

# How do you calculate the probability density function of the maximum of a sample of IID uniform random variables?

Asked 7 years, 10 months ago   Active 1 year, 3 months ago   Viewed 83k times



Given the random variable

45

$$Y = \max(X_1, X_2, \dots, X_n)$$



where  $X_i$  are IID uniform variables, how do I calculate the PDF of  $Y$ ?



pdf

maximum

24

edited Jul 20 '12 at 18:50

asked Nov 15 '11 at 19:34



Macro

33.8k

6

128

142



Mascarpone

368


1

3

6

4



If this is homework, please read the FAQ and update your question accordingly. – cardinal Nov 15 '11 at 20:01 



Can one use Vandermonde's identity to show joint function of 2 order Statistics say  $F_{-Y}(r) * G_{-Y}(r)$  ? – larry mintz Jun 9 '18 at 21:10



Out of interest, what course covers this kind of problem? It is not something that I encountered in my engineering probability course. – Alex Aug 26 at 22:27



## 4 Answers



63

*It is possible that this question is homework but I felt this classical elementary probability question was still lacking a complete answer after several months, so I'll give one here.*

From the problem statement, we want the distribution of

$$Y = \max\{X_1, \dots, X_n\}$$



where  $X_1, \dots, X_n$  are iid  $\text{Uniform}(a, b)$ . We know that  $Y < x$  if and only if every element of the sample is less than  $x$ . Then this, as indicated in @varty's hint, combined with the fact that the  $X_i$ 's are independent, allows us to deduce

$$P(Y \leq x) = P(X_1 \leq x, \dots, X_n \leq x) = \prod_{i=1}^n P(X_i \leq x) = F_X(x)^n$$

where  $F_X(x)$  is the [CDF of the uniform distribution](#). Therefore the CDF of  $Y$  is

$$F_Y(y) = P(Y \leq y) = \begin{cases} 0 & y \leq a \\ [(y-a)/(b-a)]^n & y \in (a, b) \\ 1 & y \geq b \end{cases}$$

Since  $Y$  has an absolutely continuous distribution [we can derive its density by differentiating the CDF](#). Therefore the density of  $Y$  is

$$p_Y(y) = \frac{n(y-a)^{n-1}}{(b-a)^n}$$

In the special case where  $a = 0, b = 1$ , we have that  $p_Y(y) = ny^{n-1}$ , which is the density of a [Beta distribution](#) with  $\alpha = n$  and  $\beta = 1$ , since

$$\text{Beta}(n, 1) = \frac{\Gamma(n+1)}{\Gamma(n)\Gamma(1)} = \frac{n!}{(n-1)!} = n.$$

As a note, the sequence you get if you were to sort your sample in increasing order -  $X_{(1)}, \dots, X_{(n)}$  - are called the [order statistics](#). A generalization of this answer is that [all order statistics of a Uniform\(0, 1\) distributed sample have a Beta distribution](#), as noted in @bnaul's answer.

edited Jul 17 '12 at 15:12

answered Jul 15 '12 at 22:52



**Macro**

**33.8k** 6 128 142



This actually was a homework question for me. Thanks for the explanation. – Paul P M Mar 26 '16 at 22:20



i feel like i should be able to take your insights here and answer [this question](#), but i'm not seeing how to do that. can you help me out? can you recommend a textbook or chapter that speaks to this general issue? – user28511 Mar 31 '16 at 20:38



@PaulPM Out of interest, what course covers this kind of problem? It is not something that I encountered in my engineering probability course. – Alex Aug 26 at 22:27

The maximum of a sample is one of the [order statistics](#), in particular the  $n$ th order statistic of the sample  $X_1, \dots, X_n$ . In general, computing the distribution of order statistics is difficult, as described by the Wikipedia article; for some special distributions, the order statistics are well-known (e.g. for the uniform distribution, which has Beta-distributed order statistics).

