SCHOOL of GRADUATE STUDIES
EAST TENNESSEE STATE UNIVERSITY

East Tennessee State University
Digital Commons @ East Tennessee State University

Electronic Theses and Dissertations

5-2014

Are Highly Dispersed Variables More Extreme? The Case of Distributions with Compact Support

Benedict E. Adjogah
East Tennessee State University

Follow this and additional works at: http://dc.etsu.edu/etd

Recommended Citation
Adjogah, Benedict E., "Are Highly Dispersed Variables More Extreme? The Case of Distributions with Compact Support" (2014). *Electronic Theses and Dissertations.* Paper 2382. http://dc.etsu.edu/etd/2382

This Thesis - Open Access is brought to you for free and open access by Digital Commons @ East Tennessee State University. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of Digital Commons @ East Tennessee State University. For more information, please contact dcadmin@etsu.edu.

Are Highly Dispersed Variables More Extreme? The Case of Distributions with Compact Support

---

A thesis

presented to

the faculty of the Department of Mathematics and Statistics

East Tennessee State University

In partial fulfillment

of the requirements for the degree

Master of Science in Mathematical Sciences

---

by

Benedict Adjogah

May 2014

Anant Godbole, Ph.D., Chair

Jeff Knisley, Ph.D.

Edith Seier, Ph.D.

Robert Gardner, Ph.D.

Keywords: variance, expected value, cancellation in alternating sums

ABSTRACT

Are Highly Dispersed Variables More Extreme? The Case of Distributions with Compact Support

by

Benedict Adjogah

We consider discrete and continuous symmetric random variables X taking values in $[0,1]$, and thus, having expected value 1/2. The main thrust of this investigation is to study the correlation between the variance $\mathtt{Var(X)}$ of X and the value of the expected maximum $\mathbb{E}(M_n) = \mathbb{E}(X_1, \ldots, X_n)$ of n independent and identically distributed random variables $X_1, X_2 \ldots X_n$, each distributed as X. Many special cases are studied, some leading to very interesting alternating sums, and some progress is made towards a general theory.

Copyright by

Benedict Adjogah 2014

All Rights Reserved

ACKNOWLEDGMENTS

First of all I want to thank the Almighty God for giving us life and the grace to come this far. I would like to express my profound gratitude to my Advisor, Dr. Anant Godbole. Your time, advice on both research and career have been priceless. I would also like to thank my committee members, Dr. Jeff Knisley, Dr. Edith Seier and Dr. Robert Gardner for your time and support throughout this time. Words are not enough to express how thankful I am to you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

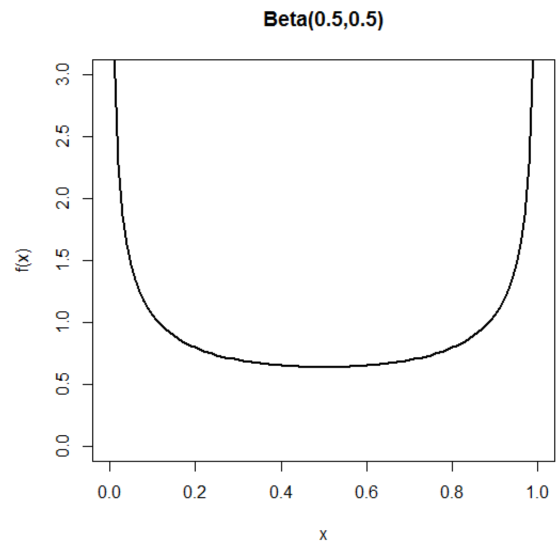
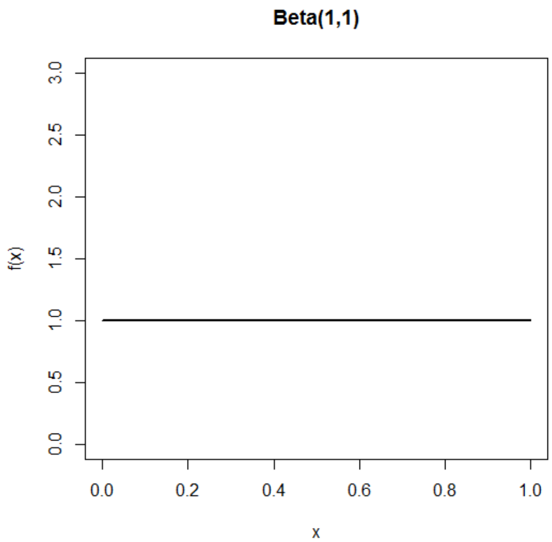


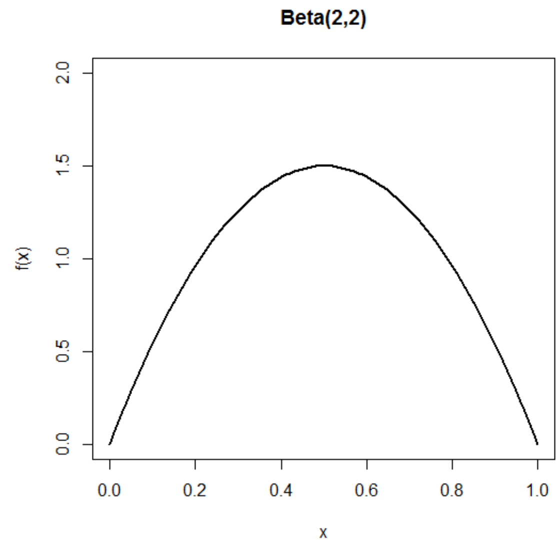
Figure 1: Distribution with `Var(X)=0.125`

We know that the variance `Var(X)` of a random variable is a measure of its spread, or dispersion. So if we have observations $X_1, \ldots, X_n$ from a distribution X on $[0, 1]$ with a high variance, then the high spread should make the maximum of the observations larger than for another distribution with a smaller variance. Consider the diagrams in Figures 1 and 2, where we get (hypothetical) expected maxima of 0.95, 0.8, and 0.7, respectively, for variables with `Var(X)` $= 0.125$; 0.0833, and 0.05.

We expect, from the figures, the largest of n values is further from the mean in Figure 1 than in Figure 2. In the same way, comparing Figure 2, we can see that the variance in our Beta(1,1) graph is higher than in our Beta(2,2) graph, and as such, this makes the Beta(1,1) graph have a higher value for its expected maximum than the Beta(2,2) graph. This thesis is intended to quantify such behavior.



Distribution with `Var(X)=0.083`



Distribution with `Var(X)=0.05`

Figure 2: Comparing Beta(1,1) and Beta(2,2) graphs

Consider a random variable X taking values in $[0, 1]$. We first show that such a random variable has bounded variance no more than $\frac{1}{4}$ and that this bound is attained. To prove the second part, first note that the discrete variable X that takes on values 0 and 1 with probability $\frac{1}{2}$ each has mean $\frac{1}{2}$ and variance given by

$$\begin{aligned}\texttt{Var(X)} &= \mathbb{E}(X^2) - \mathbb{E}^2(X) \\ &= \frac{1}{2} - \frac{1}{4} \\ &= \frac{1}{4}.\end{aligned}$$

For the first part we note that

$$\begin{aligned}\texttt{Var(X)} &= \mathbb{E}(X^2) - \mathbb{E}^2(X) \\ &\leq \mathbb{E}(X) - \mathbb{E}^2(X) \\ &= \mathbb{E}(X)(1 - \mathbb{E}(X)) \\ &\leq \frac{1}{4},\end{aligned}$$

where the last line follows from the fact that the function $\varphi(x) = x(1-x); 0 \leq x \leq 1$ takes on a maximum value of $\frac{1}{4}$ at $x = \frac{1}{2}$.

Throughout this thesis, we will consider discrete and continuous symmetric random variables X taking values in $[0, 1]$, and thus, having expected value $1/2$, and, by the above discussion, $\texttt{Var(X)} \leq \frac{1}{4}$. The main thrust of our investigation is to study the correlation between the variance $\texttt{Var(X)}$ of X and the value of the expected maximum $\mathbb{E}(M_n) = \mathbb{E}(\max(X_1, \ldots, X_n))$ of n independent and identically distributed random variables $X_1, X_2 \ldots X_n$, each distributed as X. In this chapter we consider some baseline examples, which will then be followed in later chapters by families of examples. Finally, we will give the beginnings of a general theory.

Example 1.1. *If X is uniformly distributed on $[0, 1]$, then $\texttt{Var(X)} = \frac{1}{12}$ and $\mathbb{E}(M_n) = 1 - \frac{1}{n+1}$.*

Proof. To see this we note that

$$\begin{aligned}\texttt{Var(X)} &= \mathbb{E}(X^2) - \mathbb{E}^2(X) \\ &\leq \left(\int_0^1 x^2 dx\right) - \frac{1}{4} \\ &= \frac{1}{12}.\end{aligned}$$

Next, note that $\{M_n \le x\}$ if and only if $\{X_i \le x\}$ for each $1 \le i \le n$. This is an important fact that will be used throughout the thesis. Thus, for any input variable X with CDF $F_X(x)$,

$$F_M(x) = F_{M_n}(x) = F_X^n(x) \tag{1}$$

and, in our case, this gives

$$F_M(x) = x^n \quad (0 \le x \le 1),$$

so that

$$\mathbb{E}(M) = \int_0^1 x f_M(x) dx = n \int_0^1 x^n dx = \frac{n}{n+1} = 1 - \frac{1}{n+1},$$

as asserted. This completes the proof. □

For reasons that will become clearer soon, we always will express the value of $\mathbb{E}(M_n)$ in the form $1 - \epsilon_n$, where $\epsilon_n \to 0$ as $n \to \infty$. Also, from now on, we will suppress the subscript n in expressions such as $\mathbb{E}(M_n)$ when there is no danger of confusion.

