

Min and Max Uniform Extreme Interval Values and Statistics

Marsha Jance (mjance@comcast.net)

Nick Thomopoulos (thomop@stuart.iit.edu)

Illinois Institute of Technology Stuart School of Business

Abstract

The min and max uniform extreme interval values and statistics: expected value, standard deviation, mode, median, and coefficient of variation are discussed. An extreme interval value g_α is defined as a numerical bound where a specified percentage α of the data is less than g_α . A numerical example and an analysis of the min and max extreme interval values and statistics are provided. In addition, a procedure for finding the min and max extreme interval values for different uniform parameter values and an application of this research are presented.

Introduction

The extreme interval values and statistics for the smallest (min) and largest (max) observations in a sample of size n are discussed. The observations in the sample come from a uniform population with parameters $a = 0$ and $b = 1$, and the sample sizes range from $n = 1$ to $n = 1000$. An extreme interval value g_α is defined as a numerical bound where a specified percentage α of the data is less than g_α . For example, when the probability is $P(g < g_{0.15}) = 0.15$, $g_{0.15}$ is the extreme interval value and 15% of the data is below $g_{0.15}$.

The basis for this paper comes from the dissertation and research of Jance's *Calculating Min and Max Extreme Interval Values for Various Distributions*. Jance developed Excel Visual Basic for Applications (VBA) programs to find the min and max extreme interval values and statistics for observations that come from uniformly, exponentially, and normally distributed populations. The dissertation includes tables, graphs, and applications of this research. The tables display the min and max extreme interval values and statistics for a range of observation sizes.

Min and Max Extreme Interval Values and Statistics

Suppose several samples of size n are taken from a continuous population with probability density function $f(x)$ and cumulative distribution function $F(x)$. The smallest (min) and largest (max) observations will vary from sample to sample. As a result, the min and max values will have a probability density function, cumulative distribution function, and statistics associated with them.

The min probability density function is $h(g) = nf(g)(1 - F(g))^{n-1}$ (Hines, Montgomery, Goldsman, and Borrer 215). When the observations come from a uniformly distributed population with parameters $a = 0$ and $b = 1$, the min uniform probability density

function is $h(g) = n(1 - g)^{n-1}$. In addition, the min uniform cumulative distribution function is $H(g) = \int_0^g h(g)dg$, the min uniform expected value is $E(g) = \mu = \int_0^1 gh(g)dg$, and the min uniform variance is $\sigma^2 = E(g^2) - E(g)^2$ where $E(g^2) = \int_0^1 g^2 h(g)dg$.

The max probability density function is $h(g) = ng^{n-1}$ (Hines, Montgomery, Goldsman, and Borror 215). When the observations come from a uniformly distributed population with parameters a = 0 and b = 1, the max uniform probability density function is $h(g) = ng^{n-1}$. In addition, the max uniform cumulative distribution function is $H(g) = \int_0^g h(g)dg$, the max uniform expected value is $E(g) = \mu = \int_0^1 gh(g)dg$, and the max uniform variance is $\sigma^2 = E(g^2) - E(g)^2$ where $E(g^2) = \int_0^1 g^2 h(g)dg$.

An Excel VBA application was written to find the min and max uniform extreme interval values and statistics for observation sizes ranging from n = 1 to n = 1000. The extreme interval values are found for probabilities ranging from $\alpha = 0.01$ to $\alpha = 0.99$. The VBA application works in conjunction with MATLAB's Excel Link and integration function int() to find the min expected value and min variance.

The VBA program uses interpolation to find the extreme interval values once the min and max cumulative distribution functions have been evaluated. The application searches for the largest cumulative distribution function value less than α , and the smallest cumulative distribution function value greater than α . The interpolation formula:

$$g_\alpha = g_1 + \frac{(g_2 - g_1)(\alpha - H(g_1))}{(H(g_2) - H(g_1))}$$
 with $g_1 < g_\alpha < g_2$ and $H(g_1) < \alpha < H(g_2)$ calculates the extreme interval value g_α (Law and Kelton 470).

Example: n = 25 Observations

Suppose $g = \min(x_1, \dots, x_{25})$ where x is uniformly distributed with parameters a = 0 and b = 1. The min probability density function is $h(g) = 25(1 - g)^{24}$, and the min cumulative distribution function is $H(g) = \int_0^g h(g)dg$. The min expected value is $\mu = 0.03846$, and the min standard deviation is $\sigma = 0.03701$. The min mode is 0.00000, the min median is 0.02735, and the min coefficient of variation is 0.96225.

Suppose $g = \max(x_1, \dots, x_{25})$ where x is uniformly distributed with parameters a = 0 and b = 1. The max probability density function is $h(g) = 25g^{24}$, and the max cumulative

distribution function is $H(g) = \int_c^g h(g)dg$. The max expected value is $\mu = 0.96154$, and the max standard deviation is $\sigma = 0.03701$. The max mode is 1.00000, the max median is 0.97265, and the max coefficient of variation is 0.03849.

