Springer Series in Statistics

Moshe Shaked
J. George Shanthikumar

Stochastic Orders

Springer

Moshe Shaked
Department of Mathematics
University of Arizona
Tucson, AZ 85721
shaked@math.arizona.edu

J. George Shanthikumar
Department of Industrial Engineering and Operations Research
University of California, Berkeley
Berkeley, CA 94720
shanthikumar@ieor.berkeley.edu

Library of Congress Control Number: 2006927724

ISBN-10: 0-387-32915-3
ISBN-13: 978-0387-32915-4

Printed on acid-free paper.

9 8 7 6 5 4 3 2 1

springer.com

To my wife *Edith* and to my children
Tal, Shanna, and *Lila*
M.S.

To my wife *Mellony* and to my children
Devin, *Rajan*, and *Sohan*
J.G.S.

To my wife Edith
and to my children Tal
Shanna
Lila

To my wife Mellony
and to my children Devin
Rajan
Sohan

Preface

Stochastic orders and inequalities have been used during the last 40 years, at an accelerated rate, in many diverse areas of probability and statistics. Such areas include reliability theory, queuing theory, survival analysis, biology, economics, insurance, actuarial science, operations research, and management science. The purpose of this book is to collect in one place essentially all that is known about these orders up to the present. In addition, the book illustrates some of the usefulness and applicability of these stochastic orders.

This book is a major extension of the first six chapters in Shaked and Shanthikumar [515]. The idea that led us to write those six chapters arose as follows. In our own research in reliability theory and operations research we have been using, for years, several notions of stochastic orders. Often we would encounter a result that we could easily (or not so easily) prove, but we could not tell whether it was known or new. Even when we were sure that a result was known, we would not know right away where it could be found. Also, sometimes we would prove a result for the purpose of an application, only to realize later that a stronger result (stronger than what we needed) had already been derived elsewhere. We also often have had difficulties giving a reference for *one* source that contained everything about stochastic orders that we needed in a particular paper. In order to avoid such difficulties we wrote the first six chapters in Shaked and Shanthikumar [515].

Since 1994 the theory of stochastic orders has grown significantly. We think that now is the time to put in one place essentially all that is known about these orders. This book is the result of this effort.

The simplest way of comparing two distribution functions is by the comparison of the associated means. However, such a comparison is based on only two single numbers (the means), and therefore it is often not very informative. In addition to this, the means sometimes do not exist. In many instances in applications one has more detailed information, for the purpose of comparison of two distribution functions, than just the two means. Several orders of distribution functions, that take into account various forms of possible knowledge about the two underlying distribution functions, are studied in Chapters 1 and 2.

When one wishes to compare two distribution functions that have the same mean (or that are centered about the same value), one is usually interested in the comparison of the dispersion of these distributions. The simplest way of doing it is by the comparison of the associated standard deviations. However, such a comparison, again, is based on only two single numbers, and therefore it is often not very informative. In addition to this, again, the standard deviations sometimes do not exist. Several orders of distribution functions, which take into account various forms of possible knowledge about the two underlying distribution functions (in addition to the fact that they are centered about the same value), are studied in Chapter 3. Orders that can be used for the joint comparison of both the location and the dispersion of distribution functions are studied in Chapters 4 and 5. The analogous orders for multivariate distribution functions are studied in Chapters 6 and 7.

When one is interested in the comparison of a sequence of distribution functions, associated with the random variables X_i, $i = 1, 2, \ldots$, then one can use, of course, any of the orders described in Chapters 1–7 for the purpose of comparing any two of these distributions. However, the parameter i may now introduce some patterns that connect all the underlying distributions. For example, suppose not only that the random variables X_i, $i = 1, 2, \ldots$, increase stochastically in i, but also that the increase is sharper for larger i's. Then the sequence X_i, $i = 1, 2, \ldots$, is stochastically increasing in a convex sense. Such notions of stochastic convexity and concavity are studied in Chapter 8.

Notions of positive dependence of two random variables X_1 and X_2 have been introduced in the literature in an effort to mathematically describe the property that "large (respectively, small) values of X_1 go together with large (respectively, small) values of X_2." Many of these notions of positive dependence are defined by means of some comparison of the joint distribution of X_1 and X_2 with their distribution under the theoretical assumption that X_1 and X_2 are independent. Often such a comparison can be extended to general pairs of bivariate distributions with given marginals. This fact led researchers to introduce various notions of positive dependence orders. These orders are designed to compare the strength of the positive dependence of the two underlying bivariate distributions. Many of these orders can be further extended to comparisons of general multivariate distributions that have the same marginals. In Chapter 9 we describe these orders.

We have in mind a wide spectrum of readers and users of this book. On one hand, the text can be useful for those who are already familiar with many aspects of stochastic orders, but who are not aware of all the developments in this area. On the other hand, people who are not very familiar with stochastic orders, but who know something about them, can use this book for the purpose of studying or widening their knowledge and understanding of this important area.

We wish to thank Haijun Li, Asok K. Nanda, and Taizhong Hu for critical readings of several drafts of the manuscript. Their comments led to a substantial improvement in the presentation of some of the results in these chapters. We also thank Yigal Gerchak and Marco Scarsini for some illuminating suggestions. We thank our academic advisors John A. Buzacott (of J. G. S.) and Albert W. Marshall (of M. S.) who, years ago, introduced us to some aspects of the area of stochastic orders.

Tucson, Berkeley,
August 16, 2006

Moshe Shaked
J. George Shanthikumar

Contents

Note

Throughout the book "increasing" means "nondecreasing" and "decreasing" means "nonincreasing." Expectations are assumed to exist whenever they are written. The "inverse" of a monotone function (which is not strictly monotone) means the right continuous version of it, unless stated otherwise. For example, if F is a distribution function, then the right continuous version of its inverse is $F^{-1}(u) = \sup\{x : F(x) \le u\}$, $u \in [0,1]$.

The following aging notions will be encountered often throughout the text. Let X be a random variable with distribution function F and survival function $\overline{F} \equiv 1 - F$.

(i) The random variable X (or its distribution) is said to be IFR [increasing failure rate] if $\overline{F}$ is logconcave. It is said to be DFR [decreasing failure rate] if $\overline{F}$ is logconvex.

(ii) The nonnegative random variable X (or its distribution) is said to be IFRA [increasing failure rate average] if $-\log \overline{F}$ is starshaped; that is, if $-\log \overline{F}(t)/t$ is increasing in $t \ge 0$. It is said to be DFRA [decreasing failure rate average] if $-\log \overline{F}$ is antistarshaped; that is, if $-\log \overline{F}(t)/t$ is decreasing in $t \ge 0$.

(iii) The nonnegative random variable X (or its distribution) is said to be NBU [new better than used] if $\overline{F}(s)\overline{F}(t) \ge \overline{F}(s+t)$ for all $s \ge 0$ and $t \ge 0$. It is said to be NWU [new worse than used] if $\overline{F}(s)\overline{F}(t) \le \overline{F}(s+t)$ for all $s \ge 0$ and $t \ge 0$.

