Random variables and convex functions in Stochastic orderings with its applications in mathematics

Cyril Sarsoruo

Abstract

Randomness is a phenomenon that is mathematically studied in probability theory. Random quantities obtained have to be distributed over some graphs for interpretation. Mean is a descriptive quantity in statistics and is studied well using the convexity theory in pure mathematics. The paper shows the basic relationship of probability to random variables followed with convex function. Merging probability and convexity notions, we obtain sufficient conditions in the Ohlin lemma for stochastic convex orderings. The main results of the paper show formulations of two theorems using the Ohlin lemma in their proofs.

Keywords: convex function, convex stochastic ordering, expectation, Hermite-Hadamard inequality, Jensen functional inequality mean, Ohlin lemma, random variables.

Introduction

Certain problems can be solved mathematically by translating a particular problem into its mathematical model. Developing a mathematical model requires knowledge of mathematical theories. The main purpose of this research paper is to show a process of establishing elementary mathematical definitions, theorems and intertwining them to produce complex useful lemmas or theorems. Here, the Ohlin lemma is an example of this process and it is derived from combining continuous random variables and convex functions. The Ohlin lemma is used as a tool to prove two useful theorems of Hermite-Hadamanrd and Jensen functional inequalities (Rajba, 2014). We begin the paper by stating the relationship of probability to random variables and then to its expected or mean value. Following a definition of a convex function and then intertwining it with continuous random variables, we obtain the Ohlin lemma with its sufficient conditions for stochastics convex ordering. The conditions of the Ohlin lemma are used to show simple proofs of Hermite-Hadamand and the general case of finite Jensen functional inequalities in the application part of the paper.

Probability and random variables

The purpose of probability theory is to model random experiments so that we can draw inferences about them (Rosen, 2012). An experiment is a procedure that yields one of a given set of possible outcomes. The sample space of the experiment is the set of possible outcomes. An event is a subset of the sample space. Laplace's definition of the probability of an event with finitely many possible outcomes will now be stated.

Definition 1 If S is a finite nonempty sample space of equally likely outcomes, and E is an event, that is, a subset of *S,* then the probability of E is $p(E) = \frac{|E|}{|S|}$.

According to Laplace's definition, the probability of an event is between 0 and 1. Note that if E is an event from a finite sample space *S,* then $0 \leq |E| \leq |S|$. Thus,

$$0 \leq p(E) = \frac{|E|}{|S|} \leq 1.$$

Many problems are concerned with the numerical value associated with the outcome of an experiment. For instance, we may be interested in the total number of one bits in a randomly generated string of 10 bits or in the number of times a tail comes up when a coin is flipped 15 times. To study problems of this type we introduce the concept of a random variable.

Definition 2 Let S be a sample space and a function $X: S \to X(S)$ is said to be a random variable where $X(S) \subset R$.

From Definition 2, notice that a random variable is a function which assigns each possible outcome in the sample space to a real number. Example 1 gives an illustration of an application of definition 2.

Example 1 A fair coin is flipped 3 times. Let S be the sample space of 8 possible outcomes, and let X be a random variable that assigns to an outcome the number of heads in this outcome.

Random variable X is a function X: S → X(S) where X(S) = {0, 1, 2, 3} is the range which is the number of heads and
S={(TTT), (TTH), (THH), (HTT), (HHT), (HHH), (THT) ,(HTH)}

$$X(HHH) = 3,$$
$$X(HHT) = X(HTH) = X(THH) = 2,$$
$$X(TTH) = X(THT) = X(HTT) = 1,$$
$$X(TTT) = 0.$$

Combining Definitions 1 and 2, we arrive at the notion of the probability distribution of the random variable X forming Definition 3.

Definition 3 The probability distribution of a random variable X on a sample space S is the set of pairs $(r, p(X = r))$ for all $r \in X(S)$, where $p(X = r)$ is the probability that X takes the value r.

