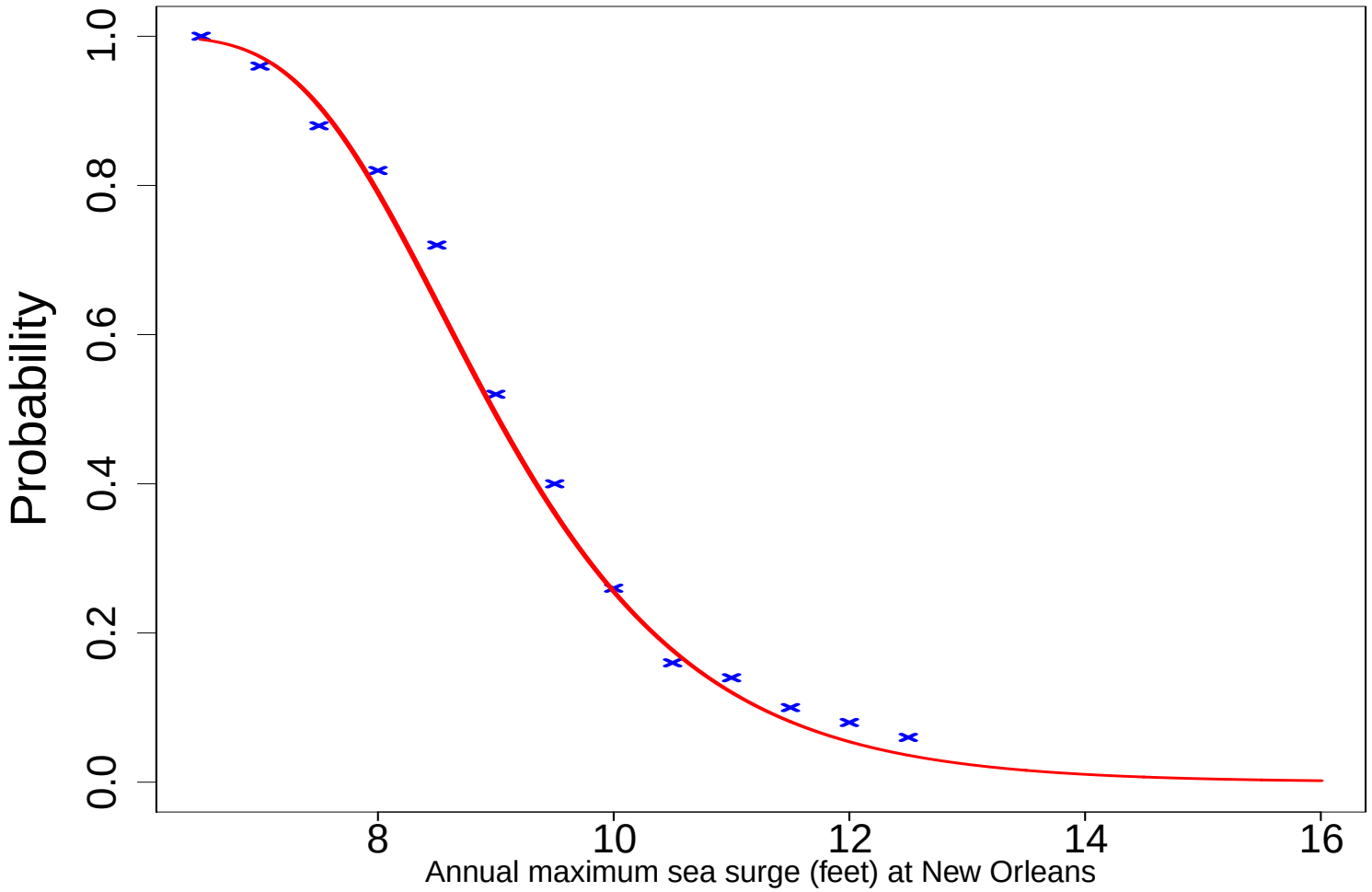
We can also estimate probabilities of events more extreme than those we have observed via **extrapolation**.