(iv) The random variable X (or its distribution) is said to be DMRL [decreasing mean residual life] if $\frac{\int_t^\infty \overline{F}(s)\mathrm{d}s}{\overline{F}(t)}$ is decreasing in t over $\{t : \overline{F}(t) > 0\}$. It is said to be IMRL [increasing mean residual life] if $\frac{\int_t^\infty \overline{F}(s)\mathrm{d}s}{\overline{F}(t)}$ is increasing in t over $\{t : \overline{F}(t) > 0\}$.

(v) The nonnegative random variable X (or its distribution) is said to be NBUE [new better than used in expectation] if $\frac{\int_t^\infty \overline{F}(s)\mathrm{d}s}{\overline{F}(t)} \le EX$ for all

$t \geq 0$. It is said to be NWUE [new worse than used in expectation] if $\frac{\int_t^\infty \overline{F}(s)\mathrm{d}s}{\overline{F}(t)} \geq EX$ for all $t \geq 0$.

The majorization order will be used in some places in the text. Recall from Marshall and Olkin [383] that a vector $\boldsymbol{a} = (a_1, a_2, \ldots, a_n)$ is said to be smaller in the majorization order than the vector $\boldsymbol{b} = (b_1, b_2, \ldots, b_n)$ (denoted $\boldsymbol{a} \prec \boldsymbol{b}$) if $\sum_{i=1}^n a_i = \sum_{i=1}^n b_i$ and if $\sum_{i=1}^j a_{[i]} \leq \sum_{i=1}^j b_{[i]}$ for $j = 1, 2, \ldots, n-1$, where $a_{[i]}$ $[b_{[i]}]$ is the ith largest element of $\boldsymbol{a}$ $[\boldsymbol{b}]$, $i = 1, 2, \ldots, n$. An n-dimensional function ϕ is called *Schur convex* [*concave*] if $\boldsymbol{a} \prec \boldsymbol{b} \Longrightarrow \phi(\boldsymbol{a}) \leq [\geq]\ \phi(\boldsymbol{b})$.

The notation $\mathbb{N} \equiv \{\ldots, -1, 0, 1, \ldots\}$, $\mathbb{N}_+ \equiv \{0, 1, \ldots\}$, and $\mathbb{N}_{++} \equiv \{1, 2, \ldots\}$ will be used in this text.

1

Univariate Stochastic Orders

In this chapter we study stochastic orders that compare the "location" or the "magnitude" of random variables. The most important and common orders that are considered in this chapter are the usual stochastic order $\le_{\rm st}$, the hazard rate order $\le_{\rm hr}$, and the likelihood ratio order $\le_{\rm lr}$. Some variations of these orders, and some related orders, are also examined in this chapter.

1.A The Usual Stochastic Order

1.A.1 Definition and equivalent conditions

Let X and Y be two random variables such that

$$P\{X > x\} \le P\{Y > x\} \quad \text{for all } x \in (-\infty, \infty). \tag{1.A.1}$$

Then X is said to be *smaller than* Y *in the usual stochastic order* (denoted by $X \le_{\rm st} Y$). Roughly speaking, (1.A.1) says that X is less likely than Y to take on large values, where "large" means any value greater than x, and that this is the case for all x's. Note that (1.A.1) is the same as

$$P\{X \le x\} \ge P\{Y \le x\} \quad \text{for all } x \in (-\infty, \infty). \tag{1.A.2}$$

It is easy to verify (by noting that every closed interval is an infinite intersection of open intervals) that $X \le_{\rm st} Y$ if, and only if,

$$P\{X \ge x\} \le P\{Y \ge x\} \quad \text{for all } x \in (-\infty, \infty). \tag{1.A.3}$$

In fact, we can recast (1.A.1) and (1.A.3) in a seemingly more general, but actually an equivalent, way as follows:

$$P\{X \in U\} \le P\{Y \in U\} \quad \text{for all upper sets } U \subseteq (-\infty, \infty). \tag{1.A.4}$$

(In the univariate case, that is on the real line, a set U is an upper set if, and only if, it is an open or a closed right half line.) In the univariate case the equivalence of (1.A.4) with (1.A.1) and (1.A.3) is trivial, but in Chapter 6 it will be seen that the generalizations of each of these three conditions to the multivariate case yield different definitions of stochastic orders.

Still another way of rewriting (1.A.1) or (1.A.3) is the following:

$$E[I_U(X)] \leq E[I_U(Y)] \quad \text{for all upper sets } U \subseteq (-\infty, \infty), \tag{1.A.5}$$

where I_U denotes the indicator function of U. From (1.A.5) it follows that if $X \leq_{\text{st}} Y$, then

$$E\Big[\sum_{i=1}^{m} a_i I_{U_i}(X)\Big] - b \leq E\Big[\sum_{i=1}^{m} a_i I_{U_i}(Y)\Big] - b \tag{1.A.6}$$

for all $a_i \geq 0$, $i = 1, 2, \ldots, m$, $b \in (-\infty, \infty)$, and $m \geq 0$. Given an increasing function ϕ, it is possible, for each m, to define a sequence of U_i's, a sequence of a_i's, and a b (all of which may depend on m), such that as $m \to \infty$ then (1.A.6) converges to

$$E[\phi(X)] \leq E[\phi(Y)], \tag{1.A.7}$$

provided the expectations exist. It follows that $X \leq_{\text{st}} Y$ *if, and only if,* (1.A.7) *holds for all increasing functions* ϕ *for which the expectations exist.*

The expressions $\int_x^\infty P\{X > y\}\mathrm{d}y$ and $\int_x^\infty P\{Y > y\}\mathrm{d}y$ are used extensively in Chapters 2, 3, and 4. It is of interest to note that $X \leq_{\text{st}} Y$ if, and only if,

$$\int_x^\infty P\{Y > y\}\mathrm{d}u - \int_x^\infty P\{X > y\}\mathrm{d}y \quad \text{is decreasing in } x \in (-\infty, \infty). \tag{1.A.8}$$

If X and Y are discrete random variables taking on values in $\mathbb{N}$, then we have the following. Let $p_i = P\{X = i\}$ and $q_i = P\{Y = i\}$, $i \in \mathbb{N}$. Then $X \leq_{\text{st}} Y$ if, and only if,

$$\sum_{j=-\infty}^{i} p_j \geq \sum_{j=-\infty}^{i} q_j, \quad i \in \mathbb{N},$$

or, equivalently, $X \leq_{\text{st}} Y$ if, and only if,

$$\sum_{j=i}^{\infty} p_j \leq \sum_{j=i}^{\infty} q_j, \quad i \in \mathbb{N}.$$

1.A.2 A characterization by construction on the same probability space

An important characterization of the usual stochastic order is the following theorem (here $=_{\text{st}}$ denotes equality in law).