EDIT: The Wikipedia article on [sample maximum and minimum](#) is also helpful and more specific to your problem.

answered Nov 15 '11 at 19:55



**bnaul**

2,221

1

13

14

5 ▲ For distributions with densities, computing the marginal distribution of a particular order statistic is quite straightforward. It is even easier for "special" order statistics like the minimum and maximum. – [cardinal](#) Nov 15 '11 at 20:03 ✎

▲ I guess it depends on what is meant by "calculate" in the original question. Certainly doing so numerically is straightforward; I interpreted the question as asking how to find a closed form solution, which is in general not easy. – [bnaul](#) Nov 15 '11 at 20:59

8 ▲ @bnaul: Let  $F(x) = \mathbb{P}(X \leq x)$  be an *arbitrary* distribution function and let  $X_1, \dots, X_n$  be an iid sample from  $F$ . Let  $X_{(k)}$  be the  $k$ th order statistic. Then

$$\mathbb{P}(X_{(k)} \leq x) = \sum_{m=k}^n \mathbb{P}(|\{i : X_i \leq x\}| = m) = \sum_{m=k}^n \binom{n}{m} F(x)^m (1 - F(x))^{n-m}.$$

**QED.** – [cardinal](#) Nov 15 '11 at 21:16 ✎

1 ▲ Perhaps a way to understand cardinals answer (given that you understand order statistic for uniform) is that because cdfs are monotonic 1-to-1 transformations of a uniform cdf, we can always express the event  $\{X < a\}$  in terms of a uniform random variable (this is why monte carlo works). So any result based on a uniform distribution will easily generalise to other random variables - just apply the transformation  $U = F_X(X)$ . – [probabilityislogic](#) Feb 14 '12 at 1:01 ✎

2 ▲ @probabilityislogic: The intuition is good, though it seems you have continuous random variables in mind in your comment. (The result in my second comment above, e.g., works for an arbitrary distribution function.) – [cardinal](#) Feb 21 '12 at 13:04

|

▲ If  $F_Y(y)$  is the CDF of  $Y$ , then

1

$$F_Y(y) = \text{Prob}(y > X_1, y > X_2, \dots, y > X_n)$$

▲ You can then use the iid property and the cdf of a uniform variate to compute  $F_Y(y)$ .

edited Jul 20 '12 at 0:48



**Macro**

33.8k

6

128

142

answered Nov 15 '11 at 20:04



**varty**

1,236

8

5

-3

▲ The maximum of a set of IID random variables when appropriately normalized will generally converge to one of the three extreme value types. This is Gnedenko's theorem, the equivalence of the central limit theorem for extremes. The particular type depends on the tail behavior of the population distribution. Knowing this you can use the limiting distribution to approximate the distribution for the maximum.

Since the uniform distribution on  $[a, b]$  is the subject of this question Macro has given the exact distribution for any  $n$  and a very nice answer. The result is rather trivial. For the normal distribution a nice closed form is not possible but appropriately normalized the maximum for the normal converges to the Gumbel distribution  $F(x) = \exp(-e^{-x})$ .

For the uniform the normalization is  $(b-a)-x/n$  and  $F^n(b-a-x/n) = (1-x/[n(b-a)])^n$

which converges to  $e^{-x/(b-a)}$ . Note here that  $y = b-a-x/n$ . and  $F^n(y)$  converges to 1 as  $y$  goes to  $b-a$ . This holds for all 0

In this case it is easy to compare the exact value to its asymptotic limit.

[Gumbel's book](#)

[Galambos' book](#)

[Leadbetter's book](#)

[Novak's book](#)

[Coles book](#)

edited Jul 24 '12 at 19:24

user10525

answered Jul 20 '12 at 0:59



Michael Chernick

35.8k 8 63 134

4 ▲ For this answer to be practicable, you need to stipulate, in detail, how one "appropriately normalizes" the values and you also need to provide some way to estimate how large  $n$  must be before the asymptotic formula becomes a reliable approximation. – whuber ♦ Jul 20 '12 at 15:27

▲ @whuber Anyone can look at Gnedenko's theorem to see the normalization. Equally as important is the tail characteristics that determine which of the three types applies. The theorem generalizes to stationary stochastic processes. So anyone who want to know the nitty gritty details can look at Leadbetter's book or my PhD thesis. When  $n$  is large enough is a difficult question to answer for any form of asymptotics. I guess the Berry-Esseen theorem helps for the central limit theorem. I don't know what is comparable for extremes. – Michael Chernick Jul 20 '12 at 16:57