Part B: A probability model for extremes

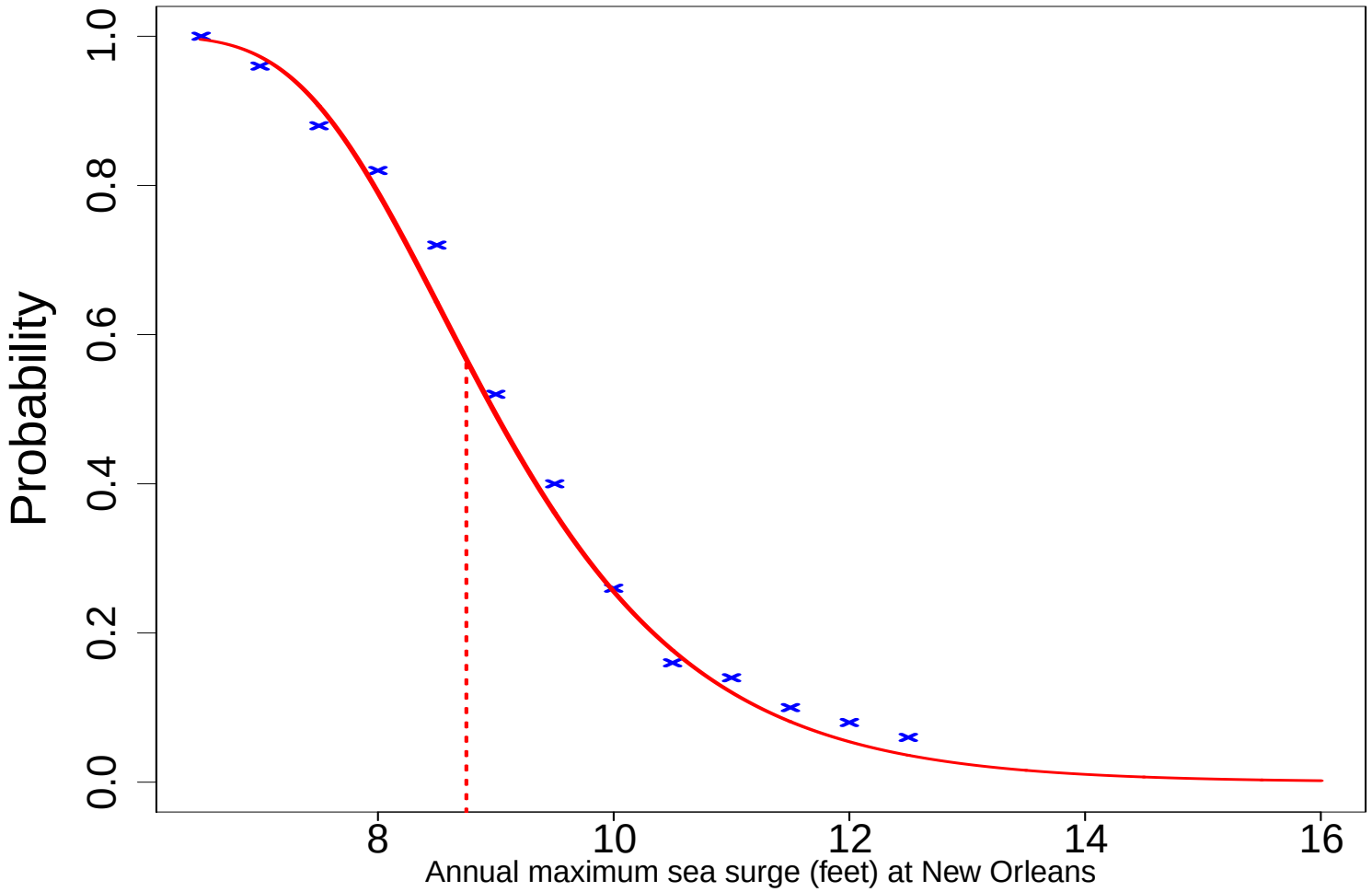
We can use our graph, or the values calculated using Gumbel's formula, to estimate the probability that, this year, the sea–surge at New Orleans will:

1. exceed 8.75 feet;
2. exceed 11.25 feet;
3. exceed 14 feet.