Theorem 1.A.1. *Two random variables X and Y satisfy $X \le_{\rm st} Y$ if, and only if, there exist two random variables $\hat{X}$ and $\hat{Y}$, defined on the same probability space, such that*

$$\hat{X} =_{\rm st} X, \tag{1.A.9}$$

$$\hat{Y} =_{\rm st} Y, \tag{1.A.10}$$

and

$$P\{\hat{X} \le \hat{Y}\} = 1. \tag{1.A.11}$$

Proof. Obviously (1.A.9), (1.A.10), and (1.A.11) imply that $X \le_{\rm st} Y$. In order to prove the necessity part of Theorem 1.A.1, let F and G be, respectively, the distribution functions of X and Y, and let F^{-1} and G^{-1} be the corresponding right continuous inverses (see Note on page 1). Define $\hat{X} = F^{-1}(U)$ and $\hat{Y} = G^{-1}(U)$ where U is a uniform $[0,1]$ random variable. Then it is easy to see that $\hat{X}$ and $\hat{Y}$ satisfy (1.A.9) and (1.A.10). From (1.A.2) it is seen that (1.A.11) also holds. □

Theorem 1.A.1 is a special case of a more general result that is stated in Section 6.B.2.

From (1.A.2) and Theorem 1.A.1 it follows that the random variables X and Y, with the respective distribution functions F and G, satisfy $X \le_{\rm st} Y$ if, and only if,

$$F^{-1}(u) \le G^{-1}(u), \quad \text{for all } u \in (0,1). \tag{1.A.12}$$

Another way of restating Theorem 1.A.1 is the following. We omit the obvious proof of it.

Theorem 1.A.2. *Two random variables X and Y satisfy $X \le_{\rm st} Y$ if, and only if, there exist a random variable Z and functions ψ_1 and ψ_2 such that $\psi_1(z) \le \psi_2(z)$ for all z and $X =_{\rm st} \psi_1(Z)$ and $Y =_{\rm st} \psi_2(Z)$.*

In some applications, when the random variables X and Y are such that $X \le_{\rm st} Y$, one may wish to construct a $\hat{Y}$ $[\hat{X}]$ on the probability space on which X $[Y]$ is defined, such that $\hat{Y} =_{\rm st} Y$ and $P\{X \le \hat{Y}\} = 1$ $[\hat{X} =_{\rm st} X$ and $P\{\hat{X} \le Y\} = 1]$. This is always possible. Here we will show how this can be done when the distribution function F $[G]$ of X $[Y]$ is absolutely continuous. When this is the case, $F(X)$ $[G(Y)]$ is uniformly distributed on $[0,1]$, and therefore $\hat{Y} = G^{-1}(F(X))$ $[\hat{X} = F^{-1}(G(Y))]$ is the desired construction $\hat{Y}$ $[\hat{X}]$.

1.A.3 Closure properties

Using (1.A.1) through (1.A.11) it is easy to prove each of the following closure results. The following notation will be used: For any random variable Z and an event A, let $[Z|A]$ denote any random variable that has as its distribution the conditional distribution of Z given A.

Theorem 1.A.3. (a) *If* $X \leq_{\rm st} Y$ *and* g *is any increasing [decreasing] function, then* $g(X) \leq_{\rm st} [\geq_{\rm st}]\ g(Y)$.

(b) *Let* $X_1, X_2, \ldots, X_m$ *be a set of independent random variables and let* $Y_1, Y_2, \ldots, Y_m$ *be another set of independent random variables. If* $X_i \leq_{\rm st} Y_i$ *for* $i = 1, 2, \ldots, m$*, then, for any increasing function* $\psi : \mathbb{R}^m \to \mathbb{R}$*, one has*

$$\psi(X_1, X_2, \ldots, X_m) \leq_{\rm st} \psi(Y_1, Y_2, \ldots, Y_m).$$

In particular,

$$\sum_{j=1}^{m} X_j \leq_{\rm st} \sum_{j=1}^{m} Y_j.$$

That is, the usual stochastic order is closed under convolutions.

(c) *Let* $\{X_j,\ j = 1, 2, \ldots\}$ *and* $\{Y_j,\ j = 1, 2, \ldots\}$ *be two sequences of random variables such that* $X_j \to_{\rm st} X$ *and* $Y_j \to_{\rm st} Y$ *as* $j \to \infty$*, where "*$\to_{\rm st}$*" denotes convergence in distribution. If* $X_j \leq_{\rm st} Y_j$*,* $j = 1, 2, \ldots$*, then* $X \leq_{\rm st} Y$.

(d) *Let* X*,* Y*, and* Θ *be random variables such that* $[X \big| \Theta = \theta] \leq_{\rm st} [Y \big| \Theta = \theta]$ *for all* θ *in the support of* Θ*. Then* $X \leq_{\rm st} Y$*. That is, the usual stochastic order is closed under mixtures.*

In the next result and in the sequel we define $\sum_{j=1}^{0} a_j \equiv 0$ for any sequence $\{a_j,\ j = 1, 2, \ldots\}$.

Theorem 1.A.4. *Let* $\{X_j,\ j = 1, 2, \ldots\}$ *be a sequence of nonnegative independent random variables, and let* M *be a nonnegative integer-valued random variable which is independent of the* X_i*'s. Let* $\{Y_j,\ j = 1, 2, \ldots\}$ *be another sequence of nonnegative independent random variables, and let* N *be a nonnegative integer-valued random variable which is independent of the* Y_i*'s. If* $X_i \leq_{\rm st} Y_i$*,* $i = 1, 2, \ldots$*, and if* $M \leq_{\rm st} N$*, then*

$$\sum_{j=1}^{M} X_j \leq_{\rm st} \sum_{j=1}^{N} Y_j.$$

Another related result is given next.

Theorem 1.A.5. *Let* $\{X_j,\ j = 1, 2, \ldots\}$ *be a sequence of nonnegative independent and identically distributed random variables, and let* M *be a positive integer-valued random variable which is independent of the* X_i*'s. Let* $\{Y_j,\ j = 1, 2, \ldots\}$ *be another sequence of independent and identically distributed random variables, and let* N *be a positive integer-valued random variable which is independent of the* Y_i*'s. Suppose that for some positive integer* K *we have that*

$$\sum_{j=1}^{K} X_j \leq_{\rm st} [\geq_{\rm st}]\ Y_1$$

and

$$M \le_{\rm st} [\ge_{\rm st}]\ KN,$$

then

$$\sum_{j=1}^{M} X_j \le_{\rm st} [\ge_{\rm st}] \sum_{j=1}^{N} Y_j.$$

Proof. The assumptions yield

$$\sum_{i=1}^{M} X_i \le_{\rm st} [\ge_{\rm st}] \sum_{i=1}^{KN} X_i = \sum_{i=1}^{N} \sum_{j=K(i-1)+1}^{Ki} X_j \le_{\rm st} [\ge_{\rm st}] \sum_{i=1}^{N} Y_i. \qquad \square$$

Consider now a family of distribution functions $\{G_\theta,\ \theta \in \mathcal{X}\}$ where $\mathcal{X}$ is a subset of the real line $\mathbb{R}$. Let $X(\theta)$ denote a random variable with distribution function G_θ. For any random variable Θ with support in $\mathcal{X}$, and with distribution function F, let us denote by $X(\Theta)$ a random variable with distribution function H given by

$$H(y) = \int_{\mathcal{X}} G_\theta(y) \mathrm{d}F(\theta), \quad y \in \mathbb{R}.$$

The following result is a generalization of both parts (a) and (c) of Theorem 1.A.3.