The set of pairs in this distribution is determined by the probabilities p(X = r) for $r \in X(S)$. The probability distribution of random variable X from Example 1 is given by p(X=3) = 1/8, p(X=2) = 3/8, p(X=1) = 3/8, p(X=0) = 1/8.

A random variable X can be either discrete or continuous. A random variable X is discrete if it has a finite or countable number of possible outcomes that can be listed as in Example 1. A random variable X is continuous if it has an uncountable number of possible outcomes, represented by the intervals. This research concerns continuous random variables so we introduce the following definitions.

Definition 4 Let X be a continuous random variable. Then a probability distribution or probability density function *(pdf) of X* is a function *f (x)* such that for any two numbers a and b*,*

$$P\ (a \leq X \leq b) = \int_{a}^{b} f(x)dx.$$

Definition 5 Let *F(b)* be the cumulative distribution function, for a continuous random variable *X* that is defined for every number $b\ \in R\ \ by$

$$F(b) = P\ (X \leq b) = \int_{-\infty}^{b} f(x)dx.$$

We now define the average or the mean of the continuous random variables, which is called expectation, denoted *E(X)* that will be used in the study.

Definition 6 The expected or mean value of a continuous random variable *X* with *pdf f (x)* is

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx.$$

Generally, the expectations of the continuous random variable *X* can be represented and analyzed mathematically using convex functions.

Convex function

At the core of the notion of convexity is the comparison of means (Niculescu & Persson, 2006). In this research, we look at the expectations of continuous random variables in convex functions. A convex function is defined as follows.

Definition 7 Let $J \subset \mathrm{R}$ be an open interval. The function f: $J \to \mathrm{R}$ is said to be convex if

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y) \qquad (1)$$

for all $x, y \in J$ and $\lambda \in [0,1]$.

It is said to be strictly convex if the inequality (1) holds strictly whenever x and y are distinct points and $\lambda \in (0,1)$.

Definition 8 Let $J \subset \mathrm{R}$ be an open interval. A function f: $J \to \mathrm{R}$ is said to be Jensen convex if and only if it satisfies the Jensen functional inequality

$$f\left(\frac{x+y}{2}\right) \leq \frac{f(x)+f(y)}{2} \qquad (2)$$

for all $x, y \in J$.

If the inequality in (2) for $x \neq y$ is sharp then f is said to be *strictly Jensen convex*.

Theorem 1 Let $J \subset \mathrm{R}$ be an open interval, and let $f : J \to \mathrm{R}$ be a given convex function. The function f satisfies (2) and is continuous if and only if it satisfies inequality (1) for all $x, y \in J$ and every $\lambda \in [0,1]$.

The proof to theorem 1 is omitted, as it will not be used in this paper.

Ohlin lemma

Definitions 1 to 8 and Theorem 1 give, the mathematical bases to formulate the Ohlin lemma (Ohlin, 1969) which provides sufficient condition for stochastic convex ordering (5) in lemma 1.

Lemma 1 Let X and Y be two continuous random variables with finite expectations such that

$$E(X) = E(Y). \qquad (3)$$

*I*f their cumulative distribution functions F_X and F_Y cross exactly one time, i.e., for there exist a point t_0 such that

$$F_X(t) \leq F_Y(t) \text{ if } t < t_0 \text{ and } F_X(t) \geq F_Y(t) \text{ if } t > t_0 \qquad (4)$$

then

$$Ef(X) \leq Ef(Y) \qquad (5)$$

for all convex functions $f : \mathrm{R} \rightarrow \mathrm{R}$.

The proof to lemma 1 is omitted because it involves theorems and other mathematical theories, which are not included in this paper and are not relevant in the context of this research.