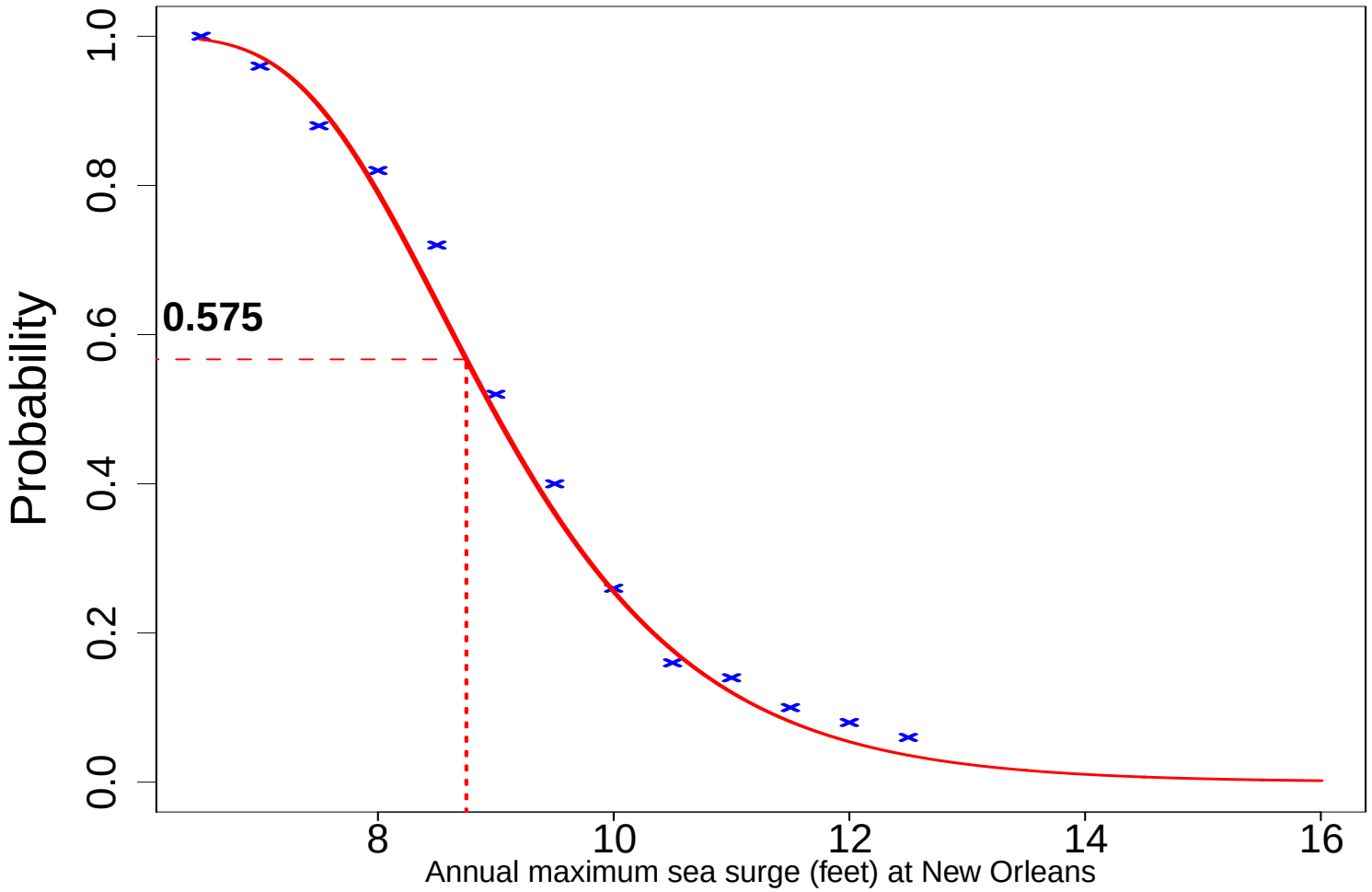
Part B: A probability model for extremes