Theorem 1.A.6. *Consider a family of distribution functions* $\{G_\theta,\ \theta \in \mathcal{X}\}$ *as above. Let* Θ_1 *and* Θ_2 *be two random variables with supports in* $\mathcal{X}$ *and distribution functions* F_1 *and* F_2*, respectively. Let* Y_1 *and* Y_2 *be two random variables such that* $Y_i =_{\rm st} X(\Theta_i)$, $i = 1, 2$*; that is, suppose that the distribution function of* Y_i *is given by*

$$H_i(y) = \int_{\mathcal{X}} G_\theta(y) \mathrm{d}F_i(\theta), \quad y \in \mathbb{R},\ i = 1, 2.$$

If

$$X(\theta) \le_{\rm st} X(\theta') \quad \textit{whenever } \theta \le \theta', \tag{1.A.13}$$

and if

$$\Theta_1 \le_{\rm st} \Theta_2, \tag{1.A.14}$$

then

$$Y_1 \le_{\rm st} Y_2. \tag{1.A.15}$$

Proof. Note that, by (1.A.13), $P\{X(\theta) > y\}$ is increasing in θ for all y. Thus

$$\begin{aligned} P\{Y_1 > y\} &= \int_{\mathcal{X}} P\{X(\theta) > y\} \mathrm{d}F_1(\theta) \\ &\le \int_{\mathcal{X}} P\{X(\theta) > y\} \mathrm{d}F_2(\theta) \\ &= P\{Y_2 > y\}, \quad \text{for all } y, \end{aligned}$$

where the inequality follows from (1.A.14) and (1.A.7). Thus (1.A.15) follows from (1.A.1). $\square$

Note that, using the notation that is introduced below before Theorem 1.A.14, (1.A.13) can be rewritten as $\{X(\theta),\ \theta \in \mathcal{X}\} \in \mathrm{SI}$.

The following example shows an application of Theorem 1.A.6 in the area of Bayesian imperfect repair; a related result is given in Example 1.B.16.

Example 1.A.7. Let Θ_1 and Θ_2 be two random variables with supports in $\mathcal{X} = (0,1]$ and distribution functions F_1 and F_2, respectively. For some survival function $\overline{K}$, define

$$\overline{G}_\theta = \overline{K}^{1-\theta}, \quad \theta \in (0,1],$$

and let $X(\theta)$ have the survival function $\overline{K}^{1-\theta}$. Note that (1.A.13) holds because $\overline{K}^{1-\theta}(y) \le \overline{K}^{1-\theta'}(y)$ for all y whenever $0 < \theta \le \theta' \le 1$. Thus, if $\Theta_1 \le_{\mathrm{st}} \Theta_2$ then Y_i, with survival function $\overline{H}_i$ defined by

$$\overline{H}_i(y) = \int_0^1 \overline{K}^{1-\theta}(y)\mathrm{d}F_i(\theta), \quad y \in \mathbb{R},\ i = 1, 2,$$

satisfy $Y_1 \le_{\mathrm{st}} Y_2$.

1.A.4 Further characterizations and properties

Clearly, if $X \le_{\mathrm{st}} Y$ then $EX \le EY$. However, as the following result shows, if two random variables are ordered in the usual stochastic order and have the same expected values, they must have the same distribution.

Theorem 1.A.8. *If $X \le_{\mathrm{st}} Y$ and if $E[h(X)] = E[h(Y)]$ for some strictly increasing function h, then $X =_{\mathrm{st}} Y$.*

Proof. First we prove the result when $h(x) = x$. Let $\hat{X}$ and $\hat{Y}$ be as in Theorem 1.A.1. If $P\{\hat{X} < \hat{Y}\} > 0$, then $EX = E\hat{X} < E\hat{Y} = EY$, a contradiction to the assumption $EX = EY$. Therefore $X =_{\mathrm{st}} \hat{X} = \hat{Y} =_{\mathrm{st}} Y$. Now let h be some strictly increasing function. Observe that if $X \le_{\mathrm{st}} Y$, then $h(X) \le_{\mathrm{st}} h(Y)$ and therefore from the above result we have that $h(X) =_{\mathrm{st}} h(Y)$. The strict monotonicity of h yields $X =_{\mathrm{st}} Y$. □

Other results that give conditions, involving stochastic orders, which imply stochastic equalities, are given in Theorems 3.A.43, 3.A.60, 4.A.69, 5.A.15, 6.B.19, 6.G.12, 6.G.13, and 7.A.14–7.A.16.

As was mentioned above, if $X \le_{\mathrm{st}} Y$, then $EX \le EY$. It is easy to find counterexamples which show that the converse is false. However, $X \le_{\mathrm{st}} Y$ implies other moment inequalities (for example, $EX^3 \le EY^3$). Thus one may wonder whether $X \le_{\mathrm{st}} Y$ can be characterized by a collection of moment inequalities. Brockett and Kahane [109, Corollary 1] showed that there exist no finite number of moment inequalities which imply $X \le_{\mathrm{st}} Y$. In fact, they showed it for many other stochastic orders that are studied later in this book.

In order to state the next characterization we define the following class of bivariate functions:

$$\mathcal{G}_{\mathrm{st}} = \{\phi : \mathbb{R}^2 \to \mathbb{R} : \phi(x,y) \text{ is increasing in } x \text{ and decreasing in } y\}.$$

Theorem 1.A.9. *Let X and Y be independent random variables. Then $X \le_{\rm st} Y$ if, and only if,*

$$\phi(X,Y) \le_{\rm st} \phi(Y,X) \quad \text{for all } \phi \in \mathcal{G}_{\rm st}. \tag{1.A.16}$$

Proof. Suppose that (1.A.16) holds. The function ϕ defined by $\phi(x,y) \equiv x$ belongs to $\mathcal{G}_{\rm st}$. Therefore $X \le_{\rm st} Y$.

In order to prove the "only if" part, suppose that $X \le_{\rm st} Y$. Let $\phi \in \mathcal{G}_{\rm st}$ and define $\psi(x,y) = \phi(x,-y)$. Then ψ is increasing on $\mathbb{R}^2$. Since X and Y are independent it follows that X and $-Y$ are independent and also that $-X$ and Y are independent. Since $X \le_{\rm st} Y$ it follows (for example, from Theorem 1.A.1) that $-Y \le_{\rm st} -X$. Therefore, by Theorem 1.A.3(b), we have

$$\psi(X,-Y) \le_{\rm st} \psi(Y,-X),$$

that is,

$$\phi(X,Y) \le_{\rm st} \phi(Y,X). \qquad \square$$

The next result is a similar characterization. In order to state it we need the following notation: Let ϕ_1 and ϕ_2 be two bivariate functions. Denote $\Delta\phi_{21}(x,y) = \phi_2(x,y) - \phi_1(x,y)$. The proof of the following theorem is omitted.

