

Discrete uniform distribution

In probability theory and statistics, the **discrete uniform distribution** is a symmetric probability distribution wherein a finite number of values are equally likely to be observed; every one of n values has equal probability $1/n$. Another way of saying "discrete uniform distribution" would be "a known, finite number of outcomes equally likely to happen".

A simple example of the discrete uniform distribution is throwing a fair die. The possible values are 1, 2, 3, 4, 5, 6, and each time the die is thrown the probability of a given score is $1/6$. If two dice are thrown and their values added, the resulting distribution is no longer uniform because not all sums have equal probability. Although it is convenient to describe discrete uniform distributions over integers, such as this, one can also consider discrete uniform distributions over any finite set. For instance, a random permutation is a permutation generated uniformly from the permutations of a given length, and a uniform spanning tree is a spanning tree generated uniformly from the spanning trees of a given graph.

The discrete uniform distribution itself is inherently non-parametric. It is convenient, however, to represent its values generally by all integers in an interval $[a,b]$, so that a and b become the main parameters of the distribution (often one simply considers the interval $[1,n]$ with the single parameter n). With these conventions, the cumulative distribution function (CDF) of the discrete uniform distribution can be expressed, for any $k \in [a,b]$, as

$$F(k; a, b) = \frac{\lfloor k \rfloor - a + 1}{b - a + 1}$$

Contents

[Estimation of maximum](#)

[Random permutation](#)

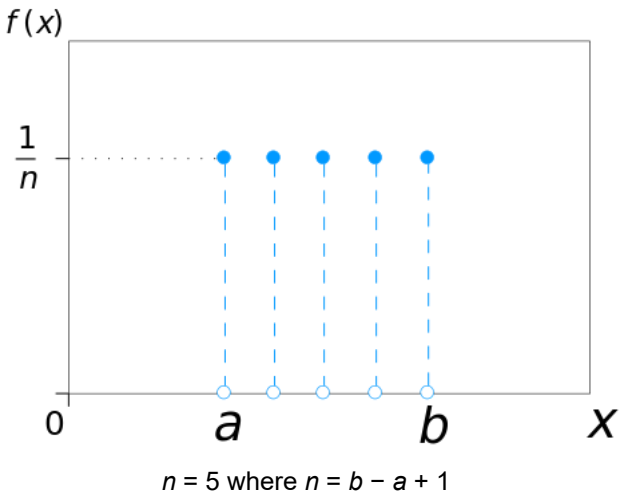
[Properties](#)

[See also](#)

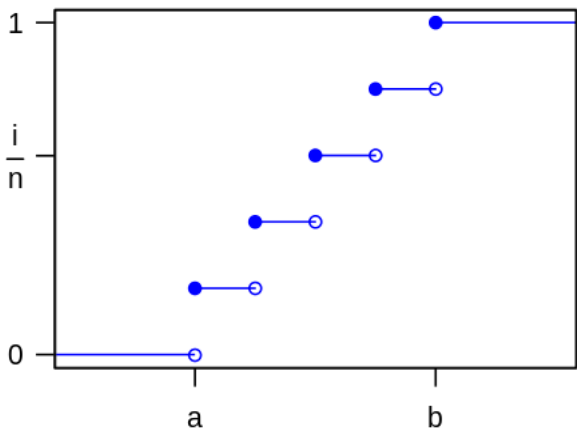
[References](#)

discrete uniform

Probability mass function



Cumulative distribution function



Notation	$\mathcal{U}\{a, b\}$ or $\text{unif}\{a, b\}$
Parameters	a, b integers with $b \geq a$ $n = b - a + 1$
Support	$k \in \{a, a + 1, \dots, b - 1, b\}$
PMF	$\frac{1}{n}$
CDF	$\frac{\lfloor k \rfloor - a + 1}{n}$
Mean	$\frac{a + b}{2}$
Median	$\frac{a + b}{2}$
Mode	N/A
Variance	$\frac{(b - a + 1)^2 - 1}{12}$

Estimation of maximum

This example is described by saying that a sample of k observations is obtained from a uniform distribution on the integers $1, 2, \dots, N$, with the problem being to estimate the unknown maximum N . This problem is commonly known as the German tank problem, following the application of maximum estimation to estimates of German tank production during World War II.

The uniformly minimum variance unbiased (UMVU) estimator for the maximum is given by

$$\hat{N} = \frac{k+1}{k}m - 1 = m + \frac{m}{k} - 1$$

where m is the sample maximum and k is the sample size, sampling without replacement.^[1] This can be seen as a very simple case of maximum spacing estimation.

This has a variance of^[1]

$$\frac{1}{k} \frac{(N-k)(N+1)}{(k+2)} \approx \frac{N^2}{k^2} \text{ for small samples } k \ll N$$

so a standard deviation of approximately $\frac{N}{k}$, the (population) average size of a gap between samples; compare $\frac{m}{k}$ above.

The sample maximum is the maximum likelihood estimator for the population maximum, but, as discussed above, it is biased.

If samples are not numbered but are recognizable or markable, one can instead estimate population size via the capture-recapture method.

Random permutation

See rencontres numbers for an account of the probability distribution of the number of fixed points of a uniformly distributed random permutation.

Properties

The family of uniform distributions over ranges of integers (with one or both bounds unknown) has a finite-dimensional sufficient statistic, namely the triple of the sample maximum, sample minimum, and sample size, but is not an exponential family of distributions, because the support varies with the parameters. For families whose support does not depend on the parameters, the Pitman–Koopman–Darmois theorem states that only exponential families have a sufficient statistic whose dimension is bounded as sample size increases. The uniform distribution is thus a simple example showing the limit of this theorem.

See also

- Dirac delta distribution

Skewness	0
Ex. kurtosis	$\frac{6(n^2 + 1)}{5(n^2 - 1)}$
Entropy	$\ln(n)$
MGF	$\frac{e^{at} - e^{(b+1)t}}{n(1 - e^t)}$
CF	$\frac{e^{iat} - e^{i(b+1)t}}{n(1 - e^{it})}$
PGF	$\frac{z^a - z^{b+1}}{n(1 - z)}$

- Uniform distribution (continuous)

References

1. Johnson, Roger (1994), "Estimating the Size of a Population", *Teaching Statistics*, **16** (2 (Summer)): 50–52, CiteSeerX 10.1.1.385.5463 (<https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.385.5463>), doi:10.1111/j.1467-9639.1994.tb00688.x (<https://doi.org/10.1111%2Fj.1467-9639.1994.tb00688.x>)
-

Retrieved from "https://en.wikipedia.org/w/index.php?title=Discrete_uniform_distribution&oldid=979586022"

This page was last edited on 21 September 2020, at 15:34 (UTC).

Text is available under the Creative Commons Attribution-ShareAlike License; additional terms may apply. By using this site, you agree to the Terms of Use and Privacy Policy. Wikipedia® is a registered trademark of the Wikimedia Foundation, Inc., a non-profit organization.