Data Science 1
Probability
Random Variables

Edward L. Boone

Introduction

From the last few videos we have seen that being able to count outcomes without enumerating them is very important.

Multiplication Rule: If E_1 can occur n_1 ways and E_2 can occur n_2 ways then the combined experiment can occur in $n_1 \times n_2$ ways.

Permutations: How many ways can I order n objects?

$$n! = n(n-1)(n-2)\cdots 3 \cdot 2 \cdot 1$$

Permutation of Subset: How many ways can I order k objects from a collection of n objects:

$$_nP_k = \frac{n!}{(n-k)!} = n(n-1)(n-2)\cdots(n-k+1)$$

Back to the beginning

We need a little more structure when we think of probability in order to make it easier to work with.

Looking at outcomes and events alone are very difficult to deal with because they change from experiment to experiment.

If we can think of these outcomes in a different manner we may be able to create some structure that helps us find commonalities between experiments.

Random Variable

A *Random Variable* is a function from the sample space $\mathcal{S}$ and the real line $\mathbb{R}$. Or more formally:

$$X : \mathcal{S} \to \mathbb{R}.$$

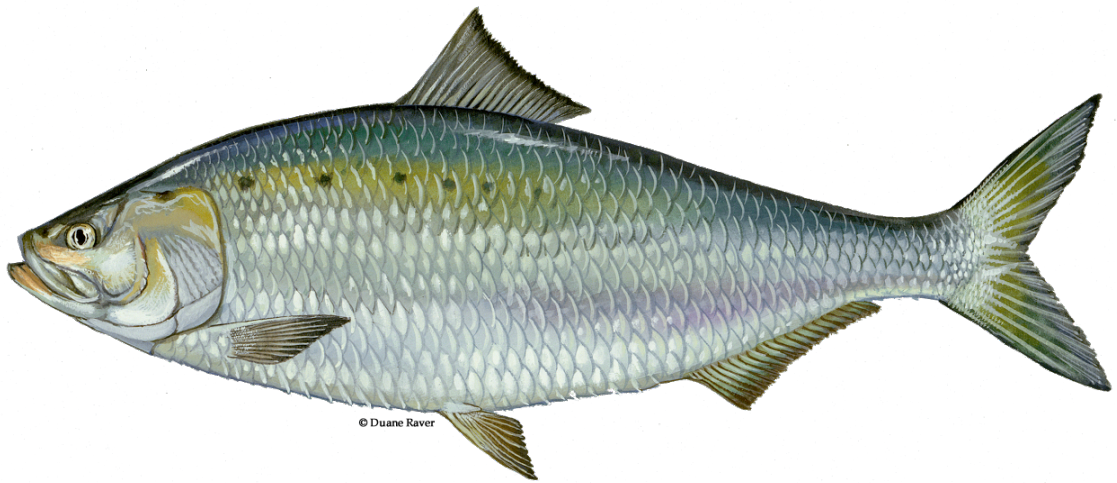
What does this mean?

This is basically how we measure the outcome.

How do we turn the outcome into a number?

Example

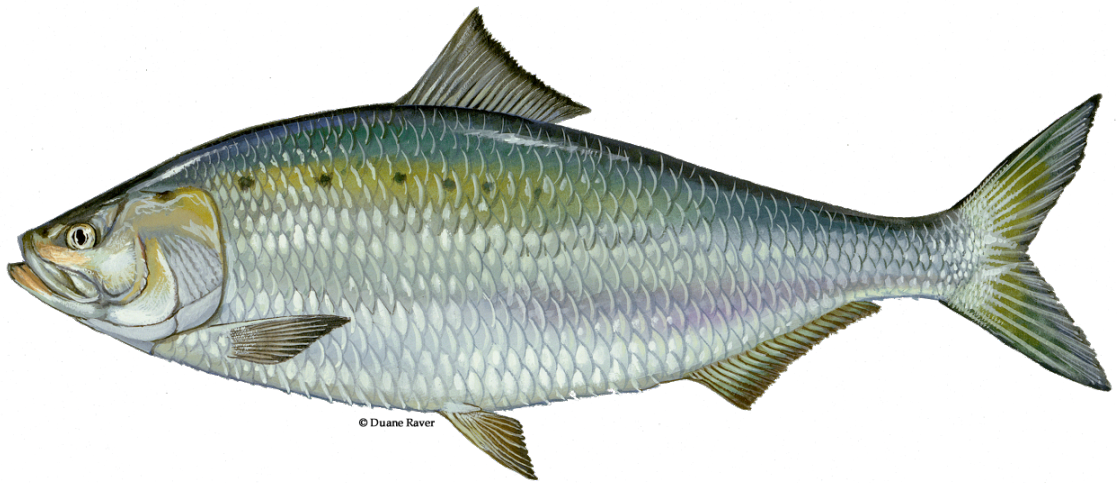
Suppose we are going fishing and we want to measure the outcome of our fishing.



How to turn this fish into a number?