Theorem 1.A.10. *Let X and Y be two independent random variables. Then $X \le_{\rm st} Y$ if, and only if,*

$$E\phi_1(X,Y) \le E\phi_2(X,Y)$$

for all ϕ_1 and ϕ_2 which satisfy that, for each y, $\Delta\phi_{21}(x,y)$ decreases in x on $\{x \le y\}$; for each x, $\Delta\phi_{21}(x,y)$ increases in y on $\{y \ge x\}$; and $\Delta\phi_{21}(x,y) \ge -\Delta\phi_{21}(y,x)$ whenever $x \le y$.

Another similar characterization is given in Theorem 4.A.36.

Let X and Y be two random variables with distribution functions F and G, respectively. Let $\mathcal{M}(F,G)$ denote the Fréchet class of bivariate distributions with fixed marginals F and G. Abusing notation we write $(\hat{X},\hat{Y}) \in \mathcal{M}(F,G)$ to mean that the jointly distributed random variables $\hat{X}$ and $\hat{Y}$ have the marginal distribution functions F and G, respectively. The Fortret-Mourier-Wasserstein distance between the finite mean random variables X and Y is defined by

$$d(X,Y) = \inf_{(\hat{X},\hat{Y}) \in \mathcal{M}(F,G)} \{E|\hat{Y} - \hat{X}|\}. \tag{1.A.17}$$

Theorem 1.A.11. *Let X and Y be two finite mean random variables such that $EX \le EY$. Then $X \le_{\rm st} Y$ if, and only if, $d(X,Y) = EY - EX$.*

Proof. Suppose that $d(X,Y) = EY - EX$. The infimum in (1.A.17) is attained for some $(\hat{X},\hat{Y})$, and we have $E|\hat{Y} - \hat{X}| = E(\hat{Y} - \hat{X})$. Therefore $P\{\hat{X} \le \hat{Y}\} = 1$, and from Theorem 1.A.1 it follows that $X \le_{\rm st} Y$.

Conversely, suppose that $X \le_{\rm st} Y$. Let $\hat{X}$ and $\hat{Y}$ be as in Theorem 1.A.1. Then, for any $(X',Y') \in \mathcal{M}(F,G)$ we have that $E|Y' - X'| \ge |EY' - EX'| = E\hat{Y} - E\hat{X}$. Therefore $d(X,Y) = EY - EX$. $\square$

A simple sufficient condition which implies the usual stochastic order is described next. The following notation will be used. Let $a(x)$ be defined on I, where I is a subset of the real line. The number of sign changes of a in I is defined by

$$S^-(a) = \sup S^-[a(x_1), a(x_2), \ldots, a(x_m)], \tag{1.A.18}$$

where $S^-(y_1, y_2, \ldots, y_m)$ is the number of sign changes of the indicated sequence, zero terms being discarded, and the supremum in (1.A.18) is extended over all sets $x_1 < x_2 < \cdots < x_m$ such that $x_i \in I$ and $m < \infty$. The proof of the next theorem is simple and therefore it is omitted.

Theorem 1.A.12. *Let X and Y be two random variables with (discrete or continuous) density functions f and g, respectively. If*

$$S^-(g - f) = 1 \quad \text{and the sign sequence is } -, +,$$

then $X \leq_{\text{st}} Y$.

Let X_1 be a nonnegative random variable with distribution function F_1 and survival function $\overline{F}_1 \equiv 1 - F_1$. Define the Laplace transform of X_1 by

$$\varphi_{X_1}(\lambda) = \int_0^\infty \mathrm{e}^{-\lambda x} \mathrm{d}F_1(x), \quad \lambda > 0,$$

and denote

$$\overline{a}_\lambda^{X_1}(n) = \frac{(-1)^n}{n!} \frac{\mathrm{d}^n}{\mathrm{d}\lambda^n} \left[\frac{1 - \varphi_{X_1}(\lambda)}{\lambda} \right], \quad n \geq 0, \ \lambda > 0,$$

and

$$\overline{\alpha}_\lambda^{X_1}(n) = \lambda^n \overline{a}_\lambda^{X_1}(n-1), \quad n \geq 1, \ \lambda > 0.$$

Similarly, for a nonnegative random variable X_2 with distribution function F_2 and survival function $\overline{F}_2 \equiv 1 - F_2$, define $\overline{\alpha}_\lambda^{X_2}(n)$. It can be shown that $\overline{\alpha}_\lambda^{X_1}$ and $\overline{\alpha}_\lambda^{X_2}$ are discrete survival functions (see the proof of the next theorem); denote the corresponding discrete random variables by $N_\lambda(X_1)$ and $N_\lambda(X_2)$. The following result gives a Laplace transform characterization of the order $\leq_{\text{st}}$.

Theorem 1.A.13. *Let X_1 and X_2 be two nonnegative random variables, and let $N_\lambda(X_1)$ and $N_\lambda(X_2)$ be as described above. Then*

$$X_1 \leq_{\text{st}} X_2 \Longleftrightarrow N_\lambda(X_1) \leq_{\text{st}} N_\lambda(X_2) \quad \text{for all } \lambda > 0.$$

Proof. First suppose that $X_1 \leq_{\text{st}} X_2$. Select a $\lambda > 0$. Let $Z_1, Z_2, \ldots,$ be independent exponential random variables with mean $1/\lambda$. It can be shown that $\overline{\alpha}_\lambda^{X_1}(n) = P\{\sum_{i=1}^n Z_i \leq X_1\}$ and that $\overline{\alpha}_\lambda^{X_2}(n) = P\{\sum_{i=1}^n Z_i \leq X_2\}$. It thus follows that $N_\lambda(X_1) \leq_{\text{st}} N_\lambda(X_2)$.

Now suppose that $N_\lambda(X_1) \leq_{\text{st}} N_\lambda(X_2)$ for all $\lambda > 0$. Select an $x > 0$. Thus $\overline{\alpha}_{n/x}^{X_1}(n) \leq \overline{\alpha}_{n/x}^{X_2}(n)$. Letting $n \to \infty$, one obtains $\overline{F}_1(x) \leq \overline{F}_2(x)$ for all continuity points x of F_1 and F_2. Therefore, $X_1 \leq_{\text{st}} X_2$ by (1.A.1). □

The implication $\Longrightarrow$ in Theorem 1.A.13 can be generalized as follows. A family of random variables $\{Z(\theta),\ \theta \in \Theta\}$ (Θ is a subset of the real line) is said to be stochastically increasing in the usual stochastic order (denoted by $\{Z(\theta),\ \theta \in \Theta\} \in \text{SI}$) if $Z(\theta) \le_{\text{st}} Z(\theta')$ whenever $\theta \le \theta'$. Recall from Theorem 1.A.3(a) that if $X_1 \le_{\text{st}} X_2$, then $g(X_1) \le_{\text{st}} g(X_2)$ for any increasing function g. The following result gives a stochastic generalization of this fact.

Theorem 1.A.14. *If $\{Z(\theta),\ \theta \in \Theta\} \in \text{SI}$ and if $X_1 \le_{\text{st}} X_2$, where X_k and $Z(\theta)$ are independent for $k = 1, 2$ and $\theta \in \Theta$, then $Z(X_1) \le_{\text{st}} Z(X_2)$.*

Note that Theorem 1.A.14 is a restatement of Theorem 1.A.6.