We now compare the variance and expected maximum of the distribution from Chapter 1.

Table 1: Comparing Variance with Expected Maximum

Example	`Var(x)`	$\mathbb{E}(M)$
1. U[0,1]	$\frac{1}{12}$	$1 - \frac{1}{n+1}$

We will be building on this example in the present chapter. Next, let us consider

$$X = \begin{cases} 0 \;\; \textit{with} \;\; \textit{probability} \;\; \frac{1}{2} \\ \\ 1 \;\; \textit{with} \;\; \textit{probability} \;\; \frac{1}{2} \end{cases}$$

Then

$$\mathbb{E}(\max(X_1, ..., X_n)) = 1 - \left(\frac{1}{2}\right)^n,$$

since $M = 0$ if and only if $X_i = 0$ for each i, and $M = 1$ otherwise. Thus,

$$\mathbb{E}(M) = \mathbb{P}(X_i \geq 1 \text{ for some } i) = 1 - \left(\frac{1}{2}\right)^n$$

.

Table 2 compares our current set of results on the variance and expected maximum given the examples above:

Table 2: Comparing Variance with Expected Maximum

Example	`Var(X)`	$\mathbb{E}(M)$
1. U[0,1]	$\frac{1}{12}$	$1-\frac{1}{n+1}$
2. $X=0$ and $X=1$	$\frac{1}{4}$	$1-(\frac{1}{2})^n$

We now proceed to the discrete uniform cases. First, let us consider the following cases of m, where $m+1$ equals the number of discrete values, uniformly spaced, that X takes.

Theorem 2.1. *This is the so-called $m=3$ case, where with $\mathbb{P}(X=x)=\frac{1}{4}$ where $x=0,\ \frac{1}{3},\ \frac{2}{3},\ 1$. In this case, we have* `Var(X)` $=$ `0.139` *and* $\mathbb{E}(M_n)=1-\frac{1}{3}\cdot\left(\frac{3}{4}\right)^n$

Proof. Now we have that

$$\begin{aligned}
\mathbb{E}(X) &= \sum_x x\mathbb{P}(x) \\
&= 0\left(\frac{1}{4}\right)+\frac{1}{3}\left(\frac{1}{4}\right)+\frac{2}{3}\left(\frac{1}{4}\right)+\left(\frac{1}{4}\right) \\
&= \frac{1}{12}+\frac{1}{6}+\frac{1}{4} \\
&= \frac{1}{2}
\end{aligned}$$

and thus,

$$\begin{aligned}
\texttt{Var(X)} &= \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 \\
&= \sum_x x^2 \mathbb{P}(x) - \frac{1}{4} \\
&= 0\left(\frac{1}{4}\right) + \frac{1}{9}\left(\frac{1}{4}\right) + \left(\frac{4}{9}\right)\frac{1}{4} + \frac{1}{4} = \frac{1}{36} + \frac{1}{9} + \frac{1}{4} - \frac{1}{4} \\
&= \frac{14}{36} - \frac{1}{4} \\
&= 0.139.
\end{aligned}$$

Thus,

$$\begin{aligned}
\mathbb{P}(X \le 0) = \frac{1}{4} &\Rightarrow \mathbb{P}(M = 0) = \left(\frac{1}{4}\right)^n \\
\mathbb{P}\left(X \le \frac{1}{3}\right) = \frac{1}{2} \Rightarrow \mathbb{P}\left(M \le \frac{1}{3}\right) = \left(\frac{1}{2}\right)^n &\Rightarrow \mathbb{P}\left(M = \frac{1}{3}\right) = \left(\frac{1}{2}\right)^n - \left(\frac{1}{4}\right)^n \\
\mathbb{P}\left(X \le \frac{2}{3}\right) = \frac{3}{4} \Rightarrow \mathbb{P}\left(M \le \frac{2}{3}\right) = \left(\frac{3}{4}\right)^n &\Rightarrow \mathbb{P}\left(M = \frac{2}{3}\right) = \left(\frac{3}{4}\right)^n - \left(\frac{1}{2}\right)^n \\
\mathbb{P}\left(X \le 1\right) = 1 \Rightarrow \mathbb{P}(M \le 1) = 1 &\Rightarrow \mathbb{P}(M = 1) = 1 - \left(\frac{3}{4}\right)^n
\end{aligned}$$

We summarize the probabilities of the maximum values with a table:

Table 3: Probabilities of Maximum Values

x	$\mathbb{P}(x)$
0	$(\frac{1}{4})^n$
$\frac{1}{3}$	$(\frac{1}{2})^n - (\frac{1}{4})^n$
$\frac{2}{3}$	$(\frac{3}{4})^n - (\frac{1}{2})^n$
1	$1 - (\frac{3}{4})^n$

It follows that

$$
\begin{aligned}
\mathbb{E}(M) &= \sum_m m\mathbb{P}(m) \\
&= 0\left(\frac{1}{4}\right)^n + \frac{1}{3}\left[\left(\frac{1}{2}\right)^n - \left(\frac{1}{4}\right)^n\right] + \frac{2}{3}\left[\left(\frac{3}{4}\right)^n - \left(\frac{1}{2}\right)^n\right] + 1 - \left(\frac{3}{4}\right)^n \\
&= 1 - \frac{1}{3}\left(\frac{3}{4}\right)^n\left[1 + \left(\frac{2}{3}\right)^n + \left(\frac{1}{3}\right)^n\right] \\
&\approx 1 - \frac{1}{3}\left(\frac{3}{4}\right)^n.
\end{aligned}
$$

□

We will now update Table 2 by adding the results of Theorem 2.1 to get Table 4. Since each family of distribution is different, it is reasonable to compare the variance of X with its expected maximum for each of the distributions according to the family they belong to. Since we have different distributions in each family, we cannot clearly see the relationship. This will become more evident as we progress.

Table 4: Comparing Variance with Expected Maximum

Example	`Var(X)`	$\mathbb{E}(M)$
1. U[0,1]	$\frac{1}{12}$	$1 - \frac{1}{n+1}$
2. $X = 0$ and $X = 1$	$\frac{1}{4}$	$1 - (\frac{1}{2})^n$
3. $X = 0, \frac{1}{3}, \frac{2}{3}, 1$ with $\mathbb{P}(x) = \frac{1}{4}$	0.139	$1 - \frac{1}{3}.\left(\frac{3}{4}\right)^n$

Theorem 2.2. *For the $m = 5$ case, with $\mathbb{P}(X = x) = \frac{1}{6}$ where $x = 0, 0.2, 0.4, 0.6, 0.8, 1,$ we have that $\mathbb{E}(X) = 0.5$, $\texttt{Var(X)} = \texttt{0.1166667}$ and $\mathbb{E}(M_n) = 1 - (0.2)\left(\frac{5}{6}\right)^n$*

Table 5: The Distribution of X where $m = 5$

x	$P(x)$
0	$\frac{1}{6}$
0.2	$\frac{1}{6}$
0.4	$\frac{1}{6}$
0.6	$\frac{1}{6}$
0.8	$\frac{1}{6}$
1	$\frac{1}{6}$

Proof. We now show that the expected value of the distribution is 0.5 and then we can proceed to find the variance and probability of the maximum which, in turn, will be used to calculate the expected maximum value.

$$\begin{aligned}
\mathbb{E}(X) &= \sum_x x\mathbb{P}(x) \\
&= 0\left(\frac{1}{6}\right) + 0.2\left(\frac{1}{6}\right) + 0.4\left(\frac{1}{6}\right) + 0.6\left(\frac{1}{6}\right) + 0.8\left(\frac{1}{6}\right) + \left(\frac{1}{6}\right) \\
&= 0.5
\end{aligned}$$

and thus,

$$\begin{aligned}
\texttt{Var(X)} &= \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 \\
&= \sum_x x^2 \mathbb{P}(x) - \frac{1}{4} \\
&= 0\left(\frac{1}{6}\right) + (0.2)^2\left(\frac{1}{6}\right) + (0.4)^2\left(\frac{1}{6}\right) + (0.6)^2\left(\frac{1}{6}\right) + (0.8)^2\left(\frac{1}{6}\right) + \frac{1}{6} - \frac{1}{4} \\
&= \frac{11}{30} - \frac{1}{4} \\
&= 0.1166667.
\end{aligned}$$

From the above it follows that

$$\begin{aligned}
\mathbb{P}(X \le 0) = \frac{1}{6} \Rightarrow \mathbb{P}(M = 0) &= \left(\frac{1}{6}\right)^n \\
\mathbb{P}(X \le 0.2) = \frac{1}{3} \Rightarrow \mathbb{P}(M \le 0.2) = \left(\frac{1}{3}\right)^n \Rightarrow \quad \mathbb{P}(M = 0.2) &= \left(\frac{1}{3}\right)^n - \left(\frac{1}{6}\right)^n \\
\mathbb{P}(X \le 0.4) = \frac{1}{2} \Rightarrow \mathbb{P}(M \le 0.4) = \left(\frac{1}{2}\right)^n \Rightarrow \quad \mathbb{P}(M = 0.4) &= \left(\frac{1}{2}\right)^n - \left(\frac{1}{3}\right)^n \\
\mathbb{P}(X \le 0.6) = \frac{4}{6} = \frac{2}{3} \Rightarrow \mathbb{P}(M \le 0.6) = \left(\frac{2}{3}\right)^n \Rightarrow \quad \mathbb{P}(M = 0.6) &= \left(\frac{2}{3}\right)^n - \left(\frac{1}{2}\right)^n \\
\mathbb{P}(X \le 0.8) = \frac{5}{6} \Rightarrow \mathbb{P}(M \le 0.8) = \left(\frac{5}{6}\right)^n \Rightarrow \quad \mathbb{P}(M = 0.8) &= \left(\frac{5}{6}\right)^n - \left(\frac{2}{3}\right)^n \\
\mathbb{P}(X \le 1) = 1 \Rightarrow \mathbb{P}(M \le 1) = 1 \Rightarrow \mathbb{P}(M = 1) &= 1 - \left(\frac{5}{6}\right)^n
\end{aligned}$$