Example

Suppose we are going fishing and we want to measure the outcome of our fishing.

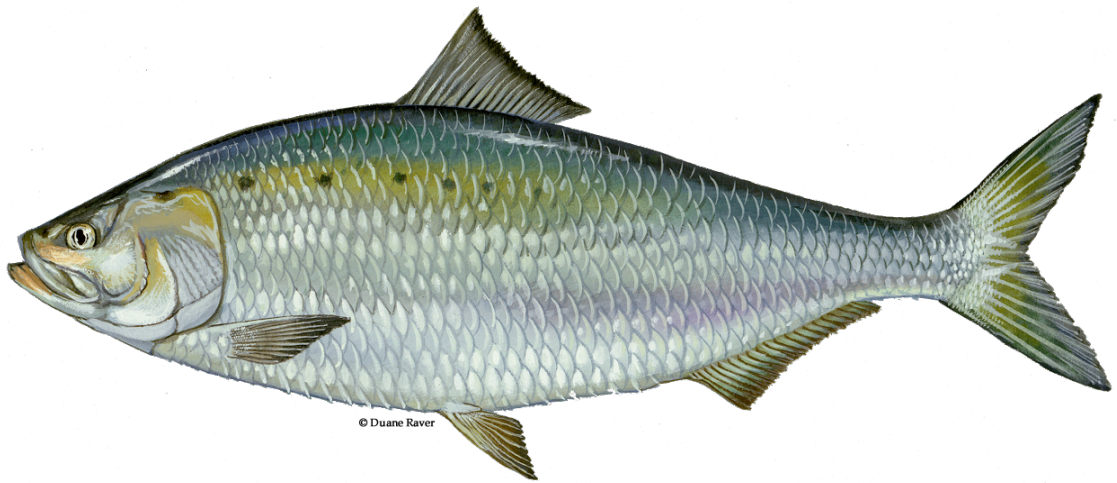


How to turn this fish into a number?

- Weight?

Example

Suppose we are going fishing and we want to measure the outcome of our fishing.

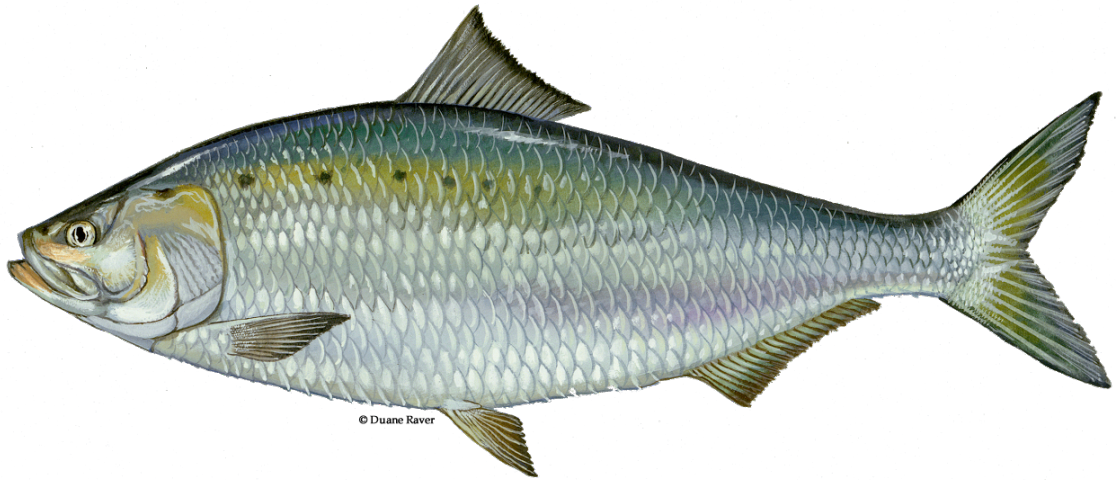


How to turn this fish into a number?

- Weight?
- Length?

Example

Suppose we are going fishing and we want to measure the outcome of our fishing.