Let X be a random variable and denote by $X_{(-\infty,a]}$ the truncation of X at a, that is, $X_{(-\infty,a]}$ has as its distribution the conditional distribution of X given that $X \le a$. $X_{(a,\infty)}$ is similarly defined. It is simple to prove the following result. Results that are stronger than this are contained in Theorems 1.B.20, 1.B.55, and 1.C.27.

Theorem 1.A.15. *Let X be any random variable. Then $X_{(-\infty,a]}$ and $X_{(a,\infty)}$ are increasing in a in the sense of the usual stochastic order.*

An interesting example in which truncated random variables are compared is the following.

Example 1.A.16. Let $X^{(1)}, X^{(2)}, \dots, X^{(n)}$ be independent and identically distributed random variables. For a fixed t, let $X^{(1)}_{(t,\infty)}, X^{(2)}_{(t,\infty)}, \dots, X^{(n)}_{(t,\infty)}$ be the corresponding truncations, and assume that they are also independent and identically distributed. Then

$$\left(\max\left\{X^{(1)}, X^{(2)}, \dots, X^{(n)}\right\}\right)_{(t,\infty)} \le_{\text{st}} \max\left\{X^{(1)}_{(t,\infty)}, X^{(2)}_{(t,\infty)}, \dots, X^{(n)}_{(t,\infty)}\right\},$$

where $\left(\max\left\{X^{(1)}, X^{(2)}, \dots, X^{(n)}\right\}\right)_{(t,\infty)}$ denotes the corresponding truncation of $\max\left\{X^{(1)}, X^{(2)}, \dots, X^{(n)}\right\}$. The proof consists of a straightforward verification of (1.A.2) for the compared random variables.

Let ϕ_1 and ϕ_2 be two functions that satisfy $\phi_1(x) \le \phi_2(x)$ for all $x \in \mathbb{R}$, and let X be a random variable. Then, clearly, $\phi_1(X) \le \phi_2(X)$ almost surely. From Theorem 1.A.1 we thus obtain the following result.

Theorem 1.A.17. *Let X be a random variable and let ϕ_1 and ϕ_2 be two functions that satisfy $\phi_1(x) \le \phi_2(x)$ for all $x \in \mathbb{R}$. Then*

$$\phi_1(X) \le_{\text{st}} \phi_2(X).$$

In particular, if ϕ is a function that satisfies $x \le [\ge]\ \phi(x)$ for all $x \in \mathbb{R}$, then $X \le_{\text{st}} [\ge_{\text{st}}]\ \phi(X)$.

Remark 1.A.18. The set of all distribution functions on $\mathbb{R}$ is a lattice with respect to the order $\le_{\text{st}}$. That is, if X and Y are random variables with distributions F and G, then there exist random variables Z and W such that $Z \le_{\text{st}} X$, $Z \le_{\text{st}} Y$, $W \ge_{\text{st}} X$, and $W \ge_{\text{st}} Y$. Explicitly, Z has the survival function $\min\{\overline{F}, \overline{G}\}$ and W has the survival function $\max\{\overline{F}, \overline{G}\}$.

The next four theorems give conditions under which the corresponding spacings are ordered according to the usual stochastic order. Let $X_1, X_2, \ldots, X_m$ be any random variables with the corresponding order statistics $X_{(1)} \le X_{(2)} \le \cdots \le X_{(m)}$. Define the corresponding spacings by $U_{(i)} = X_{(i)} - X_{(i-1)}$, $i = 2, 3, \ldots, m$. When the dependence on m is to be emphasized, we will denote the spacings by $U_{(i:m)}$.

Theorem 1.A.19. *Let $X_1, X_2, \ldots, X_m, X_{m+1}$ be independent and identically distributed* IFR (DFR) *random variables. Then*

$$(m-i+1)U_{(i:m)} \ge_{\rm st} [\le_{\rm st}]\ (m-i)U_{(i+1:m)}, \quad i = 2, 3, \ldots, m-1,$$

and

$$(m-i+2)U_{(i:m+1)} \ge_{\rm st} [\le_{\rm st}]\ (m-i+1)U_{(i:m)}, \quad i = 2, 3, \ldots, m.$$

The proof of Theorem 1.A.19 is not given here. A stronger version of the DFR part of Theorem 1.A.19 is given in Theorem 1.B.31. Some of the conclusions of Theorem 1.A.19 can be obtained under different conditions. These are stated in the next two theorems. Again, the proofs are not given. In the next two theorems we take $X_{(0)} \equiv 0$, and thus $U_{(1)} = X_{(1)}$. For the following theorem recall from page 1 the definition of Schur concavity.

Theorem 1.A.20. *Let $X_1, X_2, \ldots, X_m$ be nonnegative random variables with an absolutely continuous joint distribution function. If the joint density function of $X_1, X_2, \ldots, X_m$ is Schur concave (Schur convex), then*

$$(m-i+1)U_{(i:m)} \ge_{\rm st} [\le_{\rm st}]\ (m-i)U_{(i+1:m)}, \quad i = 1, 2, \ldots, m-1.$$

Theorem 1.A.21. *Let $X_1, X_2, \ldots, X_m$ be independent exponential random variables with possibly different parameters. Then*

$$(m-i+1)U_{(i:m)} \le_{\rm st} (m-i)U_{(i+1:m)}, \quad i = 1, 2, \ldots, m-1.$$

Theorem 1.A.22. *Let $X_1, X_2, \ldots, X_m$ be independent and identically distributed random variables with a finite support, and with an increasing [decreasing] density function over that support. Then*

$$U_{(i:m)} \ge_{\rm st} [\le_{\rm st}]\ U_{(i+1:m)}, \quad i = 2, 3, \ldots, m-1.$$

The proof of Theorem 1.A.22 uses the likelihood ratio order, and therefore it is deferred to Section 1.C, Remark 1.C.3.

Note that any absolutely continuous DFR random variable has a decreasing density function. Thus we see that the assumption in the DFR part of Theorem 1.A.19 is stronger than the assumption in the decreasing part of Theorem 1.A.22, but the conclusion in the DFR part of Theorem 1.A.19 is stronger than the conclusion in the decreasing part of Theorem 1.A.22. It is of interest to compare Theorems 1.A.19–1.A.22 with Theorems 1.B.31 and 1.C.42.

From Theorem 1.A.1 it is obvious that if $X_{(1)} \le X_{(2)} \le \cdots \le X_{(m)}$ are the order statistics corresponding to the random variables $X_1, X_2, \ldots, X_m$, then $X_{(1)} \le_{\mathrm{st}} X_{(2)} \le_{\mathrm{st}} \cdots \le_{\mathrm{st}} X_{(m)}$. Now let $X_{(1)} \le X_{(2)} \le \cdots \le X_{(m)}$ be the order statistics corresponding to the random variables $X_1, X_2, \ldots, X_m$, and let $Y_{(1)} \le Y_{(2)} \le \cdots \le Y_{(m)}$ be the order statistics corresponding to the random variables $Y_1, Y_2, \ldots, Y_m$. As usual, for any distribution function F, we let $\overline{F} \equiv 1 - F$ denote the corresponding survival function.

