Random Variable

What's a random variable?

Formally

Random Variable

$X: \Omega \to \mathbb{R}$ **is a random variable**
$X(\omega)$ **is the summary of the outcome** ω

Informally: A random variable is a way to **summarize** the important (numerical) information from your outcome.

Try It Yourself

There are 20 balls, numbered 1,2,…,20 in an urn.

You'll draw out a size-three subset. (i.e. without replacement)

$\Omega = \{$size three subsets of $\{1, \dots, 20\}\}$, $\mathbb{P}()$ is uniform measure.

Let X be the largest value among the three balls.

If outcome is $\{4,2,10\}$ then $X = 10$.

Write down the pmf of X

Fill out the poll everywhere so Robbie knows how long to explain
Go to pollev.com/robbie

Application 1: Medical Tests

Helping Doctors and Patients Make Sense of Health Statistics

A researcher posed the following scenario to a group of 160 doctors:

Assume you conduct a disease screening using a standard test in a certain region. You know the following information about the people in this region:

The probability that a person has the disease is 1% (prevalence)

If a person has the disease, the probability that she tests positive is 90% (sensitivity)

If a person does not have the disease, the probability that she nevertheless tests positive is 9% (false-positive rate)

A person tests positive. She wants to know from you whether that means that she has the disease for sure, or what the chances are. What is the best answer?

A. The probability that she has the disease is about 81%.
B. Out of 10 people with a positive test, about 9 have the disease.
C. Out of 10 people with a positive test, about 1 have the disease.
D. The probability that she has the disease is about 1%

Pause for vocabulary

Physicians have words for just about everything

Let D be has the disease; T be test is positive

$\mathbb{P}(D)$ is "prevalence"

$\mathbb{P}(T|D)$ is "sensitivity"

A 'sensitive' test is one which picks up on the disease when it's there (high sensitivity -> few false negatives)

$\mathbb{P}\left(\overline{T}\middle|\overline{D}\right)$ is "specificity"

A 'specific' test is one that is positive specifically because of the disease, and for no other reason (high specificity -> few false positives)