Table 6 shows the probability of the maximum equalling various values:

Table 6: Probability of Maximum Values

x	$\mathbb{P}(x)$
0	$\left(\frac{1}{6}\right)^n$
0.2	$\left(\frac{1}{3}\right)^n - \left(\frac{1}{6}\right)^n$
0.4	$\left(\frac{1}{2}\right)^n - \left(\frac{1}{3}\right)^n$
0.6	$\left(\frac{2}{3}\right)^n - \left(\frac{1}{2}\right)^n$
0.8	$\left(\frac{5}{6}\right)^n - \left(\frac{2}{3}\right)^n$
1	$1 - \left(\frac{5}{6}\right)^n$

Now that we have the probability of the maximum values, we can find the expected value of the maximum:

$$\begin{aligned}
\mathbb{E}(M) &= \sum_m m\mathbb{P}(m) \\
&= \sum_{j=0}^{5}(0.2j)\left[\left(\frac{j+1}{6}\right)^n - \left(\frac{j}{6}\right)^n\right] \\
&= 1-(0.2)\left(\frac{5}{6}\right)^n - (0.2)\left(\frac{2}{3}\right)^n - (0.2)\left(\frac{1}{2}\right)^n - (0.2)\left(\frac{1}{3}\right)^n - (0.2)\left(\frac{1}{6}\right)^n \\
&= 1-(0.2)\left[\left(\frac{5}{6}\right)^n + \left(\frac{2}{3}\right)^n + \left(\frac{1}{2}\right)^n + \left(\frac{1}{3}\right)^n + \left(\frac{1}{6}\right)^n\right] \\
&= 1-(0.2)\left(\frac{5}{6}\right)^n\left[1 + \left(\frac{4}{5}\right)^n + \left(\frac{3}{5}\right)^n + \left(\frac{2}{5}\right)^n + \left(\frac{1}{5}\right)^n\right] \\
&\approx 1-(0.2)\left(\frac{5}{6}\right)^n
\end{aligned}$$

□

Table 7 is an update of Table 4 from the results of Theorem 2.2

Theorem 2.3. *For the $m = 10$ case, with $\mathbb{P}(X = x) = \frac{1}{11}$ uniformly where*

Table 7: Comparing Variance with Expected Maximum

Example	Var(X)	$\mathbb{E}(M)$
1. U[0,1]	$\frac{1}{12}$	$1-\frac{1}{n+1}$
2. $X=0$ and $X=1$	$\frac{1}{4}$	$1-(\frac{1}{2})^n$
3. $m=3$ with $\mathbb{P}(x)=\frac{1}{4}$ where $x=0,\frac{1}{3},\frac{2}{3},1$	0.139	$1-\frac{1}{3}\cdot\left(\frac{3}{4}\right)^n$
4. $m=5$ with $\mathbb{P}(x)=\frac{1}{6}$ where $x=0,0.2,0.4,...,1$	0.1166667	$1-(0.2)\left(\frac{5}{6}\right)^n$

$x=0,0.1,0.2,...,1,$ *we have that* $\mathbb{E}(X)=0.5$, $\texttt{Var(X)}=0.1082$ *and* $\mathbb{E}(M_n)=1-(0.1)\left(\frac{10}{11}\right)^n$

Proof. As we did in the previous case, we show that the expected value is 0.5 and then we proceed to find the variance and probability of the maximum values which, in turn, will be used to calculate the expected maximum value. We have

$$\mathbb{E}(X)=\sum_x x\mathbb{P}(x)=\sum_{j=0}^{10}(0.1j)\left(\frac{1}{11}\right)=0.5.$$

For the variance, we have that

$$\begin{aligned}\texttt{Var(X)} &= \mathbb{E}(X^2)-[\mathbb{E}(X)]^2=\sum_x x^2\mathbb{P}(x)-\frac{1}{4}\\ &= \sum_{j=0}^{10}(0.1j)^2\left(\frac{1}{11}\right)-\left(\frac{1}{4}\right)\\ &= 0.358182-\left(\frac{1}{4}\right)\\ &= 0.1082.\end{aligned}$$

Next, we calculate the probability of the maximum values. We have

$$
\begin{aligned}
\mathbb{P}(X \le 0) &= \frac{1}{11} \Rightarrow \mathbb{P}(M = 0) = \left(\frac{1}{11}\right)^n \\
\mathbb{P}(X \le 0.1) &= \frac{2}{11} \Rightarrow \mathbb{P}(M \le 0.1) = \left(\frac{2}{11}\right)^n \Rightarrow \quad \mathbb{P}(M = 0.1) = \left(\frac{2}{11}\right)^n - \left(\frac{1}{11}\right)^n \\
\mathbb{P}(X \le 0.2) &= \frac{3}{11} \Rightarrow \mathbb{P}(M \le 0.2) = \left(\frac{3}{11}\right)^n \Rightarrow \mathbb{P}(M = 0.2) = \left(\frac{3}{11}\right)^n - \left(\frac{2}{11}\right)^n \\
\mathbb{P}(X \le 0.3) &= \frac{4}{11} \Rightarrow \mathbb{P}(M \le 0.3) = \left(\frac{4}{11}\right)^n \Rightarrow \mathbb{P}(M = 0.3) = \left(\frac{4}{11}\right)^n - \left(\frac{3}{11}\right)^n \\
\mathbb{P}\left(X \le 0.4\right) &= \frac{5}{11} \Rightarrow \mathbb{P}(M \le 0.4) = \left(\frac{5}{11}\right)^n \Rightarrow \mathbb{P}(M = 0.4) = \left(\frac{5}{11}\right)^n - \left(\frac{4}{11}\right)^n \\
\mathbb{P}(X \le 0.5) &= \frac{6}{11} \Rightarrow \mathbb{P}(M \le 0.5) = \left(\frac{6}{11}\right)^n \Rightarrow \mathbb{P}(M = 0.5) = \left(\frac{6}{11}\right)^n - \left(\frac{5}{11}\right)^n \\
\mathbb{P}(X \le 0.6) &= \frac{7}{11} \Rightarrow \mathbb{P}(M \le 0.6) = \left(\frac{7}{11}\right)^n \Rightarrow \mathbb{P}(M = 0.6) = \left(\frac{7}{11}\right)^n - \left(\frac{6}{11}\right)^n \\
\mathbb{P}(X \le 0.7) &= \mathbb{P}(M \le 0.7) = \left(\frac{8}{11}\right)^n \Rightarrow \mathbb{P}(M = 0.7) = \left(\frac{8}{11}\right)^n - \left(\frac{7}{11}\right)^n \\
\mathbb{P}(X \le 0.8) &= \frac{9}{11} \Rightarrow \mathbb{P}(M \le 0.8) = \left(\frac{9}{11}\right)^n \Rightarrow \mathbb{P}(M = 0.8) = \left(\frac{9}{11}\right)^n - \left(\frac{8}{11}\right)^n \\
\mathbb{P}(X \le 0.9) &= \frac{10}{11} \Rightarrow \mathbb{P}(M \le 0.9) = \left(\frac{10}{11}\right)^n \Rightarrow \mathbb{P}(M = 0.9) = \left(\frac{10}{11}\right)^n - \left(\frac{9}{11}\right)^n \\
\mathbb{P}(X \le 1) &= 1 \Rightarrow \mathbb{P}(M \le 1) = (1)^n \Rightarrow \mathbb{P}(M = 1) = 1 - \left(\frac{10}{11}\right)^n
\end{aligned}
$$

Table 8 shows the probability of the maximum equalling various values:

Table 8: Probability of Maximum values

x	$\mathbb{P}(x)$
0	$\left(\frac{1}{11}\right)^n$
0.1	$\left(\frac{2}{11}\right)^n - \left(\frac{1}{11}\right)^n$
0.2	$\left(\frac{3}{11}\right)^n - \left(\frac{2}{11}\right)^n$
0.3	$\left(\frac{4}{11}\right)^n - \left(\frac{3}{11}\right)^n$
0.4	$\left(\frac{5}{11}\right)^n - \left(\frac{4}{11}\right)^n$
0.5	$\left(\frac{6}{11}\right)^n - \left(\frac{5}{11}\right)^n$
0.6	$\left(\frac{7}{11}\right)^n - \left(\frac{6}{11}\right)^n$
0.7	$\left(\frac{8}{11}\right)^n - \left(\frac{7}{11}\right)^n$
0.8	$\left(\frac{9}{11}\right)^n - \left(\frac{8}{11}\right)^n$
0.9	$\left(\frac{10}{11}\right)^n - \left(\frac{9}{11}\right)^n$
1	$1 - \left(\frac{10}{11}\right)^n$

It follows that

$$\begin{aligned}
\mathbb{E}(M) &= \sum_m m\mathbb{P}(m) \\
&= \sum_{j=0}^{10} (0.1j)\left[\left(\frac{j+1}{11}\right)^n - \left(\frac{j}{11}\right)^n\right] \\
&= 0\left(\frac{1}{11}\right)^n + 0.1\left[\left(\frac{2}{11}\right)^n - \left(\frac{1}{11}\right)^n\right] + 0.2\left[\left(\frac{3}{11}\right)^n - \left(\frac{2}{11}\right)^n\right] + ... + 1 - \left(\frac{10}{11}\right)^n \\
&= 1 - (0.1)\left(\frac{10}{11}\right)^n - (0.1)\left(\frac{9}{11}\right)^n - (0.1)\left(\frac{8}{11}\right)^n - ... - (0.1)\left(\frac{1}{11}\right)^n \\
&\approx 1 - (0.1)\left(\frac{10}{11}\right)^n,
\end{aligned}$$

as asserted. □

We will now update Table 7 by adding the results of Theorem 2.3 to get Table 9.

