Cambridge University Press & Assessment
978-1-108-47279-1 — Econometric Analysis of Stochastic Dominance
Yoon-Jae Whang
Excerpt
More Information

1 Introduction

This chapter begins with the basic concepts and properties of stochastic dominance. It then gives examples of applications of stochastic dominance to various fields in economics: welfare analysis, finance, industrial organization, labor, international, health, and agricultural economics. The final subsection gives an overview of the subsequent chapters.

1.1 Concepts of Stochastic Dominance

1.1.1 Definitions

First-Order Stochastic Dominance (FSD) Let X_1 and X_2 be two (continuous) random variables with the cumulative distribution functions (CDFs) given by F_1 and F_2, respectively.[1] In economic applications, they typically correspond to incomes or financial returns of two different populations, which may vary regarding time, geographical regions or countries, or treatments. For $k = 1, 2$, let $Q_k(\tau) = \inf\{x : F_k(x) \geq \tau\}$ denote the quantile function of X_k, respectively, and let $\mathcal{U}_1$ denote the class of all monotone increasing (utility or social welfare) functions. If the functions are assumed to be differentiable, then we may write

$$\mathcal{U}_1 = \{u(\cdot) : u' \geq 0\}.$$

Definition 1.1.1 *The random variable X_1 is said to first-order stochastically dominate the random variable X_2, denoted by $F_1 \succeq_1 F_2$ (or X_1 FSD X_2),*[2] *if*

[1] Stochastic dominance can be defined for discrete or mixed continuous–discrete distributions. However, for the purpose of explanation, we shall mainly focus on continuous random variables, unless it is stated otherwise.

[2] To denote stochastic dominance relations, it is a convention to freely exchange the random variables with their respective distribution functions. For example, for first-order stochastic dominance, we may write $X_1 \succeq_1 X_2$ or F_1 FSD F_2. The same rule applies to the other concepts of stochastic dominance defined later.

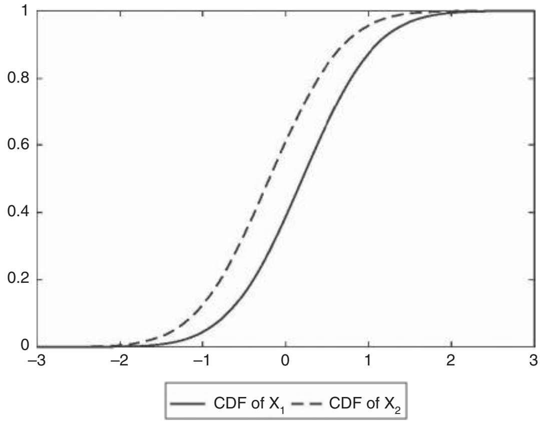

Figure 1.1 X_1 first-order stochastically dominates X_2

any of the following equivalent conditions holds: (1) $F_1(x) \leq F_2(x)$ *for all* $x \in \mathbb{R}$*; (2)* $E[u(X_1)] \geq E[u(X_2)]$ *for all* $u \in \mathcal{U}_1$*; and (3)* $Q_1(\tau) \geq Q_2(\tau)$ *for all* $\tau \in [0, 1]$.

This is the definition of *weak* stochastic dominance. If the inequalities hold with strict inequality for some $x \in \mathbb{R}$, some $u \in \mathcal{U}_1$, and some $\tau \in [0, 1]$, then the above serves as the definition of *strong* stochastic dominance, while one has *strict* stochastic dominance if the inequalities hold with strict inequality for all $x \in \mathbb{R}$, all $u \in \mathcal{U}_1$, and all $\tau \in [0, 1]$.[3] The equivalence of the three definitions will be discussed below.