The table lists some of the min and max extreme interval values for an observation size of $n = 25$. For example, when $P(g < g_{0.80}) = 0.80$, the min extreme interval value is $g_{0.80} = 0.06235$, and the max extreme interval value is $g_{0.80} = 0.99111$. The min extreme interval value $g_{0.50} = 0.02735$ is the min median. The max extreme interval value $g_{0.50} = 0.97265$ is the max median.

$P(g < g_{\alpha}) = \alpha$	Min g_{α}	Max g_{α}
0.01	0.00040	0.83176
0.05	0.00205	0.88707
0.10	0.00421	0.91201
0.20	0.00889	0.93765
0.30	0.01417	0.95298
0.40	0.02023	0.96401
0.50	0.02735	0.97265
0.60	0.03599	0.97977
0.70	0.04702	0.98583
0.80	0.06235	0.99111
0.90	0.08799	0.99579
0.95	0.11293	0.99795
0.99	0.16824	0.99960

Analysis

The min expected value, min median, min standard deviation, and min extreme interval values shift closer to zero; whereas, the min coefficient of variation shifts closer to one as the observation size increases. The min mode is zero and the min mode < min median < min expected value for all observation sizes greater than one.

The max expected value, max median, and max extreme interval values shift closer to one as the observation size increases. The max standard deviation and max coefficient of variation decrease as the observation size increases. The max expected value < max median < max mode and the max mode is one for all observation sizes larger than one.

The following tables list the min and max extreme interval values for $\alpha = 0.20$, $\alpha = 0.30$, and $\alpha = 0.40$ and observation sizes of $n = 1, 3, 5, 25, 50, 75, 100, 500$, and 1000. In addition, the min and max expected values, standard deviations, modes, medians, and coefficient of variations are displayed for these observation sizes. The min and max values are the same, and the mode

is not available when the observation size is $n = 1$. Also, the min and max standard deviations have the same values.

Min Extreme Interval Values and Statistics

n	$\alpha = 0.20$	$\alpha = 0.30$	$\alpha = 0.40$	μ	σ	Mode	Median	$\frac{\sigma}{\mu}$ CV = μ
1	0.20000	0.30000	0.40000	0.50000	0.28868	NA	0.50000	0.57735
3	0.07168	0.11210	0.15657	0.25000	0.19365	0.00000	0.20630	0.77460
5	0.04365	0.06885	0.09712	0.16667	0.14086	0.00000	0.12945	0.84515
25	0.00889	0.01417	0.02023	0.03846	0.03701	0.00000	0.02735	0.96225
50	0.00445	0.00711	0.01016	0.01961	0.01923	0.00000	0.01377	0.98058
75	0.00297	0.00474	0.00679	0.01316	0.01299	0.00000	0.00920	0.98693
100	0.00223	0.00356	0.00510	0.00990	0.00980	0.00000	0.00691	0.99015
500	0.00045	0.00071	0.00102	0.00200	0.00199	0.00000	0.00139	0.99801
1000	0.00022	0.00036	0.00051	0.00100	0.00100	0.00000	0.00069	0.99900

Max Extreme Interval Values and Statistics

n	$\alpha = 0.20$	$\alpha = 0.30$	$\alpha = 0.40$	μ	σ	Mode	Median	$\frac{\sigma}{\mu}$ CV = μ
1	0.20000	0.30000	0.40000	0.50000	0.28868	NA	0.50000	0.57735
3	0.58480	0.66943	0.73681	0.75000	0.19365	1.00000	0.79370	0.25820
5	0.72478	0.78600	0.83255	0.83333	0.14086	1.00000	0.87055	0.16903
25	0.93765	0.95298	0.96401	0.96154	0.03701	1.00000	0.97265	0.03849
50	0.96832	0.97621	0.98184	0.98039	0.01923	1.00000	0.98623	0.01961
75	0.97877	0.98408	0.98786	0.98684	0.01299	1.00000	0.99080	0.01316
100	0.98403	0.98803	0.99088	0.99010	0.00980	1.00000	0.99309	0.00990
500	0.99679	0.99759	0.99817	0.99800	0.00199	1.00000	0.99861	0.00200
1000	0.99839	0.99880	0.99908	0.99900	0.00100	1.00000	0.99931	0.00100

Other Parameter Values of a and b

The inverse-transform method for generating a uniformly distributed variable is $x = a + (b-a)u$ where u is from a uniform population with parameters $a = 0$ and $b = 1$ (Law and Kelton 460). This method can be used to find the min and max uniform extreme interval values for other values of a and b provided that $a < b$.