Table 9: Comparing Variance with Expected Maximum

Example	Var(X)	$\mathbb{E}(M)$
1. U[0,1]	$\frac{1}{12}$	$1-\frac{1}{n+1}$
2. $X=0$ and $X=1$	$\frac{1}{4}$	$1-\left(\frac{1}{2}\right)^n$
3. $m=3$ with $\mathbb{P}(x)=\frac{1}{4}$ where $X=0,\frac{1}{3},\frac{2}{3},1$	0.139	$1-\frac{1}{3}.\left(\frac{3}{4}\right)^n$
4. $m=5$ with $\mathbb{P}(x)=\frac{1}{6}$ where $X=0,0.2,0.4,...,1$	0.1166667	$1-(0.2)\left(\frac{5}{6}\right)^n$
5. $m=10$ with $\mathbb{P}(x)=\frac{1}{11}$ where $X=0,0.1,0.2,...,1$	0.1082	$1-(0.1)\left(\frac{10}{11}\right)^n$

We next consider a specific continuous uniform-like case:

Theorem 2.4. *If*

$$f(x)=\begin{cases}\frac{3}{2} & if \;\; 0<x<\frac{1}{3}\\ 0 & if \;\; \frac{1}{3}<x<\frac{2}{3}\\ \frac{3}{2} & if \;\; \frac{2}{3}<x<1\end{cases}$$

then $\texttt{Var(X)} = \texttt{.1204}$ *and* $\mathbb{E}(M)\approx 1-\frac{2}{3n+3}$.

Proof. The expected value is calculated as shown below:

$$\begin{aligned}\mathbb{E}(X) &= \int_0^1 xf(x)dx\\ &= \int_0^{\frac{1}{3}}\left(\frac{3}{2}\right)xdx+\int_{\frac{1}{3}}^{\frac{2}{3}}0xdx+\int_{\frac{2}{3}}^{1}\left(\frac{3}{2}\right)xdx\\ &= \left[\frac{3}{4}x^2\right]_0^{\frac{1}{3}}+\left[\frac{3}{4}x^2\right]_{\frac{2}{3}}^{1}\\ &= \frac{1}{12}+\frac{3}{4}-\frac{3}{4}.\frac{4}{9}\\ &= \frac{1}{2}.\end{aligned}$$

Likewise, we calculate the variance of X as shown below:

$$\begin{aligned}
\texttt{Var(X)} &= \mathbb{E}(x-\mu)^2 \\
&= \int_0^1 \left(x-\frac{1}{2}\right)^2 f(x)dx \\
&= \int_0^{\frac{1}{3}} \left(x-\frac{1}{2}\right)^2 .\frac{3}{2}dx + \int_{\frac{1}{3}}^{\frac{2}{3}} \left(x-\frac{1}{2}\right)^2 .0dx + \int_{\frac{2}{3}}^{1} \left(x-\frac{1}{2}\right)^2 .\frac{3}{2}dx \\
&= \frac{13}{108} \\
&= 0.1204.
\end{aligned}$$

In other words, to find the expected maximum value, we need to find the cumulative distribution function of the maximum. This will then be used to calculate the density function of the maximum which will, in turn, be used to calculate the expected maximum function from the definition of expected value of continuous functions. The CDF of X is obtained by conducting a case study:

$$\begin{aligned}
F(x) &= \int_0^x f(t)dt = \int_0^x \frac{3}{2}dt = \frac{3}{2}x \qquad \left(0 < x < \frac{1}{3}\right) \\
&= \int_0^{\frac{1}{3}} f(x)dx + \int_{\frac{1}{3}}^x f(t)dt \\
&= \int_0^{\frac{1}{3}} \frac{3}{2}dt + \int_{\frac{1}{3}}^x dt = \left[\frac{3}{2}x\right]_0^{\frac{1}{3}} = \frac{1}{2} \qquad \left(\frac{1}{3} < x < \frac{2}{3}\right) \\
&= \int_0^{\frac{1}{3}} f(t)dt + \int_{\frac{1}{3}}^{\frac{2}{3}} f(t)dt + \int_{\frac{2}{3}}^x f(t)dt \\
&= \frac{3}{2}\left(\frac{1}{3}\right) + \frac{3}{2}x - \frac{3}{2}\left(\frac{2}{3}\right) = \left(\frac{3x-1}{2}\right) \qquad \left(\frac{2}{3} < x < 1.\right)
\end{aligned}$$

We proceed to find the cdf of the maximum:

$$
\begin{aligned}
F_M(x) = \left(\frac{3x}{2}\right)^n &\Rightarrow f_M(x) = \frac{3}{2}n\left(\frac{3x}{2}\right)^{n-1} \qquad \left(0 < x < \frac{1}{3}\right) \\
F_M(x) = \left(\frac{1}{2}\right)^n &\Rightarrow f_M(x) = 0 \qquad \left(\frac{1}{3} < x < \frac{2}{3}\right) \\
F_M(x) = \left(\frac{3x-1}{2}\right)^n &\Rightarrow f_{M_3}(x) = \frac{3}{2}n\left(\frac{3x-1}{2}\right)^{n-1} \qquad \left(\frac{2}{3} < x < 1.\right)
\end{aligned}
$$

Therefore, the density function of the maximum is

$$
f_{M_n}(x) = \begin{cases} \frac{3}{2}n\left(\frac{3x}{2}\right)^{n-1} & if \;\; 0 < x < \frac{1}{3} \\ 0 & if \;\; \frac{1}{3} < x < \frac{2}{3} \\ \frac{3}{2}n\left(\frac{3x-1}{2}\right)^{n-1} & if \;\; \frac{2}{3} < x < 1 \end{cases}
$$

The expected value of the maximum is obtained next:

$$
\mathbb{E}(M) = \int_0^1 x f_M(x)dx \tag{2}
$$

The integral is separated into three parts, I_1, I_2, I_3, where

$$
\begin{aligned}
I_1 &= \int_0^{\frac{1}{3}} x\left[\frac{3n}{2}\left(\frac{3x}{2}\right)^{n-1}\right]dx \\
&= \frac{3n}{2}\left(\frac{3}{2}\right)^{n-1}\int_0^{\frac{1}{3}} x^n dx \\
&= \frac{1}{3(2)^n} - \frac{1}{3(2)^n(n+1)}; \\
I_2 &= 0
\end{aligned}
$$

and

$$\begin{aligned}
I_3 &= \int_{\frac{2}{3}}^{1} x.\frac{3n}{2}\left(\frac{3x-1}{2}\right)^{n-1} dx \\
&= \frac{3n}{2} \cdot \int_{\frac{1}{2}}^{1} \left(\frac{2u+1}{3}\right) u^{n-1}.\frac{2}{3}du \\
&= n\int_{\frac{1}{2}}^{1} \left(\frac{2}{3}u^n + \frac{u^{n-1}}{3}\right) du \\
&= \frac{n}{3}\left[2\frac{u^{n+1}}{n+1} + \frac{u^n}{n}\right]_{\frac{1}{2}}^{1} \\
&= \frac{n}{3}\left[\frac{3n+1}{n(n+1)} - \left(\frac{1}{2}\right)^n \left(\frac{2n+1}{n(n+1)}\right)\right].
\end{aligned}$$

Therefore, we have:

$$\begin{aligned}
I_1 + I_2 + I_3 &= \frac{n}{3(n+1)}.\frac{1}{2^n} + \frac{n}{3}\left(\frac{3n+1}{n(n+1)}\right) - \frac{1}{3}\left(\frac{1}{2^n}\right)\left(\frac{2n+1}{n+1}\right) \\
&= \frac{1}{3.(2)^n}\left(\frac{n}{n+1} - \frac{2n+1}{n+1}\right) + \frac{n}{3}\left(\frac{3n+1}{n(n+1)}\right) \\
&= \frac{1}{3(2)^n}\left(-\frac{(n+1)}{n+1}\right) + \frac{n}{3}\left(\frac{3n+1}{n(n+1)}\right) \\
&= -\frac{1}{3(2)^n} + \frac{n}{3}\left(\frac{3n+1}{n(n+1)}\right) \\
&= \frac{3n+1}{3(n+1)} - \frac{1}{3(2)^n} \\
&= \frac{3n+3-2}{3n+3} - \frac{1}{3(2)^n} \\
&= 1 - \frac{2}{3n+3} - \frac{1}{3(2)^n} \\
&\approx 1 - \frac{2}{3n+3}.
\end{aligned}$$

Therefore,

$$\mathbb{E}(M) \approx 1 - \frac{2}{3n+3},$$

which completes the proof. □

We start with the case of general m.

Theorem 3.1. *If X has the distribution $f(x) = \frac{1}{m+1}; x = \frac{1}{m}, \frac{2}{m} \dots, 1$, then* $\mathtt{Var(X)} = \frac{\mathtt{1}}{\mathtt{12}} + \frac{\mathtt{1}}{\mathtt{6m}}$ *and* $\mathbb{E}(M) = 1 - \frac{1}{m}\left(\frac{m}{m+1}\right)^n$.

