

Probability Distributions: Takeaways

by Dataquest Labs, Inc. - All rights reserved © 2020

- Probability distributions can tell us how probability is distributed across the different outcomes of a random experiment
- Fair coin tosses, dice rolls and card draws all share a common theme: all of them assume that each outcome has an equal chance of occurring. This gives them a characteristic flat shape that defines the **uniform distribution**.
- Probability distributions can give us an idea of what values the random experiment can take on and what values are not possible in a random experiment.
- As its name suggests, a **probability distribution function** allows us to represent a probability distribution as a function and create a 1-to-1 correspondence between outcomes and the probability distributed to them. Mathematically, it is represented as $P(X=x)$.
- A **cumulative probability** is all of the probability up until a certain point. Mathematically, it is represented as $P(X \leq x)$.
- We can think of the collection of data as a random experiment. We cannot know in advance what we will measure, so it meets the requirement of a random experiment. Since data is a random experiment, we can associate it with both an empirical distribution and try to approximate it with a known probability distribution.
- The normal distribution is a symmetric, bell-shaped curve that is defined by a mean and a variance.
- We use the `dnorm()` function to calculate the probabilities of specific values in a normal distribution.
- We use the `pnorm()` function to calculate cumulative probabilities up until a specific value in a normal distribution.



Takeaways by Dataquest Labs, Inc. - All rights reserved © 2020