

How to Generate a Uniform Distribution in R: A Step-by-Step Guide

Authored by
stats writer

December 31, 2025

RECOMMENDED CITATION

stats writer (2025). *How to Generate a Uniform Distribution in R: A Step-by-Step Guide*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=110142>

The **R** statistical environment provides powerful tools for simulating data and analyzing theoretical distributions. When aiming to generate random numbers that follow a uniform distribution, the primary function utilized is **runif()**. This function is foundational for any simulation where every outcome within a specified interval must have an equal chance of occurrence. Mastering its use is essential for tasks ranging from Monte Carlo simulations to simple random sampling within defined limits, ensuring the generated data faithfully represents the underlying distributional assumption.

The runif() function requires two crucial parameters: the minimum value of the desired distribution (often denoted as 'a') and the maximum value (denoted as 'b'). By default, if these parameters are omitted, **runif()** generates values between 0 and 1, representing the standard uniform distribution. However, for practical applications, we typically define specific boundaries relevant to the problem domain. The result is a set of random numbers strictly constrained between the specified minimum and maximum values, ensuring that the probability density is constant across the entire range.

While **runif()** is the dedicated function for generating continuous uniform random variables, the base R function **sample()** offers an alternative approach, particularly useful when sampling from a discrete, defined set of uniform possibilities rather than a continuous range. When using **sample()**, the parameters specify the exact range of values to select from (the population) and the required number of values to return (the sample size). Although **sample()** is highly versatile, for generating continuous uniform random deviates, the dedicated runif() function remains the standard and most efficient tool for simulating this specific type of probability distribution.

Understanding the Continuous Uniform Distribution

The uniform distribution, often referred to as a rectangular distribution, is a type of probability distribution where every outcome between a lower bound (denoted as a) and an upper bound (denoted as b) is equally probable. This means that the probability density remains perfectly flat throughout the interval. This flatness simplifies probability calculations significantly compared to distributions like the normal or exponential distribution, where the probability density varies across the range, making the uniform distribution ideal for modeling situations lacking prior information about favoring any specific value.

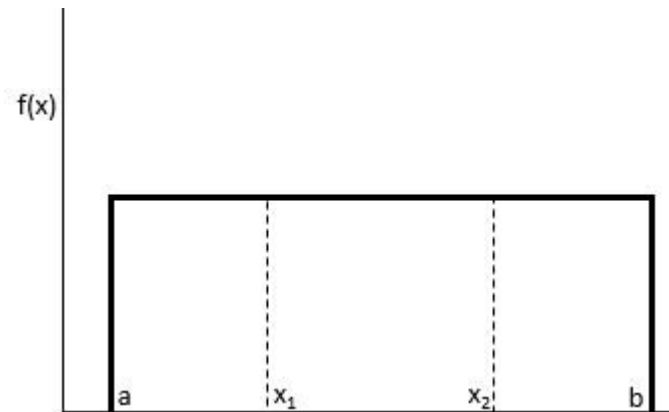
Mathematically, if we consider a continuous uniform distribution over the interval $[a, b]$, the Probability density function (PDF), $f(x)$, is defined as $1 / (b - a)$ for $a \leq x \leq b$, and 0 otherwise. This definition highlights why the total area under the curve--which must always equal 1 for any valid distribution--is maintained. If the height of the rectangle is $1 / (b - a)$ and the width is $(b - a)$, their product is exactly 1. This simple geometric interpretation forms the basis for all probability calculations derived from this model.

Calculating the probability that a value falls within a specific sub-interval within the larger range is straightforward. Because the density is constant, this probability is simply the ratio of the length of

the sub-interval to the length of the entire interval. The formula is intuitive and powerful for modeling situations where there is complete lack of knowledge about favoring any specific value within a given range, such as arrival times or measurement errors, and is expressed as:

$$P(\text{obtain value between } x_1 \text{ and } x_2) = (x_2 - x_1) / (b - a)$$

The visual representation below clearly shows the constant height and rectangular shape characteristic of a continuous uniform distribution bounded by parameters a and b .