We begin the proof by showing that t$\mathbb{E}(X) = 0.5$ and then we can proceed to find the variance and the probability of the maximum values. This will then be used to calculate the expected maximum value.

Proof. We have

$$\begin{aligned}
\mathbb{E}(X) &= \sum_x x\mathbb{P}(x) \\
&= \frac{1}{m}\left(\frac{1}{m+1}\right) + \frac{2}{m}\left(\frac{1}{m+1}\right) + ... + \left(\frac{1}{m+1}\right) \\
&= \frac{1}{m(m+1)}\left(1+2+3+4+...+m\right) \\
&= \frac{1}{m(m+1)}\left(\frac{m(m+1)}{2}\right) \\
&= 0.5,
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E}(X^2) &= \sum_x x^2\mathbb{P}(x) \\
&= \frac{1}{m^2}\left(\frac{1}{m+1}\right)\left(1+2^2+3^2+4^2+...+m^2\right) \\
&= \frac{1}{m^2}\left(\frac{1}{m+1}\right)\left(\frac{m(m+1)(2m+1)}{6}\right) \\
&= \frac{1}{m}\left(\frac{2m+1}{6}\right).
\end{aligned}$$

Therefore,

$$\begin{aligned}\texttt{Var(X)} &= \frac{1}{m}\left(\frac{2m+1}{6}\right) - (0.5)^2 \\ &= \frac{1}{12} + \frac{1}{6m}.\end{aligned}$$

We now find the probabilities of the maximum:

$$\mathbb{P}(X=0) = \frac{1}{m+1} \Rightarrow \mathbb{P}(M=0) = \left(\frac{1}{m+1}\right)^n$$

$$\mathbb{P}\left(X \leq \frac{1}{m}\right) = \frac{2}{m+1} \Rightarrow$$
$$\mathbb{P}\left(M \leq \frac{1}{m}\right) = \left(\frac{2}{m+1}\right)^n \Rightarrow \mathbb{P}\left(M = \frac{1}{m}\right) = \left(\frac{2}{m+1}\right)^n - \left(\frac{1}{m+1}\right)^n$$
$$\mathbb{P}\left(X \leq \frac{2}{m}\right) = \frac{3}{m+1} \Rightarrow$$
$$\mathbb{P}\left(M \leq \frac{2}{m}\right) = \left(\frac{3}{m+1}\right)^n \Rightarrow \mathbb{P}\left(M = \frac{2}{m}\right) = \left(\frac{3}{m+1}\right)^n - \left(\frac{2}{m+1}\right)^n$$
$$\mathbb{P}\left(X \leq \frac{3}{m}\right) = \frac{4}{m+1} \Rightarrow$$
$$\mathbb{P}\left(M \leq \frac{3}{m}\right) = \left(\frac{4}{m+1}\right)^n \Rightarrow \mathbb{P}\left(M = \frac{3}{m}\right) = \left(\frac{4}{m+1}\right)^n - \left(\frac{3}{m+1}\right)^n.$$

.

.

$$\mathbb{P}(M \leq 1) = 1 \Rightarrow \mathbb{P}(M=1) = 1 - \left(\frac{m}{m+1}\right)^n$$

Table 10 shows the probability of the maximum values.

Table 10: Probability of the Maximum Values

x	$P(x)$
0	$\left(\frac{1}{m+1}\right)^n$
$\frac{1}{m}$	$\left(\frac{2}{m+1}\right)^n - \left(\frac{1}{m+1}\right)^n$
$\frac{2}{m}$	$\left(\frac{3}{m+1}\right)^n - \left(\frac{2}{m+1}\right)^n$
$\frac{3}{m}$	$\left(\frac{4}{m+1}\right)^n - \left(\frac{3}{m+1}\right)^n$
$\frac{4}{m}$	$\left(\frac{5}{m+1}\right)^n - \left(\frac{4}{m+1}\right)^n$
.	.
.	.
.	.
1	$1 - \left(\frac{m}{m+1}\right)^n$

Now that we have the probability of the maximum values, we can calculate the general form of the expected maximum of distributions of this form.

We thus get

$$\begin{aligned}
\mathbb{E}(M) &= \sum_x x\mathbb{P}(x) \\
&= \frac{1}{m}\left[\left(\frac{2}{m+1}\right)^n - \left(\frac{1}{m+1}\right)^n\right] + \ldots + 1 - (\frac{m}{m+1})^n \\
&= 1 - (\frac{m}{m+1})^n + \frac{1}{m}\left[\left(\frac{2}{m+1}\right)^n - \left(\frac{1}{m+1}\right)^n\right] + \ldots + \\
&\quad \frac{m-1}{m}\left[\left(\frac{m}{m+1}\right)^n - \left(\frac{m-1}{m+1}\right)^n\right] \\
&\approx 1 - \left(\frac{1}{m}\right)\left(\frac{m}{m+1}\right)^n
\end{aligned}$$

□

Tables 11, 12 and 13 compares the Variance and the Expected value of the maximum for different values of n using the General m Distribution.

Case 1: $n = 10$. Here, we can see from the plot in Figure 3 that there is a strong

Table 11: Comparing Variance with Expected Maximum for n=10

m	Var(X)	$\mathbb{E}(M)$
1	0.25	0.9990
2	0.1667	0.9913
3	0.1389	0.9812
4	0.125	0.9731
5	0.1167	0.9677
6	0.1111	0.9643
7	0.1071	0.9624
8	0.1042	0.9615
9	0.1019	0.9613
10	0.1	0.9614

positive correlation between the variance of X and its expected maximum. However, the relationship does not appear to be linear.

Case 2: $n = 20$

Table 12: Comparing Variance with Expected Maximum for n=20

m	Var(X)	$\mathbb{E}(M)$
1	0.25	0.9999
2	0.1667	0.9998
3	0.1389	0.9989
4	0.125	0.9971
5	0.1167	0.9948
6	0.1111	0.9924
7	0.1071	0.9901
8	0.1042	0.9881
9	0.1019	0.9865
10	0.1	0.9851

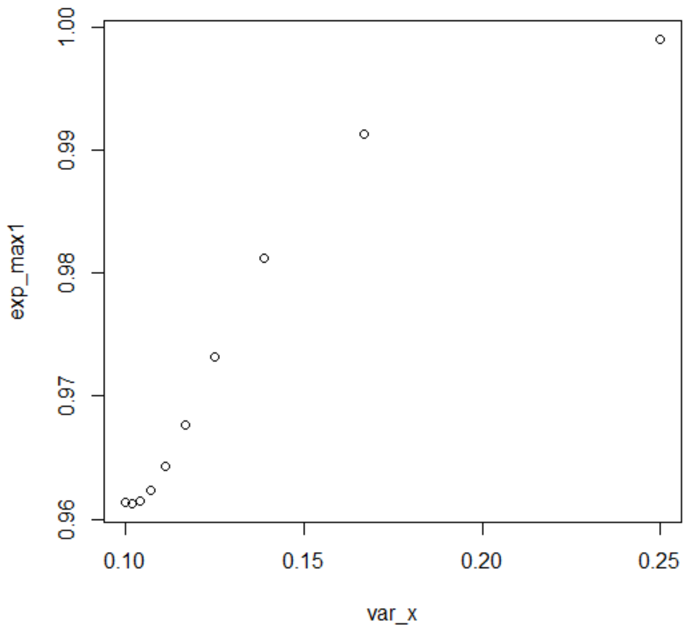


Figure 3: The Correlation between Variance and Expected Maximum

In this case, looking at Figure 4, we still have a strong relationship between the variance of X and its expected maximum. However, as with $n = 10$, the relationship is not linear.

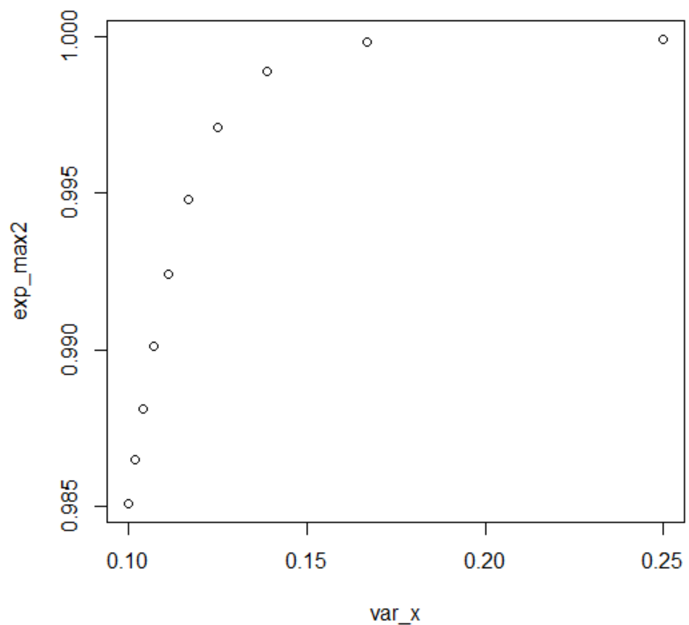


Figure 4: The Correlation between Variance and Expected Maximum

Table 13: Comparing Variance with Expected Maximum for n=30

m	Var(X)	$\mathbb{E}(M)$
1	0.25	9999
2	0.1667	0.9999
3	0.1389	0.9999
4	0.125	0.9997
5	0.1167	0.9991
6	0.1111	0.9984
7	0.1071	0.9974
8	0.1042	0.9963
9	0.1019	0.9953
10	0.1	0.9943