Figure 1.1 illustrates two distributions with a first-order stochastic dominance relation. It shows that, when X_1 FSD X_2, the CDF of X_1 lies below that of X_2. To interpret the FSD relation, suppose that the random variables correspond to incomes of two different populations. Then, the inequality $F_1(x) \leq F_2(x)$ implies that the proportion of individuals in population 1 with incomes less than or equal to an income level x is not larger than the proportion of such individuals in population 2. If we measure poverty by the proportion of individuals earning less than a predetermined level of income (poverty line) x, then this implies that, whatever poverty line we choose, we have less poverty in F_1 than in F_2.[4] Therefore, the distribution F_1 would be preferred by any social planner having a welfare function that respects monotonicity ($u \in \mathcal{U}_1$),

[3] This classification is adopted from McFadden (1989, p. 115). The distinction among weak, strong, and strict dominance could be important in theoretical arguments. However, from a statistical point of view, the theoretically distinct hypotheses often induce the same test statistic and critical region, and hence the distinction is not very important; see McFadden (1989) and Davidson and Duclos (2000) for this point.

[4] See Section 5.2 for a general discussion about the relationship between poverty and SD concepts.

explaining the fact that we say that F_1 first-order stochastically dominates F_2 when the dominance of the CDFs as functions is the other way around.

To explain the FSD relation in an alternative perspective, write the (weak) first-order stochastic dominance relation $F_1 \succeq_1 F_2$ as

$$P(X_1 > x) \geq P(X_2 > x) \text{ for all } x \in \mathbb{R}. \tag{1.1.1}$$

Consider a portfolio choice problem of an investor and suppose that the random variables denote returns of some financial assets. Then, (1.1.1) implies that, for all values of x, the probability of obtaining returns not less than x is larger under F_1 than under F_2. Such a probability would be desired by every investor who prefers higher returns, explaining again the first-order stochastic dominance of F_1 over F_2. Conversely, if the two CDFs intersect, then (1.1.1) does not hold. In this case, one could find an investor with utility function $u \in \mathcal{U}_1$ such that $E[u(X_1)] > E[u(X_2)]$, and another investor with utility function $v \in \mathcal{U}_1$ such that $E[v(X_1)] < E[v(X_2)]$, violating the FSD of F_1 over F_2.

Second-Order Stochastic Dominance (SSD) To define the second-order stochastic dominance, let $\mathcal{U}_2$ denote the class of all monotone increasing and concave (utility or social welfare) functions. If the functions are assumed to be twice differentiable, then we may write

$$\mathcal{U}_2 = \{u(\cdot) : u' \geq 0,\ u'' \leq 0\}.$$

Definition 1.1.2 *The random variable X_1 is said to second-order stochastically dominate the random variable X_2, denoted by $F_1 \succeq_2 F_2$ (or X_1 SSD X_2), if any of the following equivalent conditions holds: (1) $\int_{-\infty}^{x} F_1(t)dt \leq \int_{-\infty}^{x} F_2(t)dt$ for all $x \in \mathbb{R}$; (2) $E[u(X_1)] \geq E[u(X_2)]$ for all $u \in \mathcal{U}_2$; and (3) $\int_0^{\tau} Q_1(p)dp \geq \int_0^{\tau} Q_2(p)dp$ for all $\tau \in [0, 1]$.*

For SSD, the accumulated area under F_1 must be smaller than the counterpart under F_2 below any value of x. If X_1 first-order dominates X_2, or equivalently, if $F_1(x)$ is smaller than $F_2(x)$ for all x, then it is easy to see that X_1 second-order dominates X_2, but the converse is not true.

Figure 1.2 illustrates that, even when there is no first-order stochastic dominance between them (i.e., when the two CDFs intersect), X_1 may second-order stochastically dominate X_2.