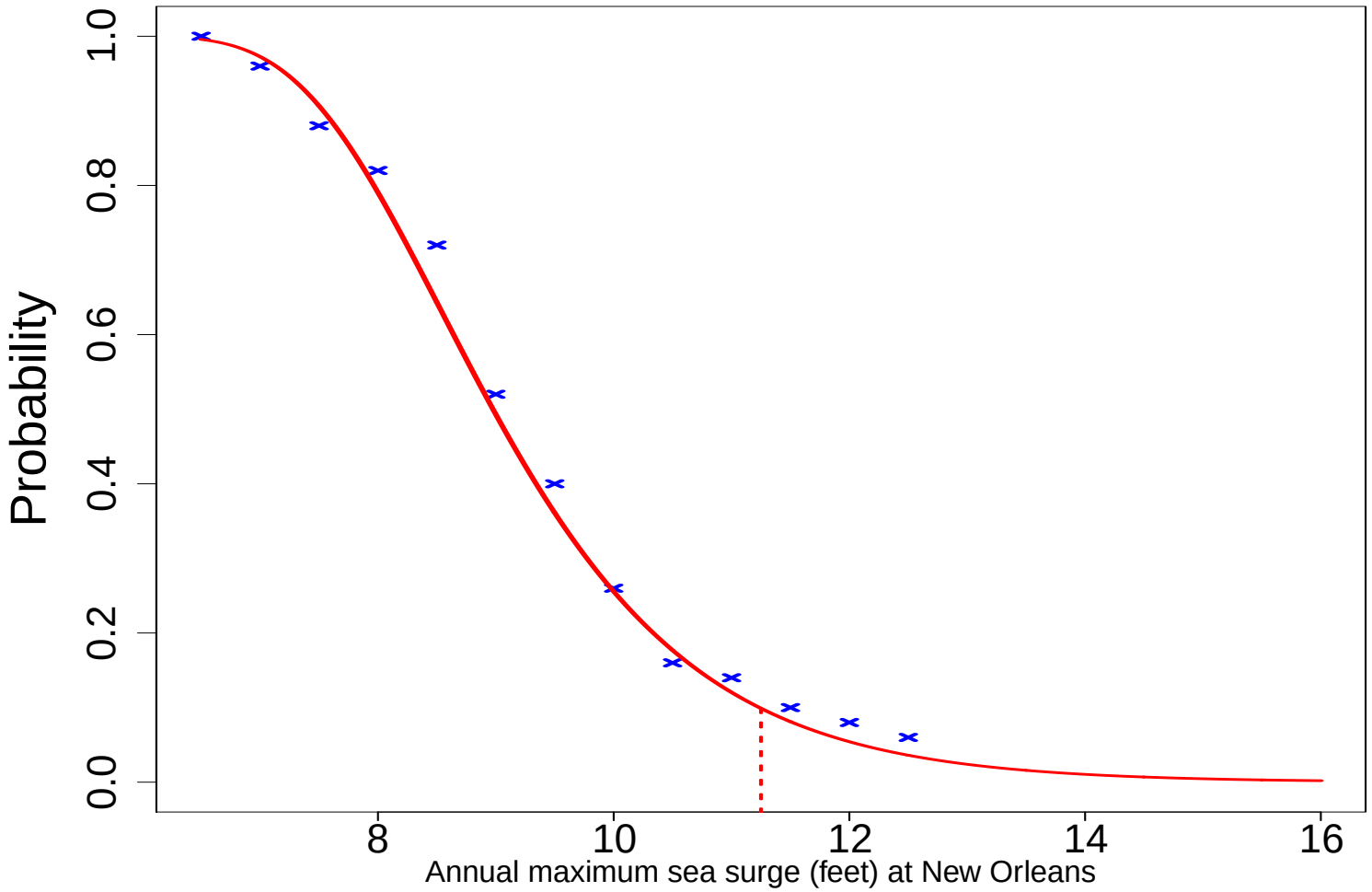
Part B: A probability model for extremes