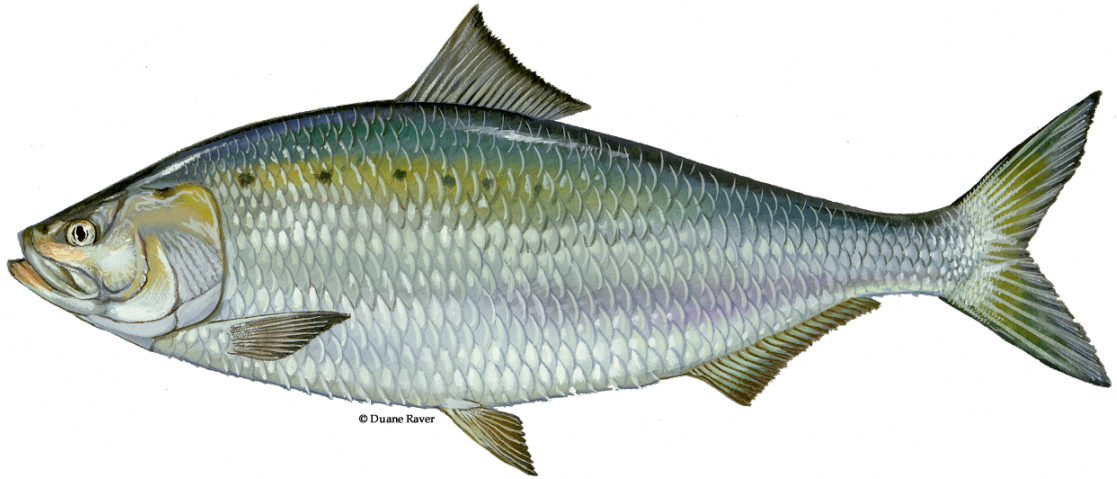


How to turn this fish into a number?

- Weight?
- Length?
- Width?

Example

Suppose we are going fishing and we want to measure the outcome of our fishing.

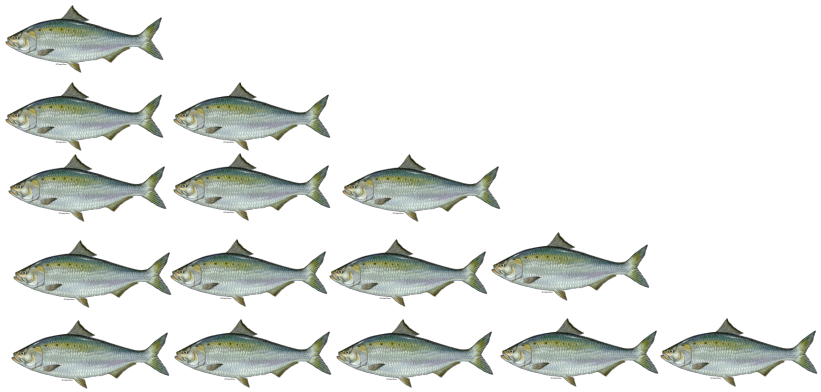


How to turn this fish into a number?

- Weight?
- Length?
- Width?
- Age?

Example

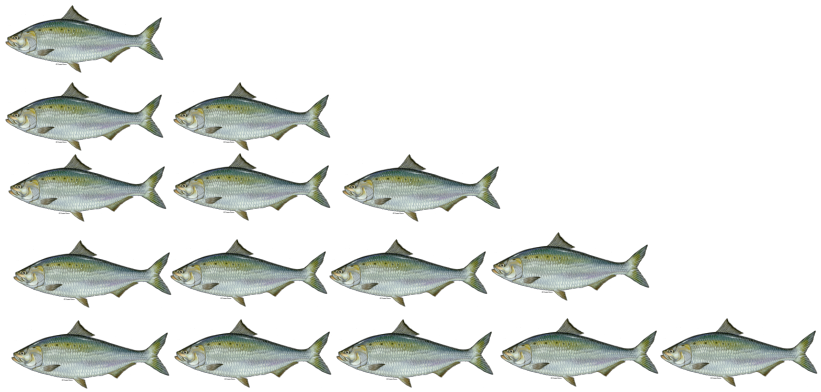
Suppose we are going fishing and we want to measure the outcome of our fishing.



How to turn these fish into a number?

Example

Suppose we are going fishing and we want to measure the outcome of our fishing.

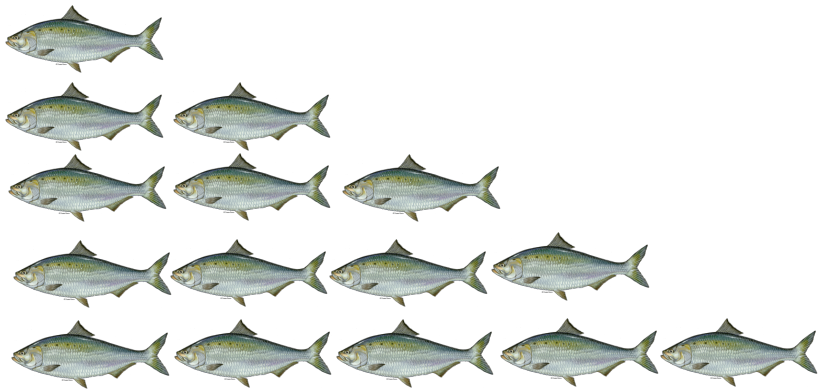


How to turn these fish into a number?

- The count of the number of fish?

Example

Suppose we are going fishing and we want to measure the outcome of our fishing.



How to turn these fish into a number?

- The count of the number of fish?
- Weight of all the fish? Weight of largest fish? Weight of smallest fish?

Random Variables

The random variable X is defined by the experimenter to be the attribute(s) that is most relevant to the study.

- It is not unique.
- Often depends on units of measurement.
- It is considered **random** because it varies for each trial of the experiment in an unpredictable way.

This will provide us with structure to help us understand probability better.