Theorem 1.A.23. (a) *Let $X_1, X_2, \ldots, X_m$ be independent random variables with distribution functions $F_1, F_2, \ldots, F_m$, respectively. Let $Y_1, Y_2, \ldots, Y_m$ be independent and identically distributed random variables with a common distribution function G. Then $X_{(i)} \le_{\mathrm{st}} Y_{(i)}$ for all $i = 1, 2, \ldots, m$ if, and only if,*

$$\prod_{j=1}^{m} F_j(x) \ge G^m(x) \quad \textit{for all } x;$$

that is, if, and only if, $X_{(m)} \le_{\mathrm{st}} Y_{(m)}$.

(b) *Let $X_1, X_2, \ldots, X_m$ be independent random variables with survival functions $\overline{F}_1, \overline{F}_2, \ldots, \overline{F}_m$, respectively. Let $Y_1, Y_2, \ldots, Y_m$ be independent and identically distributed random variables with a common survival function $\overline{G}$. Then $X_{(i)} \ge_{\mathrm{st}} Y_{(i)}$ for all $i = 1, 2, \ldots, m$ if, and only if,*

$$\prod_{j=1}^{m} \overline{F}_j(x) \ge \overline{G}^m(x) \quad \textit{for all } x;$$

that is, if, and only if, $X_{(1)} \ge_{\mathrm{st}} Y_{(1)}$.

The proof of Theorem 1.A.23 is not given here.

More comparisons of order statistics in the usual stochastic order can be found in Theorem 6.B.23 and in Corollary 6.B.24.

The following neat example compares a sum of independent heterogeneous exponential random variables with an Erlang random variable; it is of interest to compare it with Examples 1.B.5 and 1.C.49. We do not give the proof here.

Example 1.A.24. Let X_i be an exponential random variable with mean $\lambda_i^{-1} > 0$, $i = 1, 2, \ldots, m$, and assume that the X_i's are independent. Let Y_i, $i = 1, 2, \ldots, m$, be independent, identically distributed, exponential random variables with mean η^{-1}. Then

$$\sum_{i=1}^{m} X_i \ge_{\mathrm{st}} \sum_{i=1}^{m} Y_i \iff \sqrt[m]{\lambda_1 \lambda_2 \cdots \lambda_m} \le \eta.$$

The next example may be compared with Examples 1.B.6, 1.C.51, and 4.A.45.

Example 1.A.25. Let X_i be a binomial random variable with parameters n_i and p_i, $i = 1, 2, \ldots, m$, and assume that the X_i's are independent. Let Y be a binomial random variable with parameters n and p where $n = \sum_{i=1}^m n_i$. Then

$$\sum_{i=1}^m X_i \geq_{\text{st}} Y \Longleftrightarrow p \leq \sqrt[n]{p_1^{n_1} p_2^{n_2} \cdots p_m^{n_m}},$$

and

$$\sum_{i=1}^m X_i \leq_{\text{st}} Y \Longleftrightarrow 1 - p \leq \sqrt[n]{(1-p_1)^{n_1}(1-p_2)^{n_2} \cdots (1-p_m)^{n_m}}.$$

The following example gives necessary and sufficient conditions for the comparison of normal random variables; it is generalized in Example 6.B.29. See related results in Examples 3.A.51 and 4.A.46.

Example 1.A.26. Let X be a normal random variable with mean μ_X and variance σ_X^2, and let Y be a normal random variable with mean μ_Y and variance σ_Y^2. Then $X \leq_{\text{st}} Y$ if, and only if, $\mu_X \leq \mu_Y$ and $\sigma_X^2 = \sigma_Y^2$.

Example 1.A.27. Let the random variable X have a unimodal density, symmetric about 0. Then

$$(X + a)^2 \leq_{\text{st}} (X + b)^2 \quad \text{whenever } |a| \leq |b|.$$

Example 1.A.28. Let $\boldsymbol{X}$ have a multivariate normal density with mean vector $\mathbf{0}$ and variance-covariance matrix $\boldsymbol{\Sigma}_1$. Let $\boldsymbol{Y}$ have a multivariate normal density with mean vector $\mathbf{0}$ and variance-covariance matrix $\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2$, where $\boldsymbol{\Sigma}_2$ is a nonnegative definite matrix. Then

$$\|\boldsymbol{X}\|^2 \leq_{\text{st}} \|\boldsymbol{Y}\|^2,$$

where $\|\cdot\|$ denotes the Euclidean norm.

The next result involves the total time on test (TTT) transform and the observed TTT random variable. Let F be the distribution function of a nonnegative random variable, and suppose, for simplicity, that 0 is the left endpoint of the support of F. The TTT transform associated with F is defined by

$$H_F^{-1}(u) = \int_0^{F^{-1}(u)} \overline{F}(x)\mathrm{d}x, \quad 0 \leq u \leq 1, \tag{1.A.19}$$

where $\overline{F} \equiv 1 - F$ is the survival function associated with F. The inverse, H_F, of the TTT transform is a distribution function. If the mean $\mu = \int_0^\infty x\mathrm{d}F(x) = \int_0^\infty \overline{F}(x)\mathrm{d}x$ is finite, then H_F has support in $[0, \mu]$. If X has the distribution function F, then let X_{ttt} be any random variable that has the distribution H_F. The random variable X_{ttt} is called the *observed total time on test.*

Theorem 1.A.29. *Let X and Y be two nonnegative random variables. Then*

$$X \leq_{\text{st}} Y \Longrightarrow X_{\text{ttt}} \leq_{\text{st}} Y_{\text{ttt}}.$$

See related results in Theorems 3.B.1, 4.A.44, 4.B.8, 4.B.9, and 4.B.29.

1.A.5 Some properties in reliability theory

Recall from page 1 the definitions of the IFR, DFR, NBU, and NWU properties. The next result characterizes random variables that have these properties by means of the usual stochastic order. The statements in the next theorem follow at once from the definitions. Recall from Section 1.A.3 that for any random variable Z and an event A we denote by $[Z \big| A]$ any random variable that has as its distribution the conditional distribution of Z given A.

Theorem 1.A.30. (a) *The random variable X is* IFR [DFR] *if, and only if,* $[X - t \big| X > t] \geq_{\rm st} [\leq_{\rm st}] [X - t' \big| X > t']$ *whenever* $t \leq t'$.
(b) *The nonnegative random variable X is* NBU [NWU] *if, and only if,* $X \geq_{\rm st} [\leq_{\rm st}] [X - t \big| X > t]$ *for all* $t > 0$.

Note that if X is the lifetime of a device, then $[X - t \big| X > t]$ is the residual life of such a device with age t. Theorem 1.A.30(a), for example, characterizes IFR and DFR random variables by the monotonicity of their residual lives with respect to the order $\leq_{\rm st}$. Theorem 1.A.30 should be compared to Theorem 1.B.38, where a similar characterization is given. Some multivariate analogs of Theorem 1.A.30(a) are used in Section 6.B.6 to introduce some multivariate IFR notions.