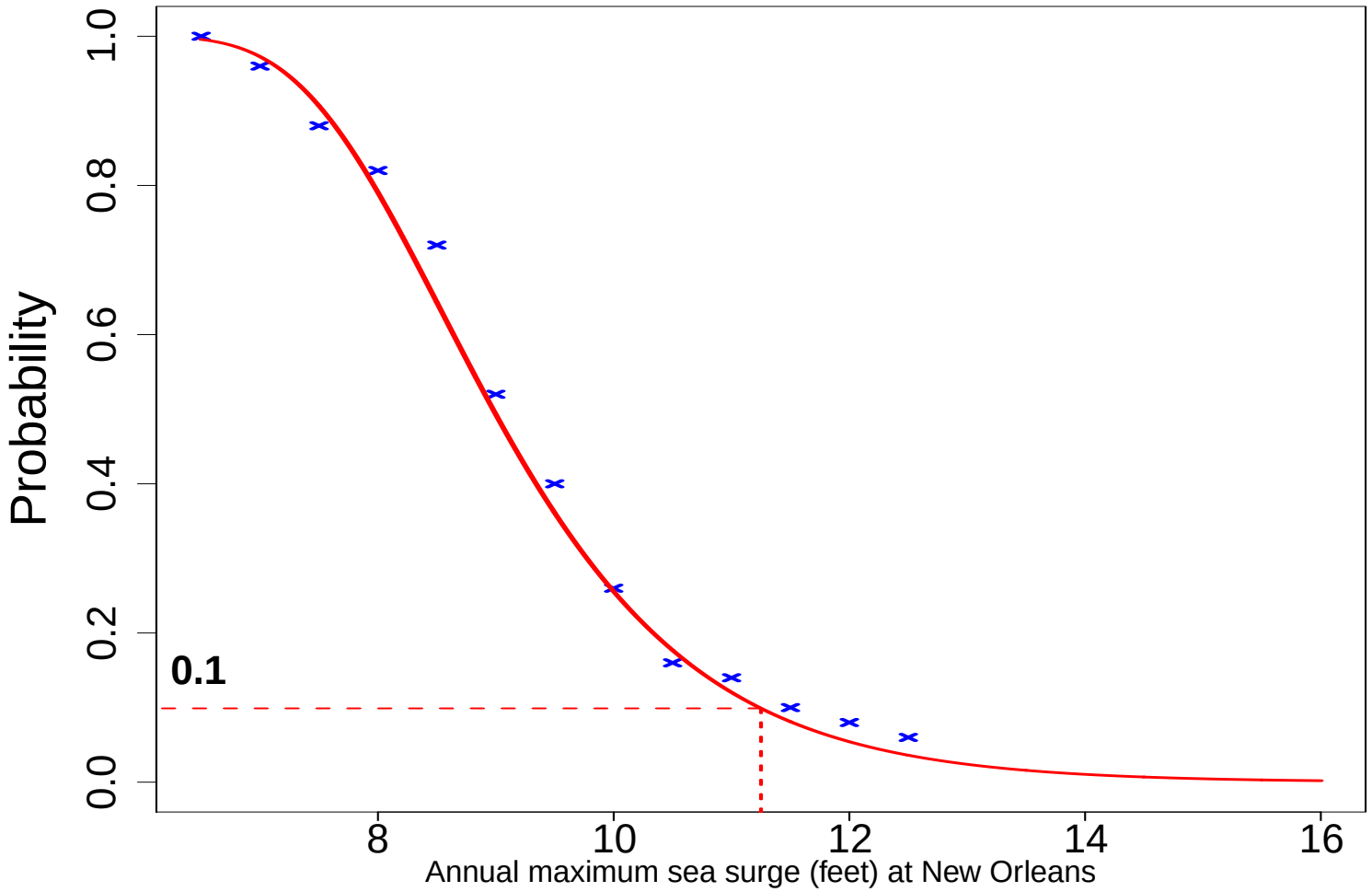
Part B: A probability model for extremes