Applications

The Ohlin lemma or lemma 1 is used as a tool, which is applied to the proof of the following theorems mainly used in convex stochastic convex ordering to compare the Means of the probability distributions of continuous variables. Firstly, the Ohlin lemma is used to give a simple proof of the Hermite-Hadamard inequality. This inequality gives us an estimate of the integral mean value of a continuous convex function (Rajba, 2014). Secondly, we introduce the general finite case of Jensen inequality and use the Ohlin lemma to prove it. The Jensen inequality in its simplest form states that the convex transformation of a mean is less than or equal to the mean applied after convex transformation (Kuzma, 1985). We formulate two theorems and apply conditions of Ohlin lemma to prove them.

Theorem 2 Let $J \subset \mathrm{R}$ be an open interval, and $f : J \rightarrow \mathrm{R}$ be a convex function with $a,b \in J$, $a < b$ then

$$f\left(\frac{a+b}{2}\right) \leq \frac{1}{b-a}\int_a^b f(x)dx \leq \frac{f(a)+f(b)}{2}. \qquad (6)$$

Proof We will prove the double inequality (6) by expanding conditions (3), (4) and (5) of the Ohlin lemma. Take $a,b \in J$ and $a < b$. Let X, Y, Z be three continuous random variables with measures $\mu_X = \delta_{(a+b)/2}$, μ_Y which is equally distributed in $[a,b]$ and $\mu_Z = \frac{1}{2}(\delta_a + \delta_b$, respectively.

1. We calculate the expectations of the three continuous random variables are X, Y and Z as follow:

$$E(X) = \int_{\mathbb{R}} td\mu_X(t) = \mu_X(\ \{\frac{a+b}{2}\}) \cdot \frac{a+b}{2} = \frac{a+b}{2};$$

$$E(Y) = \int_{\mathbb{R}} td\mu_Y(t) = \int_{[a,b]} t \cdot \frac{1}{b-a}dt = (\frac{1}{b-a}) \cdot \frac{a^2-b^2}{2} = \frac{a+b}{2};$$

$$E(Z) = \int_{\mathbb{R}} td\mu_Z(t) = a \cdot \mu_z(\{a\}) + b \cdot \mu_Z(\{b\}) = \frac{a+b}{2}.$$

From the calculations, we have shown that the expectations are equal with

$$E(X) = E(Y) = E(Z) = \tfrac{a+b}{2}$$

and we have satisfied condition (3).

2. Observe that the cumulative distributive functions are given by:

$$F_X(t) = \begin{cases} 0, & \text{t } \leq \frac{a+b}{2}, \\ 1, & 1 > \frac{a+b}{2} \end{cases},$$

and

$$F_Y(t) = \begin{cases} 0, & \text{t } \leq a, \\ \frac{1}{b-a}t - \frac{a}{b-a}, & \text{t } \in (a,b], \\ 1, & \text{t } > b, \end{cases}$$

and also

$$F_Z(t) = \begin{cases} 0, & \text{t } \leq a, \\ \frac{1}{2}, & \text{t } \in (a, b], \\ 1, & \text{t } > b. \end{cases}$$

Using the above calculated cumulative distributions functions $F_X(t)$, $F_Y(t)$ and $F_Z(t)$ we may write the following. For cumulative distributions functions

$F_X(t)$ and $F_Y(t)$ at point $t_0 = \frac{a+b}{2}$ we have

$$t \in (-\infty, \frac{a+b}{2}) \Longrightarrow F_X(t) \leq F_Y(t),$$

and

$$t \in (\frac{a+b}{2}, \infty) \Longrightarrow F_X(t) \geq F_Y(t).$$

Furthermore for cumulative distributions functions $F_Y(t)$ and $F_Z(t)$ at point $t_0 = \frac{a+b}{2}$ we get

$$t \in (-\infty, \frac{a+b}{2}) \Longrightarrow F_Y(t) \leq F_Z(t)$$

and

$$t \in (\frac{a+b}{2}, \infty) \Longrightarrow F_Y(t) \geq F_Z(t).$$

From the above observations, we have shown that (4) is satisfied.