Key Statistical Properties of the Uniform Distribution

The uniform distribution possesses distinct and easily calculable statistical measures that depend only on the boundary parameters, a and b . These properties are crucial for understanding the central tendency and spread of the data generated by this distribution. Since the distribution is perfectly symmetrical around its center, the mean and median will always coincide, simplifying descriptive analysis and ensuring that the center of the distribution is precisely halfway between its bounds.

The mean (μ) represents the average expected value in the distribution. Due to the symmetrical nature and constant density across the range, the mean simply lies exactly halfway between the minimum and maximum values. This is one of the most straightforward calculations in probability theory and provides immediate insight into the central location of the distributed values, contrasting sharply with skewed distributions where the mean, median, and mode often diverge.

In addition to the mean, understanding the variability is essential. The variance (σ^2) and standard deviation (σ) quantify the spread of the data points around the mean. Because the probability density is uniform, the variance calculation accounts for the width of the interval but is scaled by a factor of $1/12$, a result derived through integration of the expected value function. The standard deviation, being the square root of the variance, provides a measure of spread in the original units of the variable.

The uniform distribution has the following mathematical properties:

The mean of the distribution is $\mu = (a + b) / 2$

The variance of the distribution is $\sigma^2 = (b - a)^2 / 12$

The standard deviation of the distribution is $\sigma = \sqrt{\sigma^2}$

Core R Functions for Uniform Distribution Analysis

Beyond generating random deviates using **runif()**, the R statistical environment provides a suite of functions specifically designed to analyze the theoretical properties of the uniform distribution. These functions allow users to calculate densities, probabilities, and quantiles without needing complex manual integration or summation, making statistical problem-solving efficient and accurate. The standard naming convention in R uses prefixes (d, p, q, r) followed by the distribution name (**unif**).

The two most frequently used functions for analytical tasks are **dunif()** and **pnunif()**. The function **dunif(x, min, max)** calculates the height of the curve, or the Probability density function (PDF), for a given value x . This function tells us the relative likelihood of observing x , although for continuous variables, the density itself is not a probability. The parameters **min** and **max** define the bounds (a and b) of the uniform distribution being evaluated. Since the density is constant across the range, **dunif()** will return $1 / (\max - \min)$ for any x within the interval, and 0 otherwise.

In contrast, **pnunif(x, min, max)** calculates the Cumulative distribution function (CDF). The CDF represents the total probability that a random variable will take a value less than or equal to x . This is often the most critical function for solving real-world probability problems, as it directly gives cumulative probabilities. Similar to **dunif()**, the parameters **min** and **max** define the distributional boundaries. Understanding the difference between density (d-functions) and cumulative probability (p-functions) is vital for applying these tools correctly in statistical analysis.

dunif(x, min, max) - calculates the probability density function (pdf) for the uniform distribution where x is the value of a random variable, and *min* and *max* are the minimum and maximum numbers for the distribution, respectively.

pnunif(x, min, max) - calculates the cumulative distribution function (cdf) for the uniform distribution where x is the value of a random variable, and *min* and *max* are the minimum and maximum numbers for the distribution, respectively.

Applying pnunif() for Cumulative Probability (Example 1)

When approaching problems concerning arrival times, manufacturing tolerances, or other phenomena where outcomes are equally likely across a finite interval, the CDF function **pnunif()**

provides an elegant and concise solution. The primary utility of **punif()** is determining the probability accumulated up to a specific point, x , from the distribution's minimum boundary. This directly translates to solving "less than or equal to" probability questions, which are common in real-world statistical modeling.

Consider a practical scenario: a bus arrives at a specific stop every 20 minutes, meaning the waiting time follows a uniform distribution between 0 (immediate arrival) and 20 minutes (just missed the bus). If you arrive randomly, we want to know the probability that the bus shows up in 8 minutes or less. In this case, our interval is $[0, 20]$, and we are interested in the cumulative probability up to $x=8$. The fact that the distribution is uniform simplifies the theoretical calculation, but using R's dedicated function standardizes the procedure.

Since we require the cumulative probability from the lower boundary (0) up to our desired point (8), **punif()** is the perfect tool. We input the cutoff value (8), the minimum time (0), and the maximum time (20). R automatically calculates the ratio of the interval length $(8 - 0)$ to the total distribution length $(20 - 0)$, providing the exact probability. This application demonstrates the efficiency of using R's built-in distributional functions for rapid statistical computation.