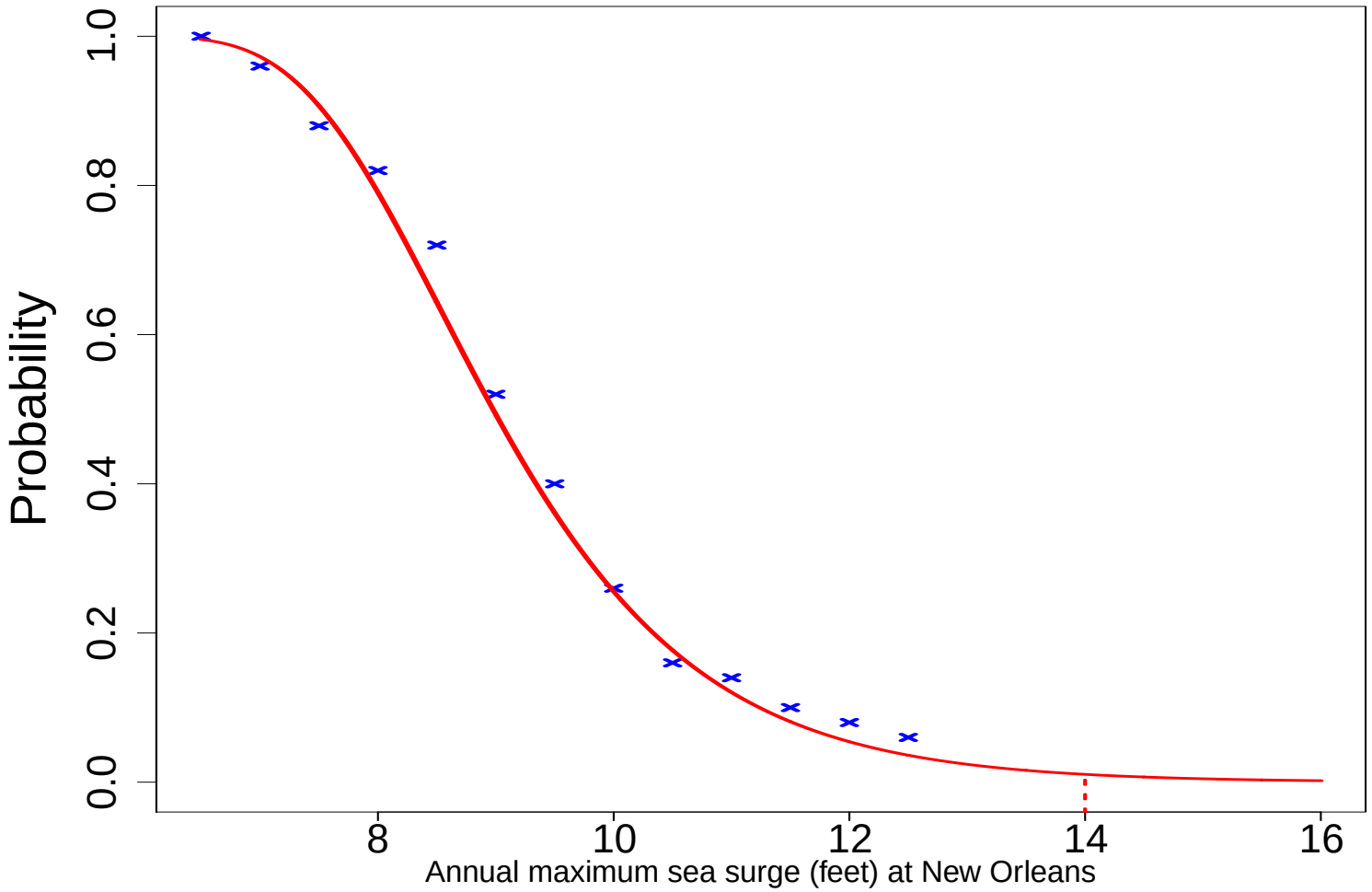
Part B: A probability model for extremes