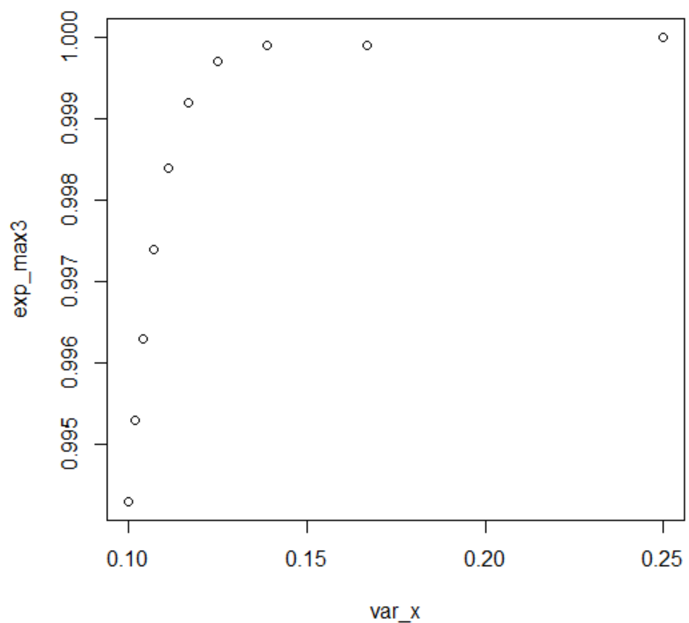


Figure 5: The Correlation between Variance and Expected Maximum

Case 3: $n = 30$. From Figure 5 above, we can see that there is still a strong positive correlation between the variance and the expected value of the maximum. Thus, for all the values of n considered, we see that as the variance of X increases, its expected maximum value also increases, leading to a direct (but complicated) relationship between the variance and the expected maximum.

We next consider the general continuous uniform-like case:

Theorem 3.2. *If*

$$f(x) = \begin{cases} \frac{1}{2a} \quad if \quad 0 < x < a \\ \\ 0 \quad if \quad a < x < 1-a \\ \\ \frac{1}{2a} \quad if \quad 1-a < x < 1 \end{cases}$$

then $\texttt{Var(X)} = \frac{1}{12}(\texttt{4a}^2 - \texttt{6a} + \texttt{3})$ *and* $\mathbb{E}(M) \approx 1 - \frac{2a}{n+1}$

Proof. We first find $\mathbb{E}(X)$:

$$\mathbb{E}(X) = \int_0^1 x f(x) dx$$

Therefore,

$$\begin{aligned} \mathbb{E}(X) &= \int_0^a \left(\frac{1}{2a}\right) x dx + \int_a^{1-a} 0 x dx + \int_{1-a}^1 \left(\frac{1}{2a}\right) x dx \\ &= \frac{1}{2}. \end{aligned}$$

Turning to the variance, we see that

$$\begin{aligned} \texttt{Var(X)} &= \int_0^1 \left(x - \frac{1}{2}\right)^2 f(x) dx \\ &= \frac{1}{2a} \int_0^a \left(x - \frac{1}{2}\right)^2 dx + \frac{1}{2a} \int_{1-a}^1 \left(x - \frac{1}{2}\right)^2 dx. \end{aligned}$$

Letting $u = x - \frac{1}{2}$, we see that

$$\frac{1}{2a} \int \left(x - \frac{1}{2}\right)^2 = \frac{1}{2a} \int u^3 du = \frac{1}{2a} \frac{u^3}{3} = \frac{u^3}{6a},$$

so that

$$
\begin{aligned}
\texttt{Var(X)} &= \frac{1}{6a}\left[(x-\frac{1}{2})^3\right]_0^a + \frac{1}{6a}\left[(x-\frac{1}{2})^3\right]_{1-a}^1 \\
&= \frac{1}{48a}\left[8x^3 - 12x + 6x + 1\right]_0^a + \frac{1}{48a}\left[8x^3 - 12x + 6x + 1\right]_{1-a}^1 \\
&= \frac{1}{24}(4a^2 - 6a + 3) + \frac{1}{24}(4a^2 - 6a + 3) = \frac{1}{12}(4a^2 - 6a + 3).
\end{aligned}
$$

Therefore, we can confirm our result from the previous chapter that when $a = \frac{1}{3}$,

$$\texttt{Var(X)} = \frac{1}{12}\left(\frac{4}{9} - 2 + 3\right) = \frac{13}{108} \approx 0.12037.$$

We now proceed to find the cumulative distribution function

$$
\begin{aligned}
F(x) &= \int_0^x f(t)dt \\
&= \int_0^x \frac{1}{2a}dt \\
&= \frac{1}{2a}x; \qquad\qquad 0 < x < a \\
F(x) &= \int_0^a f(x)dx + \int_a^x f(t)dt \\
&= \int_0^a \frac{1}{2a}dx + \int_a^x 0dt \\
&= \left[\frac{1}{2a}x\right]_0^a \\
&= \frac{1}{2}; \qquad\qquad a < x < 1-a \\
F(x) &= \int_0^a f(t)dt + \int_a^{1-a} f(t)dt + \int_{1-a}^x f(t)dt \\
&= 1 - \frac{1}{2a} + \frac{1}{2a}x \qquad\qquad 1-a < x < 1.
\end{aligned}
$$

Therefore, we can write the cumulative distribution function, $F(x)$ as:

$$
F(x) = \begin{cases}
\frac{1}{2a}x & if \;\; 0 < x < a \\
\frac{1}{2} & if \;\; a < x < 1-a \\
1 - \frac{1}{2a} + \frac{1}{2a}x & if \;\; 1-a < x < 1.
\end{cases}
$$

We proceed to find the properties of the maximum functions:

$$
F(x) = \left(\frac{1}{2}x\right)^n \Rightarrow f(x) = \frac{n}{2a}\left(\frac{x}{2a}\right)^{n-1}
$$

$$
F(x) = \left(\frac{1}{2}\right)^n \Rightarrow f(x) = 0
$$

$$F(x) = \left(1 - \frac{1}{2a} + \frac{1}{2a}x\right)^n \Rightarrow f(x) = \frac{n}{2a}\left(1 - \frac{1}{2a} + \frac{1}{2a}x\right)^{n-1}.$$

The expected maximum is calculated as follows:

$$\mathbb{E}(M) = \int_0^1 x f_M(x) dx.$$

Separating the integral into three parts, I_1, I_2, I_3, we get

$$\begin{aligned} I_1 &= \int_0^a x f(x) dx \\ &= \int_0^a x \left[\frac{n}{2a}\left(\frac{x}{2a}\right)^{n-1}\right] dx \\ &= n\left(\frac{1}{2a}\right)^n \frac{a^{n+1}}{n+1} \\ &= \left(\frac{1}{2}\right)^n (1 + o(1)); \end{aligned}$$

$$I_2 = \int_a^{1-a} x f(x) dx = \int_a^{1-a} x \cdot 0 dx = 0$$

and

$$I_3 = \int_{\frac{2}{3}}^1 x . \frac{3n}{2}\left(\frac{3x-1}{2}\right)^{n-1} dx.$$

In I_3 we let $u = 1 + \frac{x}{2a} - \frac{1}{2a}$, which leads to

$$\begin{aligned} I_3 &= \frac{n}{2a}\int_{\frac{1}{2}}^1 (2au + 1 - 2a)\, u^{n-1}(2a) du \\ &= n\int_{\frac{1}{2}}^1 (2au^n + u^{n-1} - 2au^{n-1}) du \\ &= n\left[\frac{2au^{n+1}}{n+1} + \frac{u^n}{n} - \frac{2au^n}{n}\right]_{\frac{1}{2}}^1 \\ &= \frac{n+1-2a}{n+1} - \left(\frac{1}{2}\right)^n\left(\frac{n+1-an-2a}{n+1}\right). \end{aligned}$$

Therefore, we have:

$$I_1+I_2+I_3 = \left(\frac{1}{2}\right)^n (1+o(1))+\frac{n+1-2a}{n+1}-\left(\frac{1}{2}\right)^n \left(\frac{n+1-an-2a}{n+1}\right) = 1-\frac{2a}{n+1}(1+o(1)),$$

as claimed. Note that when $a = \frac{1}{3}$, we recover the previous result.

$\square$

4 THE BETA FAMILY CASE

Here we look at the correlation between the variance and the expected value of the maximum for various values of n, and compare the results for different beta distributions:

In general, for the beta (α, β) distribution, we have

$$f(x) = \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1}.$$

Theorem 4.1. *For $\alpha = \beta = 1/2$, we have $f(x) = \frac{1}{\pi} x^{-\frac{1}{2}}(1-x)^{-\frac{1}{2}}$ with $\mathtt{Var(x)} = \frac{1}{8}$ and*

$$\mathbb{E}(M) = \int_0^1 nx \frac{1}{\pi\sqrt{x(1-x)}} \left(\frac{2}{\pi}\sin^{-1}\sqrt{x}\right)^{n-1} dx.$$

Proof.

$$\begin{aligned} f(x) &= \frac{\Gamma(\frac{1}{2}+\frac{1}{2})}{\Gamma(\frac{1}{2})\Gamma(\frac{1}{2})} x^{-\frac{1}{2}}(1-x)^{-\frac{1}{2}} \\ &= \frac{\Gamma(1)}{\Gamma(\frac{1}{2})\Gamma(\frac{1}{2})} x^{-\frac{1}{2}}(1-x)^{-\frac{1}{2}} \\ &= \frac{1}{\pi} x^{-\frac{1}{2}}(1-x)^{-\frac{1}{2}} \qquad (0 < x < 1). \end{aligned}$$

The graph of Beta(0.5,0.5) can be seen in Figure 1 on page 8.