To have second-order stochastic dominance $F_1 \succeq_2 F_2$, for any negative area ($F_2 < F_1$) there should be a positive area ($F_1 < F_2$) which is greater than or equal to the negative area and which is located before the negative area. To relate this to the second definition (2) of SSD, consider the expression

$$E[u(X_1)] - E[u(X_2)] = \int_{-\infty}^{\infty} [F_2(x) - F_1(x)]\, u'(x)dx,$$

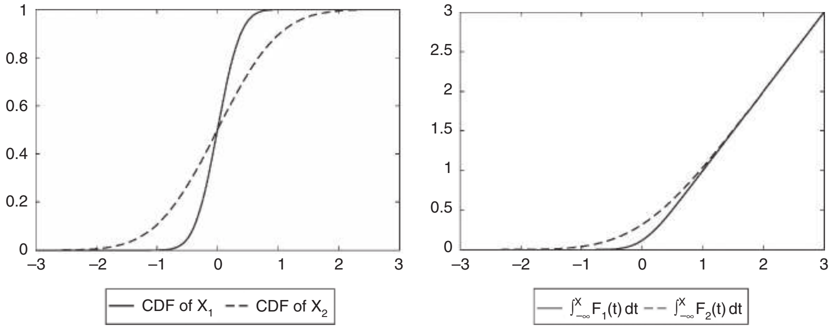

Figure 1.2 X_1 does not first-order stochastically dominate X_2, but X_1 second-order stochastically dominates X_2

which follows from integration by parts under regularity conditions (lemma 1 of Hanoch and Levy 1969; see also Equation 1.1.6). Whenever u' is a decreasing function (i.e., $u'' < 0$), the positive area is multiplied by a larger number $u'(x) > 0$ than the negative area which comes later on, so that the total integral becomes non-negative, establishing the second-order stochastic dominance of X_1 over X_2 under Definition 1.1.2 (2).

In the analysis of income distributions, the concavity assumption $u'' \leq 0$ implies that a transfer of income from a richer to a poorer individual always increases social welfare, which is a weaker form of the transfer principle (Dalton 1920). In the portfolio choice problem, on the other hand, the concavity assumption reflects risk aversion of an investor. That is, a risk-averse investor would prefer a portfolio with a guaranteed payoff to a portfolio without the guarantee, provided they have the same expected return. Therefore, the definition implies that any risk-averse investor would prefer a portfolio which dominates the other in the sense of SSD, because it yields a higher expected utility.

Higher-Order Stochastic Dominance The concept of stochastic dominance can be extended to higher orders. Higher-order SD relations correspond to increasingly smaller subsets of utility functions. Davidson and Duclos (2000) offer a very useful characterization of stochastic dominance of any order.

For $k = 1, 2$, define the *integrated CDF* and the *integrated quantile function* to be

$$F_k^{(s)}(x) = \begin{cases} F_k(x) & \text{for } s = 1 \\ \int_{-\infty}^{x} F_k^{(s-1)}(t)dt & \text{for } s \geq 2 \end{cases} \tag{1.1.2}$$

and

$$Q_k^{(s)}(x) = \begin{cases} Q_k(x) & \text{for } s = 1 \\ \int_{0}^{x} Q_k^{(s-1)}(t)dt & \text{for } s \geq 2. \end{cases}, \tag{1.1.3}$$

respectively. For $s \geq 1$, let

$$\mathcal{U}_s = \{u(\cdot) : u' \geq 0, u'' \leq 0, \ldots, (-1)^{s+1}u^{(s)} \geq 0\}$$

denote a class of (utility or social welfare) functions, where $u^{(s)}$ denotes the sth-order derivative of u (assuming that it exists).

Definition 1.1.3 *The random variable X_1 is said to stochastically dominate the random variable X_2 at order s, denoted by $F_1 \succeq_s F_2$, if any of the following equivalent conditions holds: (1) $F_1^{(s)}(x) \leq F_2^{(s)}(x)$ for all $x \in \mathbb{R}$ and $F_1^{(r)}(\infty) \leq F_2^{(r)}(\infty)$ for all $r = 1, \ldots, s-1$; (2) $E[u(X_1)] \geq E[u(X_2)]$ for all $u \in \mathcal{U}_s$; and (3) $Q_1^{(s)}(\tau) \geq Q_2^{(s)}(\tau)$ for all $\tau \in [0,1]$ and $Q_1^{(r)}(1) \geq Q_2^{(r)}(1)$ for all $r = 1, \ldots, s-1$.*