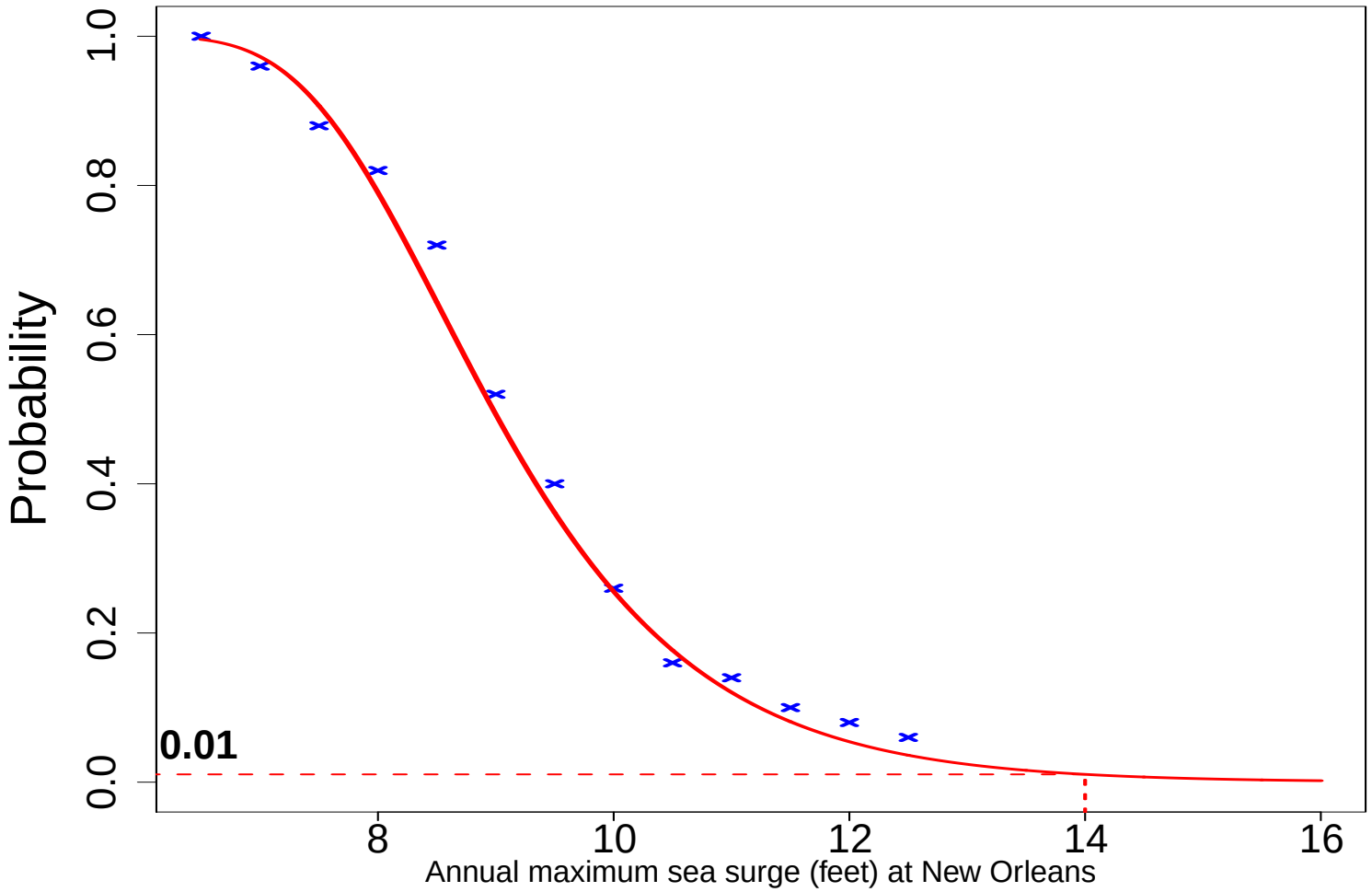
Part B: A probability model for extremes