First, the extreme interval value g_α when $a = 0$ and $b = 1$ is found. Then, the extreme interval value for other values of a and b is $g'_\alpha = a + g_\alpha(b - a)$. The following example finds the min and max extreme interval values for $\alpha = 0.30$, $n = 25$, $a = 13$, and $b = 15$.

Let $g'_\alpha = \min(x_1, \dots, x_{25})$ where x is uniformly distributed with parameters $a = 13$ and $b = 15$. When $n = 25$, $\alpha = 0.30$, $a = 0$, and $b = 1$, the min extreme interval value is

$g_{0.30} = 0.01417$. Thus, the min extreme interval value is $g_{0.30}' = 13 + (15-13)(0.01417) = 13.02834$ when $n = 25$, $\alpha = 0.30$, $a = 13$, and $b = 15$.

The following shows that $g_{0.30}' = 13.02834$ is the min extreme interval value when $n = 25$, $\alpha = 0.30$, $a = 13$, and $b = 15$. The min probability density function is

$$h(g') = 25 \left(\frac{1}{(15 - 13)} \right) \left(1 - \frac{(g' - 13)}{(15 - 13)} \right)^{24} = 12.5 \left(1 - \frac{g' - 13}{2} \right)^{24} \quad \text{when } n = 25, a = 13, \text{ and } b = 15.$$

When $g' = 13.02834$, the min cumulative distribution function is

$$H(g') = \int_{13}^{13.02834} h(g') dg' \approx 0.30.$$

Now let $g' = \max(x_1, \dots, x_{25})$ where x is uniformly distributed with parameters $a = 13$ and $b = 15$. When $n = 25$, $\alpha = 0.30$, $a = 0$, and $b = 1$, the max extreme interval value is $g_{0.30} = 0.95298$. Thus, the max extreme interval value is $g_{0.30}' = 13 + (15-13)(0.95298) = 14.90596$ when $n = 25$, $\alpha = 0.30$, $a = 13$, and $b = 15$.

The following shows that $g_{0.30}' = 14.90596$ is the max extreme interval value when $n = 25$, $\alpha = 0.30$, $a = 13$, and $b = 15$. The max probability density function is

$$h(g') = 25 \left(\frac{1}{(15 - 13)} \right) \left(\frac{(g' - 13)}{(15 - 13)} \right)^{24} = 12.5 \left(\frac{g' - 13}{2} \right)^{24} \quad \text{when } n = 25, a = 13, \text{ and } b = 15.$$

When $g' = 14.90596$, the max cumulative distribution function is

$$H(g') = \int_{13}^{14.90596} h(g') dg' \approx 0.30.$$

Application

A system consists of three components running in parallel. The system will fail when the last component stops running. One wants to determine the probability that the system will still be running after 8.5 hours.

Find the probability $P(g' > 8.5)$ where $g' = \max(x_1, x_2, x_3)$ and x is uniformly distributed with parameters $a = 7.5$ hours and $b = 9$ hours. The probability $P(g' > 8.5)$ is equivalent to $P(g' < 8.5) = \alpha$. This probability $P(g' < 8.5) = \alpha$ is the same as the probability $P(g < g_\alpha) = \alpha$

when $n = 3$, $a = 0$, $b = 1$, and $g_\alpha = \frac{(g_\alpha - a)}{(b - a)} = \frac{(8.5 - 7.5)}{(9.0 - 7.5)} = 0.66667$.

Jance's max uniform table will be used to find the probability for $g_\alpha = 0.66667$, $n = 3$, $a = 0$, and $b = 1$. The closest entry in Jance's table gives a probability of $\alpha = 30\%$. Thus, there is a 70% chance that the system will still be running after 8.5 hours.

Conclusions

The min and max extreme interval values and statistics for uniformly distributed variables with parameters $a = 0$ and $b = 1$ are discussed. An example showing some min and max extreme interval values and statistics for an observation size of $n = 25$ is presented. In addition, an analysis of the min and max extreme interval values and statistics for different observation sizes is provided. Finally, a procedure for finding the extreme interval values for different parameter values of a and b (provided $a < b$), and an application of this research are presented.

References

Hines, William W., Douglas C. Montgomery, David M. Goldsman, and Connie M. Borrer. *Probability and Statistics in Engineering Fourth Edition*. John Wiley & Sons, Inc., 2003.

Jance, Marsha L. *Calculating Min and Max Extreme Interval Values for Various Distributions*. Doctoral dissertation. Chicago, IL: Illinois Institute of Technology Stuart School of Business, 2007.

Law, Averill M., and W. David Kelton. *Simulation Modeling and Analysis Third Edition*. McGraw-Hill, 2000.

MATLAB: Version 7.1, The MathWorks, 2005.

Microsoft Excel: Professional Edition 2003, Redmond, WA: Microsoft Corporation, 2003.