3. Since (3) and (4) are fulfilled, we know from the Ohlin lemma that (5) is satisfied. Using (5) we will obtain (6). Observe that for convex orderings $Ef(X) \leq Ef(Y) \leq Ef(Z)$ we have

$$Ef(X) = \int_{\mathbb{R}} f(t)d\mu_X(t) = \mu_X(\ \{\frac{a+b}{2}\}) \cdot f(\frac{a+b}{2}) = f(\frac{a+b}{2}) \leq$$

$$\leq Ef(Y) = \int_{\mathbb{R}} f(t)d\mu_Y(t) = \int_a^b f(t)\frac{1}{b-a}dt = \frac{1}{b-a}\int_a^b f(t)dt \leq$$

$$\leq Ef(Z) = \int_{\mathbb{R}} f(t)d\mu_Z(t) = \int_a^b f(t)d\mu_Z(t) = f(a)\cdot\mu_z(\{a\}) + f(b)\cdot\mu_Z(\{b\}) =$$

$$= \frac{f(a)+f(b)}{2}.$$

This completes the proof of the double Hermite-Hadamard inequality.

General finite case of Jensen inequality

Theorem 3 is the general finite case of Jensen inequality that is derived from definition 8. This theorem replaces variables $a,b \in J$ with finite variables $x_1,x_2,...,x_n \in J$. We state the theorem and show the proof of it satisfies the conditions (3), (4) and (5) of the Ohlin lemma.

Theorem 3 Let $J \subset$ R be an open interval. If a function f: $J \rightarrow$ R is convex, then it satisfies Jensen's functional inequality

$$f(\frac{x_1+x_2+...+x_n}{n}) \leq \frac{f(x_1)+f(x_2)+...+f(x_n)}{n} \quad (7)$$

for all $x_1,x_2,...,x_n \in J$, $x_1 < x_2 < ... < x_n$ and $n \in$ N.

Proof We will prove (7) with help of the Ohlin lemma by verifying conditions (3), (4) and (5). Consider $x_1, x_2,..., x_n \in J$ such that $x_1 < x_2 < ... < x_n$. Let X and Y be two continuous random variables with measures $\mu_X = \delta_{(x1+x2+...+xn)/n}$ and $\mu_Y = \frac{1}{n}(\delta_{x_1} + \delta_{x_2} + ... + \delta_{x_n}$, respectively.

1. We calculate the expected values of condition (3) for all $t \in$ R as follows:

$$E(X) = \int_{\mathbb{R}} td\mu_X(t) = \mu_X(\ \{\frac{x_1+x_2+...+x_n}{n}\}) \cdot \frac{x_1+x_2+...+x_n}{n} = \\ = \frac{x_1+x_2+...+x_n}{n};$$

$$E(Y) = \int_{\mathbb{R}} td\mu_Y(t) = x_1 \cdot \mu_Y(\{x_1\}) + x_2 \cdot \mu_Y(\{x_2\}) + ... + x_n \cdot \mu_Y(\{x_n\}) = \\ = \frac{x_1+x_2+...+x_n}{n}.$$

From the calculations, we have shown that the expected values are equal with

$$E(X) = E(Y) = \tfrac{x_1+x_2+...+x_n}{n}.$$

2. We want to prove (4), the cumulative distributive functions are given by:

$$F_X(t) = \begin{cases} 0, & \text{t } \le \frac{x_1+x_2+...+x_n}{n}, \\ 1, & \text{t } > \frac{x_1+x_2+...+x_n}{n}, \end{cases}$$

and

$$F_Y(t) = \begin{cases} 0, & \text{t } \le x_1, \\ \frac{1}{n}, & \text{t } \in (x_1, x_2], \\ \frac{2}{n}, & \text{t } \in (x_2, x_3], \\ \vdots & \vdots \\ \frac{n-1}{n}, & \text{t } \in (x_{n-1}, x_n], \\ 1, & \text{t } > x_n. \end{cases}$$