Part B: A probability model for extremes



Part B: A probability model for extremes



Part B: A probability model for extremes

So we get:

Probabilities	Exceeds		
	8.75 feet	11.25 feet	14 feet
Data alone	0.66	0.12	0
Gumbel model	0.575	0.1	0.01

Part B: A probability model for extremes

So we get:

Probabilities	Exceeds		
	8.75 feet	11.25 feet	14 feet
Data alone	0.66	0.12	0
Gumbel model	0.575	0.1	**0.01**

Can you see why Katrina was billed as the "**storm of the century**"?

Part C: Application to structural design

As you might remember from the first part of this session, during Hurricane Katrina sea–surges exceeded 14 feet and parts of the sea wall system protecting the city were breached.

A new flood defence system is to be built; as the mathematician, you are asked how tall the sea wall should be to protect against the storm we might expect to see, on average, **once every 500 years**.

Part C: Application to structural design

Use the Gumbel model to help estimate the height of the new sea wall.

We want x such that

$$\boldsymbol{P}(\text{Sea–surge} > x) = \frac{1}{500}.$$

Using Gumbel's formula, this gives

$$1 - \exp\left[-\exp\left\{-\left(\frac{x - 8.536}{1.2}\right)\right\}\right] = \frac{1}{500}.$$

Part C: Application to structural design

Rearranging for x, we get:

$$\exp\left[-\exp\left\{-\left(\frac{x-8.536}{1.2}\right)\right\}\right] = \frac{499}{500}$$

$$-\exp\left\{-\left(\frac{x-8.536}{1.2}\right)\right\} = \ell_n\left(\frac{499}{500}\right)$$

Part C: Application to structural design

Rearranging for x, we get:

$$\exp\left[-\exp\left\{-\left(\frac{x-8.536}{1.2}\right)\right\}\right] = \frac{499}{500}$$

$$\exp\left\{-\left(\frac{x-8.536}{1.2}\right)\right\} = -\ell_n\left(\frac{499}{500}\right)$$

$$-\left(\frac{x-8.536}{1.2}\right) = \ell_n\left[-\ell_n\left(\frac{499}{500}\right)\right]$$

Part C: Application to structural design

Rearranging for x, we get:

$$\begin{aligned}
\exp\left[-\exp\left\{-\left(\frac{x-8.536}{1.2}\right)\right\}\right] &= \frac{499}{500} \\
\exp\left\{-\left(\frac{x-8.536}{1.2}\right)\right\} &= -\ell_n\left(\frac{499}{500}\right) \\
\frac{x-8.536}{1.2} &= -\ell_n\left[-\ell_n\left(\frac{499}{500}\right)\right] \\
x-8.536 &= -1.2\times\ell_n\left[-\ell_n\left(\frac{499}{500}\right)\right] \\
x &= -1.2\ell_n\left[-\ell_n\left(\frac{499}{500}\right)\right]+8.536 \\
&= 15.99233 \approx 16 \text{ feet.}
\end{aligned}$$

Further study in Maths

Why take Maths any further than GCSE?

- A level Maths gives you further “grounding” in the basics, and is seen as a ‘prestigious’ A level to have!
- Doing a degree in Maths/Statistics can be more “interesting”:
  - Lecturers are also (usually) **researchers**
  - Real–life research problems brought into the classroom
  - Excellent employment prospects if you do well!
- All areas of science require good mathematicians!