Now, we have that

$$\begin{aligned}
\mathbb{E}(X) &= \int_0^1 x\frac{1}{\pi}x^{-\frac{1}{2}}(1-x)^{-\frac{1}{2}}dx \\
&= \frac{1}{\pi}\int_0^1 x^{\frac{1}{2}}(1-x)^{-\frac{1}{2}} \\
&= \frac{1}{\pi}\int_0^1 \left(\frac{x}{1-x}\right)^{\frac{1}{2}} dx \\
&= \frac{1}{\pi}\left[\sin^{-1}\sqrt{x} - \sqrt{-(x-1)x}\right]_0^1 \\
&= \frac{1}{2},
\end{aligned}$$

and also

$$\begin{aligned}
\texttt{Var(X)} &= \frac{1}{\pi}\int_0^1 x^{-\frac{1}{2}}\left(x-\frac{1}{2}\right)^2(1-x)^{-\frac{1}{2}}\,dx \\
&= \frac{1}{\pi}\int_0^1 x^{-\frac{1}{2}}\left(x^2-x+\frac{1}{4}\right)(1-x)^{-\frac{1}{2}}\,dx \\
&= \frac{1}{\pi}\int_0^1 \left(x^{\frac{3}{2}}-x^{\frac{1}{2}}+\frac{1}{4}x^{-\frac{1}{2}}\right)(1-x)^{-\frac{1}{2}}\,dx \\
&= \left[\frac{2x^3-3x^2+x+\sqrt{x-1}\sqrt{x}\log\left(\sqrt{x-1}+\sqrt{x}\right)}{4\pi\sqrt{-(x-1)x}}\right]_0^1 \\
&= \frac{1}{8} \\
&= 0.125.
\end{aligned}$$

Thus,

$$\begin{aligned}
F(x) &= \int_0^x \frac{1}{\pi}y^{-\frac{1}{2}}(1-y)^{-\frac{1}{2}}dy \\
&= \frac{1}{\pi}\int_0^x \frac{1}{(y-y^2)^{\frac{1}{2}}}dy \\
&= \frac{2}{\pi}\sin^{-1}\sqrt{x}
\end{aligned}$$

from which it follows that

$$\mathbb{P}(M \leq x) = F_M(x) = [F(x)]^n = \left(\frac{2}{\pi}\sin^{-1}\sqrt{x}\right)^n.$$

Therefore,

$$f_M(x) = n\frac{1}{\pi\sqrt{x(1-x)}}\left(\frac{2}{\pi}\sin^{-1}\sqrt{x}\right)^{n-1} dx,$$

and hence

$$\mathbb{E}(M) = \int_0^1 nx\frac{1}{\pi\sqrt{x(1-x)}}\left(\frac{2}{\pi}\sin^{-1}\sqrt{x}\right)^{n-1} dx.$$

□

Theorem 4.2. *The case of $\alpha = \beta = 1$ was considered in Example 1.1 where we had a uniform distribution on [0,1] with* $\texttt{Var(x)} = \frac{1}{12}$ *and* $\mathbb{E}(M) = 1 - \frac{1}{n+1}$.

Theorem 4.3. *When $\alpha = \beta = 2$ we have $f(x) = 6x(1-x)$ with* $\texttt{Var(x)} = \frac{1}{20}$ *and* $\mathbb{E}(M) = \int_0^1 nx(3x^2 - 2x^3)^{n-1}(6x - 6x^2)dx$.

Proof.

$$\begin{aligned} f(x) &= \frac{\Gamma(2+2)}{\Gamma(2)\Gamma(2)}x(1-x) \\ &= \frac{\Gamma(4)}{\Gamma(2)\Gamma(2)}x(1-x) \\ &= 6x(1-x) \qquad\qquad (0 < x < 1). \end{aligned}$$

Figure 6 shows the distribution of Beta(2,2).

Now, we have that,

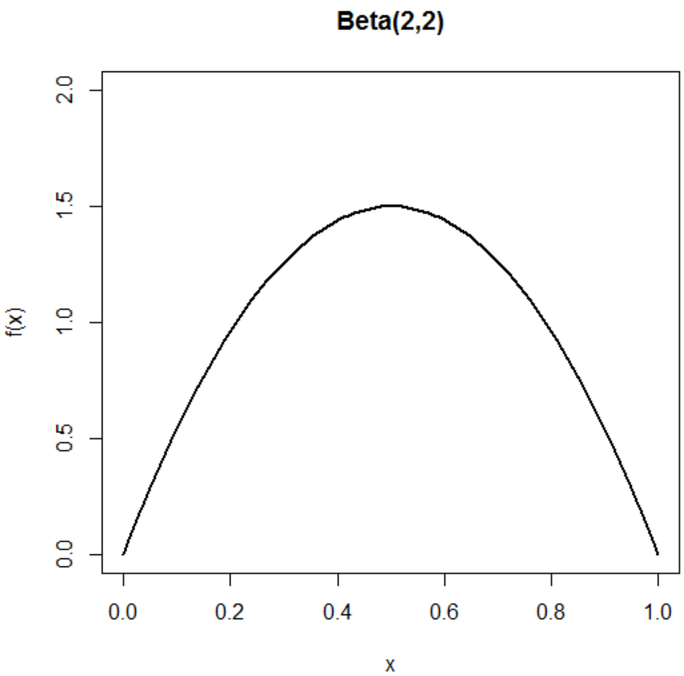


Figure 6: Distribution of Beta(2,2)

$$\begin{aligned}
\mathbb{E}(X) &= \int_0^1 x\,[6x(1-x)]\,dx = \int_0^1 \left(6x^2 - 6x^3\right) dx \\
&= \left(2x^3 - \frac{3}{2}x^4\right)_0^1 \\
&= \frac{1}{2},
\end{aligned}$$

and also

$$
\begin{aligned}
\texttt{Var(X)} &= \int_0^1 \left(x - \frac{1}{2}\right)^2 (6x(1-x))\, dx \\
&= \int_0^1 \left(x^2 - x + \frac{1}{4}\right)\left(6x - 6x^2\right) dx \\
&= \int_0^1 \left(\frac{3}{2}x - \frac{15}{2}x^2 + 12x^3 - 6x^4\right) dx \\
&= \left[\frac{3}{4}x^2 - \frac{5}{2}x^3 + 3x^4 - \frac{6}{5}x^5\right]_0^1 \\
&= \frac{1}{20}.
\end{aligned}
$$

Thus,

$$
\begin{aligned}
F(x) &= \int_0^x 6y(1-y)dy = 3y^2 - 2y^3|_0^x \\
&= 3x^2 - 2x^3,
\end{aligned}
$$

from which it follows that

$$F_M(x) = \mathbb{P}(M \leq x) = [F(x)]^n = \left(3x^2 - 2x^3\right)^n.$$

Therefore,

$$f_M(x) = n(6x - 6x^2)(3x^2 - 2x^3)^{n-1},$$

and hence

$$\mathbb{E}(X) = \int_0^1 nx(6x - 6x^2)(3x^2 - 2x^3)^{n-1},$$

as claimed. □

Theorem 4.4. *When $\alpha = \beta = 3$, we have $f(x) = 30x^2(1-x)^2$ with $\mathtt{Var(x)} = \frac{1}{28}$ and $\mathbb{E}(M) = \int_0^1 nx(30x^2 - 60x^3 + 30x^4)(10x^3 - 15x^4 + 6x^5)^{n-1}dx$.*

Proof.

$$\begin{aligned} f(x) &= \frac{\Gamma(3+3)}{\Gamma(3)\Gamma(3)}x^2(1-x)^2 \\ &= \frac{\Gamma(6)}{\Gamma(3)\Gamma(3)}x^2(1-x)^2 \\ &= 30x^2(1-x)^2 \qquad\qquad (0 < x < 1). \end{aligned}$$

Figure 7 shows the distribution of Beta(3,3). Also, we have that,

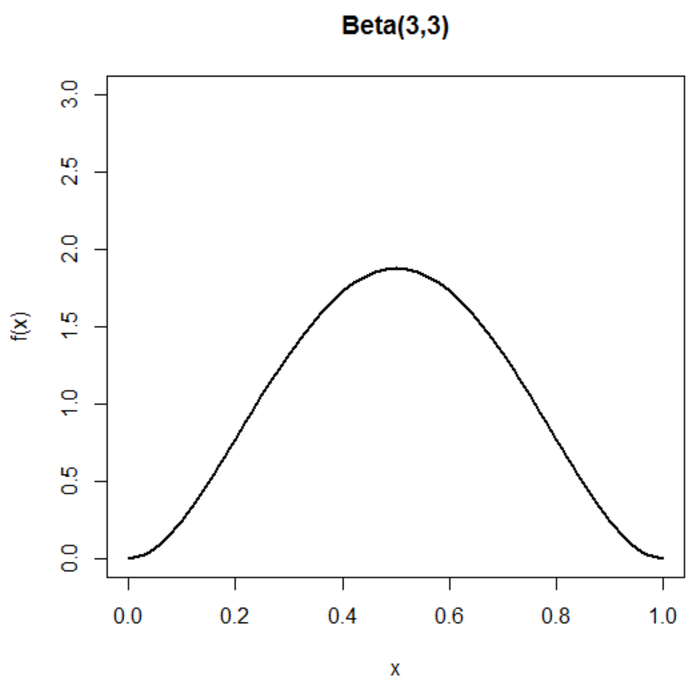


Figure 7: Distribution of Beta(3,3)

$$
\begin{aligned}
\mathbb{E}(X) &= \int_0^1 x\left[30x^2(1-x)^2\right]dx = 30\int_0^1 \left(x^3 - 2x^4 + x^5\right)dx \\
&= 30\left[\frac{x^4}{4} - \frac{2}{5}x^5 + \frac{x^6}{6}\right]_0^1 \\
&= \frac{1}{2},
\end{aligned}
$$

and

$$
\begin{aligned}
\texttt{Var(X)} &= 30\int_0^1 x^2\left(x-\frac{1}{2}\right)^2\left(6x(1-x)^2\right)dx \\
&= 30\int_0^1 \left(x^4 - x^3 + \frac{1}{4}x^2\right)\left(1-2x+x^2\right)dx \\
&= 30\int_0^1 \left(\frac{1}{4}x^2 - \frac{3}{2}x^3 + \frac{13}{4}x^4 - 3x^5 + x^6\right)dx \\
&= \left[\frac{1}{12}x^3 - \frac{3}{8}x^4 + \frac{13}{20}x^5 - \frac{x^6}{2} + \frac{x^7}{7}\right]_0^1 \\
&= \frac{1}{28} \\
&\approx 0.0357143.
\end{aligned}
$$