Now $F_X(t)$ and $F_Y(t)$ are cumulative distribution functions and for every $t \in \mathbb{R}$ there exists $k \in \{0, 1, ..., n\}$ such that $F_Y(t) = \frac{k}{n}$, we have

$$F_X(t) = 0 \le \frac{k}{n} = F_Y(t), \text{ for } t < \frac{x_1 + ... + x_n}{n}$$

and

$$F_X(t) = 1 \ge \frac{k}{n} = F_Y(t), \text{ for } t > \frac{x_1 + ... + x_n}{n}.$$

We have satisfied (4) as claimed.

3. Since (3) and (4) are fulfilled above, then we can use (5) to obtain (7). We write

$$f(\frac{x_1 + x_2 + ... + x_n}{n}) = \mu_X(\{\frac{x_1 + x_2 + ... + x_n}{n}\}) \cdot f(\frac{x_1 + x_2 + ... + x_n}{n}) =$$

$$\int_{\mathbb{R}} f(t) d\mu_X(t) = Ef(X) \le$$

$$\le Ef(Y) = \int_{\mathbb{R}} f(t) d\mu_Y(t) = \int_{x_1}^{x_2} f(t) d\mu_Y(t) + \int_{x_2}^{x_3} f(t) d\mu_Y(t) +$$

$$... + \int_{x_n-1}^{x_n} f(t) d\mu_Y(t) = f(x_1) \cdot \mu_Y(\{x_1\}) + f(x_2) \cdot \mu_Y(\{x_2\}) + ... +$$

$$+ f(x_n) \cdot \mu_Y(\{x_n\}) = \frac{f(x_1) + f(x_2) + ... + f(x_n)}{n}.$$

Hence we have arrived at (7) and so we have proved that for convex ordering $Ef(X) \le Ef(Y)$.

All conditions (3), (4) and (5) are fulfilled. Thus, this completes the proof to (7).

Conclusion

In this paper, we observed a fundamental process of establishing elementary knowledge of mathematics, which can be used to derive complex lemmas like the Ohlin lemma. We further observed the application of the Ohlin lemma in the proofs to Theorems 2 and 3 which are the Hermite-Hadamard and the finite case of Jensen functional inequalities. These two theorems are applied to solve problems in areas such as optimization, finance, economics, and operation research where random continuous variables are involved. Having such proved theorems, we can utilize them to model and study the means of the probability distribution functions.

Reference

Kuczma. M. (1985). An introduction to the theory of functional fquations and inequalities. *Prace Naukowe Uniwersytetu Sla¸skiego w Katowicach*, 489, Pan´stwowe Wydawnictwo Naukowe, Uniwersytet Sla¸ski, Warszawa,´ Krako´w, Katowice.

Niculescu C. P. & Persson L. E. (2006). *Convex functions and their applications. A contemporary approach*. Springer: New York.

Ohlin. J. (1969). On a class of measures of dispersion with application to optimal reinsurance. *ASTIN Bulletin 5*, 249-266.

Rajba. T. (2014). On the Ohlin lemma for Hermite-Hadamard-Fejer type inequalities, *Math. Ineq. Appl. 17* (2), 557-571.

Rosen, H. K. (2012). *Discrete mathematics and its applications* (7 ed.) McGraw Hill: New York.

Acknowledgment

I would like to acknowledge my Master's thesis supervisor Dr Hab. Tomasz Szostok for introducing me to this research and giving me valuable research advice. I also acknowledge Professor Peter K. Anderson and Associate Professor Dr Martin Daniel for reviewing this paper in checking English grammatical errors.

However, responsibility for any errors of fact or opinion remains with the author.

Author

Cyril Sarsoruo is a lecturer in the Department of Mathematics and Computing Science at DWU. He holds a Master's degree in Theoretical Mathematics and specializes in mathematical analysis. His research interests include functional equations and inequalities. Email: csarsoruo@dwu.ac.pg