For a nonnegative random variable X with a finite mean, let A_X denote the corresponding asymptotic equilibrium age. That is, if the distribution function of X is F, then the distribution function $F_{\rm e}$ of A_X is defined by

$$F_{\rm e}(x) = \frac{1}{EX} \int_0^x \overline{F}(y) {\rm d}y, \quad x \geq 0, \tag{1.A.20}$$

where $\overline{F} \equiv 1 - F$ is the corresponding survival function. Recall from page 1 the definitions of the NBUE and the NWUE properties. The following result is immediate.

Theorem 1.A.31. *The nonnegative random variable X with finite mean is* NBUE [NWUE] *if, and only if,* $X \geq_{\rm st} [\leq_{\rm st}] A_X$.

Another characterization of NBUE random variables is the following. Recall from Section 1.A.4 the definition of the observed total time on test random variable $X_{\rm ttt}$.

Theorem 1.A.32. *Let X be a nonnegative random variable with finite mean μ. Then X is* NBUE *if, and only if,*

$$X_{\rm ttt} \geq_{\rm st} \mathcal{U}(0, \mu),$$

where $\mathcal{U}(0, \mu)$ denotes a uniform random variable on $(0, \mu)$.

Let X be a nonnegative random variable with finite mean and distribution function F, and let A_X be the corresponding asymptotic equilibrium age having the distribution function $F_{\rm e}$ given in (1.A.20). The requirement

$$X \geq_{\rm st} [A_X - t | A_X > t] \quad \text{for all } t \geq 0, \tag{1.A.21}$$

has been used in the literature as a way to define an aging property of the lifetime X. It turns out that this aging property is equivalent to the new better than used in convex ordering (NBUC) notion that is defined in (4.A.31) in Chapter 4.

1.B The Hazard Rate Order

1.B.1 Definition and equivalent conditions

If X is a random variable with an absolutely continuous distribution function F, then the hazard rate of X at t is defined as $r(t) = (\mathrm{d}/\mathrm{d}t)(-\log(1-F(t)))$. The hazard rate can alternatively be expressed as

$$r(t) = \lim_{\Delta t \downarrow 0} \frac{P\{t < X \leq t + \Delta t | X > t\}}{\Delta t} = \frac{f(t)}{\overline{F}(t)}, \quad t \in \mathbb{R}, \tag{1.B.1}$$

where $\overline{F} \equiv 1 - F$ is the survival function and f is the corresponding density function. As can be seen from (1.B.1), the hazard rate $r(t)$ can be thought of as the intensity of failure of a device, with a random lifetime X, at time t. Clearly, the higher the hazard rate is the smaller X should be stochastically. This is the motivation for the order discussed in this section.

Let X and Y be two nonnegative random variables with absolutely continuous distribution functions and with hazard rate functions r and q, respectively, such that

$$r(t) \geq q(t), \quad t \in \mathbb{R}. \tag{1.B.2}$$

Then X is said to be *smaller than Y in the hazard rate order* (denoted as $X \leq_{\rm hr} Y$).

Although the hazard rate order is usually applied to random lifetimes (that is, nonnegative random variables), definition (1.B.2) may also be used to compare more general random variables. In fact, even the absolute continuity, which is required in (1.B.2), is not really needed. It is easy to verify that (1.B.2) holds if, and only if,

$$\frac{\overline{G}(t)}{\overline{F}(t)} \quad \text{increases in } t \in (-\infty, \max(u_X, u_Y)) \tag{1.B.3}$$

(here $a/0$ is taken to be equal to ∞ whenever $a > 0$). Here F denotes the distribution function of X and G denotes the distribution function of Y, and u_X and u_Y denote the corresponding right endpoints of the supports of X and of Y. Equivalently, (1.B.3) can be written as

$$\overline{F}(x)\overline{G}(y) \geq \overline{F}(y)\overline{G}(x) \quad \text{for all } x \leq y. \tag{1.B.4}$$

Thus (1.B.3) or (1.B.4) can be used to define the order $X \le_{\rm hr} Y$ even if X and/or Y do not have absolutely continuous distributions. A useful further condition, which is equivalent to $X \le_{\rm hr} Y$ when X and Y have absolutely continuous distributions with densities f and g, respectively, is the following:

$$\frac{f(x)}{\overline{F}(y)} \ge \frac{g(x)}{\overline{G}(y)} \quad \text{for all } x \le y. \tag{1.B.5}$$

Rewriting (1.B.4) as

$$\frac{\overline{F}(t+s)}{\overline{F}(t)} \le \frac{\overline{G}(t+s)}{\overline{G}(t)} \quad \text{for all } s \ge 0 \text{ and all } t,$$

it is seen that $X \le_{\rm hr} Y$ if, and only if,

$$P\{X - t > s \big| X > t\} \le P\{Y - t > s \big| Y > t\} \quad \text{for all } s \ge 0 \text{ and all } t; \tag{1.B.6}$$

that is, if, and only if, the residual lives of X and Y at time t are ordered in the sense $\le_{\rm st}$ for all t. Equivalently, (1.B.6) can be written as

$$[X \big| X > t] \le_{\rm st} [Y \big| Y > t] \quad \text{for all } t. \tag{1.B.7}$$

Substituting $u = \overline{F}^{-1}(t)$ in (1.B.3) shows that $X \le_{\rm hr} Y$ if, and only if,

$$\frac{\overline{G}\,\overline{F}^{-1}(u)}{u} \ge \frac{\overline{G}\,\overline{F}^{-1}(v)}{v} \quad \text{for all } 0 < u \le v < 1.$$

Simple manipulations show that the latter condition is equivalent to

$$\frac{1 - FG^{-1}(1-u)}{u} \le \frac{1 - FG^{-1}(1-v)}{v} \quad \text{for all } 0 < u \le v < 1. \tag{1.B.8}$$

For discrete random variables that take on values in $\mathbb{N}$ the definition of $\le_{\rm hr}$ can be written in two different ways. Let X and Y be such random variables. We denote $X \le_{\rm hr} Y$ if

$$\frac{P\{X = n\}}{P\{X \ge n\}} \ge \frac{P\{Y = n\}}{P\{Y \ge n\}}, \quad n \in \mathbb{N}. \tag{1.B.9}$$

Equivalently, $X \le_{\rm hr} Y$ if

$$\frac{P\{X = n\}}{P\{X > n\}} \ge \frac{P\{Y = n\}}{P\{Y > n\}}, \quad n \in \mathbb{N}.$$

The discrete analog of (1.B.4) is that (1.B.9) holds if, and only if,

$$P\{X \ge n_1\}P\{Y \ge n_2\} \ge P\{X \ge n_2\}P\{Y \ge n_1\} \quad \text{for all } n_1 \le n_2. \tag{1.B.10}$$

In a similar manner (1.B.3) and (1.B.5) can be modified in the discrete case. Unless stated otherwise, we consider only random variables with absolutely continuous distribution functions in the following sections.