

 Lagunita is retiring and will shut down at 12 noon Pacific Time on March 31, 2020. A few courses may be open for self-enrollment for a limited time. We will continue to offer courses on other online learning platforms; visit <http://online.stanford.edu>.

Course > Probability: Discrete Random Variables > Discrete Random Variables > Probability Distribution Defined

 Bookmark this page

Probability Distribution Defined

Learning Objective: Find the probability distribution of discrete random variables, and use it to find the probability of events of interest.

Probability Distribution

When we learned how to find probabilities by applying the basic principles, we generally focused on just one particular outcome or event, like the probability of getting exactly one tail when a coin is tossed twice, or the probability of getting a 5 when a die is rolled. Now that we have mastered the solution of individual probability problems, we'll proceed to look at the big picture by considering all the possible values of a discrete random variable, along with their associated probabilities. This list of possible values and probabilities is called the **probability distribution** of the random variable.

Comment

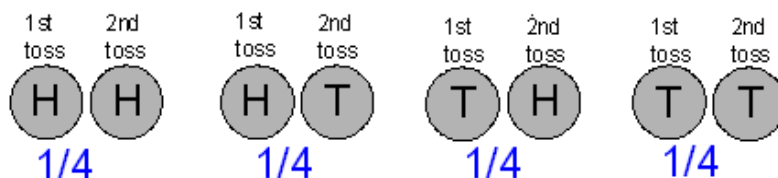
In the Exploratory Data Analysis sections of this course, we often looked at the distribution of sample values in a quantitative data set. We would display the values with a histogram, and summarize them by reporting their mean. In this section, when we look at the probability distribution of a random variable, we consider all its possible values and their overall probabilities of occurrence. Thus, we have in mind an entire population of values for a variable. When we display them with a histogram or summarize them with a mean, these are representing a population of values, not a sample. The distinction between sample and population is an essential concept in statistics, because an ultimate goal is to draw conclusions about unknown values for a population, based on what is observed in the sample.

Recall our first example, when we introduced the idea of a random variable. In this example we tossed a coin twice.

Example: Flipping a Coin Twice

What is the probability distribution of X , where the random variable X is the number of tails appearing in two tosses of a fair coin?

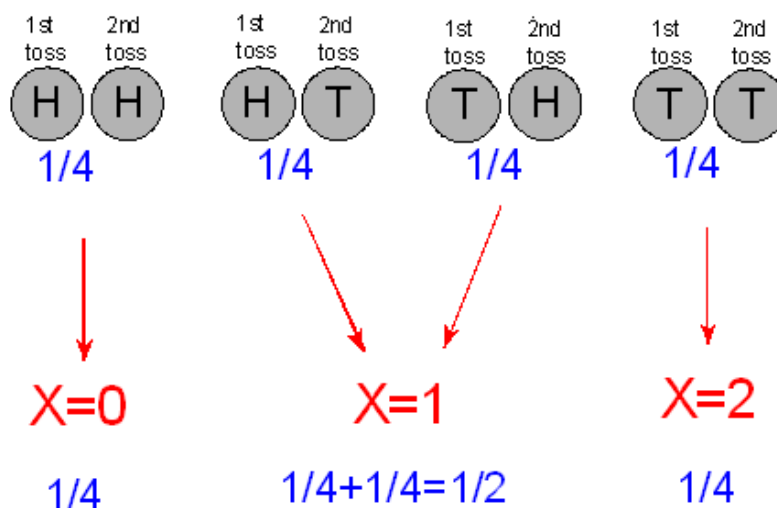
We first note that since the coin is fair, each of the four outcomes HH, HT, TH, TT in the sample space S is equally likely, and so each has a probability of $1/4$. (Alternatively, the multiplication principle can be applied to find the probability of each outcome to be $1/2 * 1/2 = 1/4$.)



X takes the value 0 only for the outcome HH, so the probability that $X = 0$ is $1/4$.

X takes the value 1 for outcomes HT or TH. By the addition principle, the probability that $X = 1$ is $1/4 + 1/4 = 1/2$.

Finally, X takes the value 2 only for the outcome TT, so the probability that $X = 2$ is $1/4$.



The **probability distribution of the random variable X** is easily summarized in a table:

List of possible values	x	0	1	2
Probability of each value	$P(X=x)$	$1/4$	$1/2$	$1/4$

As mentioned before, we write " $P(X = x)$ " to denote "the probability that the random variable X takes the value x ."

The way to interpret this table is:

X takes the values 0, 1, 2 and $P(X = 0) = 1/4$, $P(X = 1) = 1/2$, $P(X = 2) = 1/4$.

Note that events of the type $(X = x)$ are subject to the principles of probability established earlier, and will provide us with a way of systematically exploring the behavior of random variables. In particular, the first two principles in the context of probability distributions of random variables will now be stated.

Any probability distribution of a discrete random variable must satisfy:

1. $0 \leq P(X = x) \leq 1$

2. $\sum_x P(X = x) = 1$

Open Learning Initiative [↗](#)



[↗](#) Unless otherwise noted this work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License [↗](#).

© All Rights Reserved