Whitmore (1970) introduces the concept of third-order stochastic dominance ($s = 3$, TSD) in finance; see also Whitmore and Findlay (1978). Levy (2016, section 3.8) relates the additional requirement $u''' \geq 0$ to the skewness of distributions and shows that TSD may reflect the preference for "positive skewness," i.e., investors dislike negative skewness but like positive skewness. Shorrocks and Foster (1987) show that the addition of a "transfer sensitivity" requirement leads to TSD ranking of income distributions. This requirement is stronger than the Pigou–Dalton principle of transfers since it makes regressive transfers less desirable at lower income levels.

If we let $s \to \infty$, then the class $\mathcal{U}_\infty$ of utility functions has marginal utilities that are completely monotone. This leads to the concept of infinite-order stochastic dominance, which is the weakest notion of stochastic dominance; see Section 5.4.3 for details.

Equivalence of the Definitions We now show the equivalence of the conditions that appear in the definitions of SD. For simplicity, we discuss the case of FSD and SSD, and assume that X_1 and X_2 have a common compact support, say $\mathcal{X} = [0, 1]$.[5]

We first establish the following lemma:

Lemma 1.1.1 *If $F_1(x) \leq F_2(x)$ for all $x \in \mathbb{R}$, then $EX_1 \geq EX_2$.*

Proof: Recall that, for any nonnegative random variable X with CDF F,

$$EX = \int_0^\infty P(X > t)\,dt = \int_0^\infty [1 - F(t)]\,dt; \tag{1.1.4}$$

[5] The equivalence results can be extended to general random variables, possibly with unbounded supports; see Hanoch and Levy (1969) and Tesfatsion (1976). The proofs in this subsection are based on Wolfstetter (1999, chapter 4) and Ross (1996, chapter 9). For a proof of strong stochastic dominance, see Levy (2016, section 3).

see, e.g., Billingsley (1995, equation 21.9). Therefore,

$$
\begin{aligned}
EX_1 - EX_2 &= \int_0^{\infty} [P(X_1 > t) - P(X_2 > t)]\, dt \\
&= \int_0^{\infty} [F_2(t) - F_1(t)]\, dt \geq 0. \qquad (1.1.5)
\end{aligned}
$$

□

The following theorem establishes the equivalence of (1) and (2) in Definition 1.1.1:[6]

Theorem 1.1.2 *$F_1(x) \leq F_2(x)$ for all $x \in \mathbb{R}$ if and only if $E[u(X_1)] \geq E[u(X_2)]$ for all $u \in \mathcal{U}_1$.*

Proof: Suppose that $F_1(x) \leq F_2(x)$ for all $x \in \mathbb{R}$ and let $u \in \mathcal{U}_1$ be an increasing function. Let $u^{-1}(z) = \inf\{x : u(x) > z\}$. For any $z \in \mathbb{R}$, we have

$$
\begin{aligned}
P(u(X_1) > z) &= P\left(X_1 > u^{-1}(z)\right) \\
&= 1 - F_1\left(u^{-1}(z)\right) \\
&\geq 1 - F_2\left(u^{-1}(z)\right) \\
&= P\left(X_2 > u^{-1}(z)\right) \\
&= P(u(X_2) > z).
\end{aligned}
$$

Therefore, by Lemma 1.1.1, we have $E[u(X_1)] \geq E[u(X_2)]$ for any $u \in \mathcal{U}_1$.