Thus,

$$
\begin{aligned}
F(x) &= \int_0^x 30y^2(1-y)^2dy = 30\int_0^x (y^2 - 2y^3 + y^4)dy \\
&= 30\left(\frac{y^3}{3} - \frac{1}{2}y^4 + \frac{1}{5}y^5\right)_0^x \\
&= 10x^3 - 15x^4 + 6x^5,
\end{aligned}
$$

which imples that

$$
F_M(x) = \mathbb{P}(M \le x) = [F(x)]^n = \left(10x^3 - 15x^4 + 6x^5\right)^n.
$$

Therefore,

$$f_M(x) = n(30x^2 - 60x^3 + 30x^4)(10x^3 - 15x^4 + 6x^5)^{n-1},$$

and thus,

$$\Rightarrow \mathbb{E}(x) = \int_0^1 nx(30x^2 - 60x^3 + 30x^4)(10x^3 - 15x^4 + 6x^5)^{n-1},$$

as asserted. $\square$

Table 14 shows the correlation between the variance and the expected value of the maximum for various values of n.

Table 14: Comparing Variance with Expected Maximum at different values of n

Model	Var	$\mathbb{E}(M)$									
1. $f(x) = \frac{1}{\pi}x^{-\frac{1}{2}}(1-x)^{-\frac{1}{2}}$	$\frac{1}{8}$	.9646	.9895	.9951	.9971	.9981	.9987	.9990	.9993	.9994	.9995
2. U[0,1]	$\frac{1}{12}$	.9091	.9524	.9677	.9756	.9804	.9836	.9859	.9877	.9890	.9901
3. $f(x) = 6x(1-x)$	$\frac{1}{20}$	.8312	.8815	.9037	.9168	.9258	.9324	.9375	.9416	.9450	.9478
4. $f(x) = 30x^2(1-x)^2$	$\frac{1}{28}$	.7839	.8324	.8553	.8695	.8795	.8871	.8911	.8980	.9022	.9058

For each value of n, we can see that as the variance decreases, there is a decrease in the corresponding expected maximum.

5 THE GENERAL CASE

Here we have $X \sim F, f$ on $[0, 1]$;

$$\mathbb{E}(X) = \frac{1}{2},$$

$$\texttt{Var}(\texttt{X}) = \sigma^2.$$

Now we want to find the expected value of $\mathbb{E}(\max(X_1, ...X_n))$. Here is a preliminary result that may be found, e.g., in a probability textbook such as Feller [1]

Theorem 5.1. *Given $X \geq 0$,*

$$\mathbb{E}(X) = \int_0^\infty (1 - F(x))dx$$

and

$$\mathbb{E}(X) = \sum_0^\infty \mathbb{P}(X > x)$$

in the continuous and discrete cases respectively.

Proof. In the discrete case the right hand side equals

$$\begin{aligned} \mathbb{P}(X = 1) + \mathbb{P}(X = 2) + \quad & \mathbb{P}(X = 3) \quad ... \\ + \mathbb{P}(X = 2) + \quad & \mathbb{P}(X = 3) \quad ... \\ + \quad & \mathbb{P}(X = 3) \quad ... \end{aligned}$$

The result follows by considering the above sum from top to bottom. In the continuous case,

$$\mathbb{E}(X) = \int_0^\infty x f(x) dx,$$

so on integrating by parts, we have

$$\begin{aligned}
\mathbb{E}(X) &= [xF(x)]_0^\infty - \int_0^\infty F(x)dx \\
&= \lim_{t\to\infty} tF(t) - \lim_{t\to\infty} \int_0^t F(x)dx \\
&= \lim_{t\to\infty} F(t) \lim_{t\to\infty} \int_0^t [1 - F(x)]\, dx \\
&= \int_0^\infty (1 - F(x))\, dx.
\end{aligned}$$

In particular, if

$$0 \leq X \leq 1,$$

$$\mathbb{E}(X) = \int_0^1 (1 - F(x))\, dx.$$

□

In general, we can prove that in the continuous case,

Theorem 5.2.

$$\mathbb{E}(X^r) = \int_0^\infty r x^{r-1} (1 - F(x))\, dx.$$

Proof. We know that by definition of expected value,

$$\mathbb{E}(X^r) = \int_0^\infty x^r f(x)dx.$$

Letting $u = x^r$ and $dv = f(x)dx$, we see that $du = rx^{r-1}$, and $v = F(x)$, so that

$$\begin{aligned}
\mathbb{E}(X^r) &= x^r F(x)_0^\infty - \int_0^\infty F(x) r x^{r-1} dx \\
&= \lim_{t\to\infty} t^r F(t) - \lim_{t\to\infty} \int_0^t F(x) r x^{r-1} dx \\
&= \lim_{t\to\infty} F(t) \int_0^t \left[r x^{r-1} - F(x) r x^{r-1} \right] \\
&= \lim_{t\to\infty} F(t) \int_0^t (1 - F(x)) r x^{r-1} \\
&= \lim_{t\to\infty} F(t) \lim_{t\to\infty} \int_0^t (1 - F(x)) r x^{r-1} \\
&= 1. \int_0^\infty (1 - F(x)) r x^{r-1},
\end{aligned}$$

as claimed. □

Our aim is to find the correlation between the variance and the expected value of the maximum.

Therefore, using the above, we need to study the correlation between

$a)$

$$\begin{aligned}
\mathtt{Var(X)} &= \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 \\
&= \int_0^\infty 2x\,(1 - F(x)) - \left(\int_0^\infty (1 - F(x) dx \right)^2, \\
&= \phi(x, F(x)).
\end{aligned}$$

and

$$b)$$

$$\begin{aligned}
\mathbb{E}(M_n) &= \int_0^\infty (1 - F_{MAX})dx \\
&= \int_0^\infty (1 - F^n(x))\, dx \\
&= \psi_n(x, F(x)).
\end{aligned}$$

This seems like a very hard problem for general n, but let us explore the simplest case, i.e., $n = 2$. Here

$$\mathtt{Var(X)} = \int_0^\infty \mathtt{2x\,(1 - F(x))} - \left(\int_0^\infty \mathtt{(1 - F(x)dx}\right)^2$$

$$= \int_0^\infty (1 - F(x)) \left(2x - \int_0^\infty (1 - F(x)\right) dx$$

and

$$\psi_2(x, F(x)) = \int_0^\infty \left(1 - F^2(x)\right) dx$$

This seems like a good starting point for a theoretical study in the general case.

6 CONCLUSIONS AND COMMENTS

We have seen that there exists a strong positive correlation between $\texttt{Var(X)}$ and $\mathbb{E}(M)$ for (i) the discrete uniform family; (ii) continuous uniform-like variables; and (iii) the beta family. This is not surprising or unexpected, but as seen in the last Chapter, the general case is quite difficult, even for $n = 2$. Here are some other issues:

(i) In our case, when the variables are symmetric, the expected value $\mathbb{E}(m)$ of the minimum can easily be seen to be ϵ_n whenever $\mathbb{E}(M) = 1 - \epsilon_n$. In general this is not the case for non-symmetric variables.

(ii) Our results hold easily when the variables X_i are symmetrically distributed on any compact interval $[a, b]$. However, it is hard to predict what may be true when the distribution of X is not symmetric. This is an area for further investigation.

BIBLIOGRAPHY

[1] W. Feller, "An Introduction to Probability Theory and its Applications", *2nd Ed*, Wiley, Princeton, NJ, 1971.

[2] Hogg, R.V. and A.T. Craig, "Introduction to Mathematical Statistics", *7th Ed*, Prentice Hall, NJ, 2013.

[3] Galambos Janos, J. Lechner and Emil Simiu, "Extreme Value Theory and Applications: Proceedings of the Conference on Extreme Value Theory and Applications", *Vol 1*, Gaithersburg, Maryland, 1993.

[4] Reiss R.-D. and M. Thomas, "Statistical Analysis of Extreme Values: with Applications to Insurance, Finance, Hydrology and Other Fields", *3rd ed*, Springer Verlag, Berlin, 2007.

VITA

BENEDICT ADJOGAH

Education:	B.A. Mathematics and Computer Science, University of Ghana, Legon, Legon Ghana 2009
	M.S. Mathematical Sciences, East Tennessee State University, Johnson City, Tennessee 2014
Professional Experience:	Research Analyst, Random House Ghana Limited, North Legon, Ghana, 2006–2009
	Teaching Assistant, Computer Science Department, University of Ghana, Legon, Ghana, 2009–2012
	Teaching Assistant,East Tennessee State University Johnson City, Tennessee, 2012–2014