Example 1: A bus shows up at a bus stop every 20 minutes. If you arrive at the bus stop, what is the probability that the bus will show up in 8 minutes or less?

Solution: We use the **punif()** function to find the cumulative probability that the bus arrives within 8 minutes, given the parameters $\text{min}=0$ and $\text{max}=20$:

```
punif(8, min=0, max=20)
```

```
## 0.4
```

The probability that the bus shows up in 8 minutes or less is **0.4**, or 40%.

Calculating Interval Probabilities (Example 2)

Statistical analysis often requires finding the likelihood that a random variable falls within a defined intermediate range, say between x_1 and x_2 . For the uniform distribution, while we could use the simple ratio formula, employing the CDF function **punif()** offers a standardized and highly reliable method. This technique relies on the general probability rule that the probability $P(x_1 < X < x_2)$ is equal to the difference between the cumulative probabilities: $P(X < x_2) - P(X < x_1)$.

To solve for an intermediate interval, we calculate the cumulative probability up to the upper limit (x_2) and then subtract the cumulative probability accumulated up to the lower limit (x_1). This effectively isolates the probability mass contained only within the segment (x_1, x_2) .

removing the probability associated with the lower tail of the distribution. This subtraction technique is universally applicable across all continuous probability distributions supported by R's p-functions, making it a valuable skill for any data analyst.

Consider a scenario involving biological measurement, where the weight of a frog species is uniformly distributed between 15 and 25 grams. If we want the probability that a randomly selected frog weighs between 17 and 19 grams, we set $a=15$, $b=25$, $x_1=17$, and $x_2=19$. We first calculate $P(X < 19)$ using **punif(19, 15, 25)** and then subtract $P(X < 17)$ using **punif(17, 15, 25)**. This powerful approach ensures precision and adherence to rigorous statistical methodology.

Example 2: The weight of a certain species of frog is uniformly distributed between 15 and 25 grams. If you randomly select a frog, what is the probability that the frog weighs between 17 and 19 grams?

Solution: To find the solution, we calculate the cumulative probability of a frog weighing less than 19 grams and subtract the cumulative probability of a frog weighing less than 17 grams:

```
punif(19, 15, 25) - punif(17, 15, 25)
```

```
## 0.2
```

Thus, the probability that the frog weighs between 17 and 19 grams is **0.2**.

Determining Upper-Tail Probabilities (Example 3)

A final common problem type involves calculating the probability that a random variable exceeds a certain value, x . This is known as calculating an upper-tail probability, or $P(X > x)$. Since the total probability space for any continuous distribution must sum to 1, this probability is found by taking the complement of the cumulative probability: $P(X > x) = 1 - P(X \leq x)$. This complementary rule is a fundamental concept in probability theory and is highly effective when using R's p-functions.

Using the **punif()** function, we calculate $P(X \leq x)$ just as before, supplying the point x and the distribution bounds. We then subtract this result from 1. This method avoids potential pitfalls or complexities that might arise from manual calculation, especially when dealing with non-standard distribution bounds, and leverages the built-in precision of the R environment, which handles the necessary normalization inherently within the **punif()** call.

Consider a scenario where the length of an NBA game is uniformly distributed between 120 and 170 minutes. We are interested in the probability that a game lasts more than 150 minutes. We define the parameters as $a=120$ and $b=170$. We calculate $P(X \leq 150)$ using **punif(150,**

120, 170) and then apply the complement rule. This provides the probability of the event occurring in the upper tail of the uniform distribution, yielding the desired result with minimal calculation effort.

Example 3: The length of an NBA game is uniformly distributed between 120 and 170 minutes. What is the probability that a randomly selected NBA game lasts more than 150 minutes?

Solution: To answer this question, we use the formula $1 - P(X \leq 150)$, where $P(X \leq 150)$ is calculated using the cumulative distribution function:

```
1 - punif(150, 120, 170)
```

```
## 0.4
```

The probability that a randomly selected NBA game lasts more than 150 minutes is **0.4**.