Conversely, suppose that $E[u(X_1)] \geq E[u(X_2)]$ for all $u \in \mathcal{U}_1$. Let

$$
u_x(z) = \begin{cases} 1 & \text{if } z > x \\ 0 & \text{if } z \leq x \end{cases}.
$$

Clearly, $u_x(\cdot) \in \mathcal{U}_1$ for each x. Therefore, for each $x \in \mathbb{R}$,

$$
\begin{aligned}
P(X_1 > x) &= E[u_x(X_1)] \\
&\geq E[u_x(X_2)] \\
&= P(X_2 > x).
\end{aligned}
$$

□

For SSD, the following theorem establishes the equivalence of (1) and (2) in Definition 1.1.2:[7]

Theorem 1.1.3 *$\int_0^x F_1(t)dt \leq \int_0^x F_2(t)dt$ for all $x \in \mathcal{X}$ if and only if $E[u(X_1)] \geq E[u(X_2)]$ for all $u \in \mathcal{U}_2$.*

[6] The equivalence of the conditions (1) and (3) easily follows from monotonicity of the CDFs.
[7] For a proof of the equivalence of the conditions (1) and (3), see Thistle (1989, proposition 4).

Proof: Suppose that $E[u(X_1)] \geq E[u(X_2)]$ for all $u \in \mathcal{U}_2$. Consider the following function:

$$u_x(z) = \begin{cases} z & \text{if } z \leq x \\ x & \text{if } z > x \end{cases}.$$

Obviously, for each $x \in \mathcal{X}$, $u_x(\cdot) \in \mathcal{U}_2$ so that

$$\begin{aligned} 0 &\leq E\left[u_x(X_1)\right] - E\left[u_x(X_2)\right] \\ &= \int_0^x \left[1 - F_1(t)\right] dt - \int_0^x \left[1 - F_2(t)\right] dt \\ &= \int_0^x \left[F_2(t) - F_1(t)\right] dt. \end{aligned}$$

Conversely, suppose that $\int_0^x F_1(t)dt \leq \int_0^x F_2(t)dt$ for all $x \in \mathcal{X}$. Since monotonicity implies differentiability almost everywhere (a.e.), we have $u' > 0$ and $u'' \leq 0$ a.e. for each $u \in \mathcal{U}_2$. Therefore, by integration by parts, we have

$$\begin{aligned} \Delta u &:= E\left[u(X_1)\right] - E\left[u(X_2)\right] \\ &= -\int_0^1 u(x) d\left[F_2(x) - F_1(x)\right] \\ &= \int_0^1 u'(x)\left[F_2(x) - F_1(x)\right] dx \qquad (1.1.6) \\ &= u'(1)\int_0^1 \left[F_2(t) - F_1(t)\right] dt \qquad (1.1.7) \\ &\quad - \int_0^1 u''(t) \int_0^t \left[F_2(s) - F_1(s)\right] ds dt. \end{aligned}$$

Since $u' > 0$ and $u'' \leq 0$, the assumed condition $\int_0^x \left[F_2(t) - F_1(t)\right] dt \geq 0$ for all $x \in \mathcal{X}$ implies immediately $\Delta u \geq 0$. This establishes Theorem 1.1.3. □

1.1.2 Basic Properties of Stochastic Dominance

While stochastic dominance relations compare *whole* distribution functions, they are also related to the moments and other aspects of distributions.

Let $supp(F)$ denote the support of distribution F. The following theorem gives sufficient and necessary conditions for the first-order stochastic dominance.

Theorem 1.1.4 *Let X_1 and X_2 be random variables with distribution functions F_1 and F_2, respectively. (1) If $P(X_2 \leq X_1) = 1$, then X_1 FSD X_2; (2) If $\min\{supp(F_1)\} \geq \max\{supp(F_2)\}$, then X_1 FSD X_2; (3) If X_1 FSD X_2, then $EX_1 \geq EX_2$ and $\min\{supp(F_1)\} \geq \min\{supp(F_2)\}$.*

(1) and (2) in the above theorem give sufficient conditions for the FSD. (1) holds because, if X_1 is not smaller than X_2 (with probability 1), then

$$\begin{aligned}&X_1 \le x \text{ implies } X_2 \le x \text{ for all } x\\ &\quad\Longrightarrow \{X_1 \le x\} \subseteq \{X_2 \le x\} \text{ for all } x\\ &\quad\Longrightarrow P(X_1 \le x) \le P(X_2 \le x) \text{ for all } x\\ &\quad\Longrightarrow F_1(x) \le F_2(x) \text{ for all } x.\end{aligned}$$ [8]

For example, if $X_1 = X_2 + a$ for a constant $a > 0$, then (1) implies that X_1 FSD X_2. (2) says that if the minimum of the support of F_1 is not less than the maximum of the support of F_2, then we have first-order stochastic dominance of X_1 over X_2. This follows directly from (1).

On the other hand, (3) gives necessary conditions for FSD. That is, if X_1 FSD X_2, then the mean of X_1 is not smaller than the mean of X_2. This follows from the expression[9]

$$EX_1 - EX_2 = \int_{-\infty}^{\infty} [F_2(x) - F_1(x)]\, dx, \tag{1.1.8}$$

which is nonnegative, provided the integral exists; see also Lemma 1.1.1. Also, if X_1 FSD X_2, then the minimum of the support of F_1 is not smaller than that of F_2. This is called the "left tail" condition because it implies that F_2 has a thicker left tail than F_1. This result holds because, otherwise, there would exist a value x_0 such that $F_1(x_0) > F_2(x_0)$, and hence X_1 could not first-order stochastically dominate X_2.

For second-order stochastic dominance, analogous conditions can be established (the proofs are also similar):

Theorem 1.1.5 *Let X_1 and X_2 be random variables with distribution functions F_1 and F_2, respectively. (1) If X_1 FSD X_2, then X_1 SSD X_2; (2) If* $\min\{supp(F_1)\} \ge \max\{supp(F_2)\}$, *then X_1 SSD X_2; (3) If X_1 SSD X_2, then $EX_1 \ge EX_2$ and* $\min\{supp(F_1)\} \ge \min\{supp(F_2)\}$.

In the above theorem, (3) shows that $EX_1 \ge EX_2$ is a necessary condition for the SSD. Is there any general condition on variances which is also a necessary condition for the SSD? In general, the answer is no. However, for distributions with an equal mean, we can state a necessary condition for the SSD using their variances.

8 Here, the notation '$A \Longrightarrow B$' means 'A implies B'.
9 This holds because, for any random variable X with CDF F,

$$EX = \int_0^{\infty} [1 - F(x) - F(-x)]\, dx,$$

provided the integral exists.

Theorem 1.1.6 *Let X_1 and X_2 be random variables with $EX_1 = EX_2$. If X_1 SSD X_2, then $Var(X_1) \leq Var(X_2)$.*

To see this, take, for example, a quadratic utility function $u(x) = x + \beta x^2$ for $\beta < 0$, which certainly lies in $\mathcal{U}_2$. Then, $Eu(X_1) \geq Eu(X_2)$ and $EX_1 = EX_2$ together immediately imply that $Var(X_1) \leq Var(X_2)$. The mean-variance approach in the portfolio choice problem compares only the first two moments of distributions. A natural question would be whether F_1 second-order stochastically dominates F_2, if F_1 has larger mean and smaller variance than F_2. The answer is no, in general. This can be illustrated using the following counterexample (Levy, 1992, p. 567):

x	$P(X_1 = x)$	x	$P(X_2 = x)$
1	0.80	10	0.99
100	0.20	1000	0.01

Note that $EX_1 = 20.8 > EX_2 = 19.9$ and $Var(X_1) = 1468 < Var(X_2) = 9703$. Hence, X_1 dominates X_2 by the mean-variance criterion. However, X_1 does not second-order stochastically dominate X_2 because a risk-averse investor with utility function $u(x) = \log(x)$ would prefer X_2 over X_1, since $Eu(X_1) = 0.4 < Eu(X_2) = 1.02$; see the next subsection for another example with continuously distributed random variables.

The foregoing discussion implies that there is no direct relationship between the mean-variance approach and the stochastic dominance approach in general. However, in the special case of normal distributions, stochastic dominance can be related to mean-variance in the following sense:

Theorem 1.1.7 *Let X_1 and X_2 be random variables with normal distributions. Then, (1) $EX_1 > EX_2$ and $Var(X_1) = Var(X_2)$ if and only if X_1 FSD X_2; (2) if $EX_1 > EX_2$ or $Var(X_1) < Var(X_2)$, then X_1 SSD X_2.*

For more complete discussions on the properties of stochastic dominance, the reader may refer to Levy (2016) and Wolfstetter (1999, chapters 4–5).

1.1.3 A Numerical Example

The mean-variance criterion has been widely adopted in portfolio choice problems. It is a simple performance indicator comparing only the first two moments of distributions; whenever the mean is higher and the variance is lower for one distribution than for the other, the former distribution is preferred. However, it is well known that the criterion is valid only in certain cases: (1) when the utility function is quadratic, and (2) when the distributions of the portfolios are all members of a two-parameter family; see Hanoch and Levy (1969). In reality, however, the assumptions are restrictive and the stochastic dominance approach provides an ordering of prospects under much less restrictive conditions.

To illustrate how the two approaches yield different results, we present a simple numerical example using two prospects, X and Y, with probability density functions (PDFs) and cumulative distribution functions (CDFs) given by

$$f_X(x) = 0.1 \cdot 1(0 \le x < 1 \text{ or } 2 \le x \le 3) + 0.8 \cdot 1(1 \le x < 2),$$
$$f_Y(x) = 0.5 \cdot 1(0.5 \le x \le 2.5)$$

and

$$F_X(x) = 0.1x \cdot 1(0 \le x < 1) + (0.8x - 0.7) \cdot 1(1 \le x < 2)$$
$$\qquad + (0.1x + 0.7) \cdot 1(2 \le x \le 3) + 1(x > 3),$$
$$F_Y(x) = 0.5(x - 0.5) \cdot 1(0.5 \le x \le 2.5) + 1(x > 2.5),$$

respectively, where $1(\cdot)$ denotes the indicator function. Figure 1.3 depicts the PDFs and the CDFs of the prospects. Their expected values and variances are given by $E(X) = 3/2$, $Var(X) = 17/60$, $E(Y) = 3/2$, and $Var(Y) = 1/3$.

In terms of the mean-variance criterion, the prospect X is more efficient than the prospect Y. However, X does not second-order stochastically dominate Y, which can easily be observed from Figure 1.3. Since the value of the CDF of X is greater than that of Y over the region [0, 0.5], the integrated area of the distribution of X is greater than that of the distribution of Y. This violates the second-order stochastic dominance of X over Y.

In reality, we do not observe the population distributions F_X and F_Y, but rather a sample randomly drawn from the distributions. This motivates us

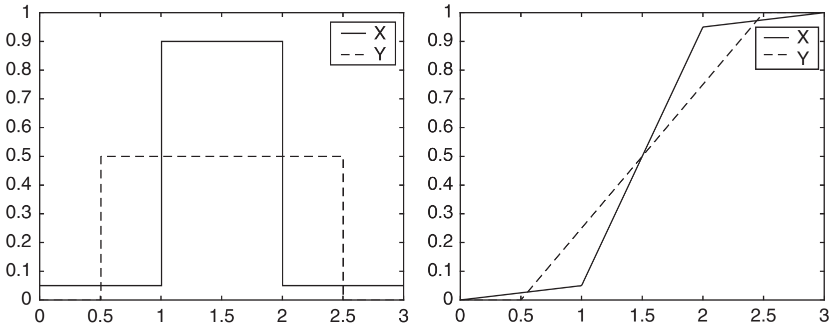

Figure 1.3 PDFs (left) and CDFs (right) for the simulation design