2 Definitions and Basic Properties

In the Introduction, we referred to copulas as "functions that join or couple multivariate distribution functions to their one-dimensional marginal distribution functions" and as "distribution functions whose one-dimensional margins are uniform." But neither of these statements is a definition—hence we will devote this chapter to giving a precise definition of copulas and to examining some of their elementary properties.

But first we present a glimpse of where we are headed. Consider for a moment a pair of random variables X and Y, with distribution functions $F(x) = P[X \le x]$ and $G(y) = P[Y \le y]$, respectively, and a joint distribution function $H(x,y) = P[X \le x, Y \le y]$ (we will review definitions of random variables, distribution functions, and other important topics as needed in the course of this chapter). To each pair of real numbers (x,y) we can associate three numbers: $F(x)$, $G(y)$, and $H(x,y)$. Note that each of these numbers lies in the interval $[0,1]$. In other words, each pair (x,y) of real numbers leads to a point $\left(F(x),G(y)\right)$ in the unit square $[0,1]\times[0,1]$, and this ordered pair in turn corresponds to a number $H(x,y)$ in $[0,1]$. We will show that this correspondence, which assigns the value of the joint distribution function to each ordered pair of values of the individual distribution functions, is indeed a function. Such functions are copulas.

To accomplish what we have outlined above, we need to generalize the notion of "nondecreasing" for univariate functions to a concept applicable to multivariate functions. We begin with some notation and definitions. In Sects. 2.1-2.9, we confine ourselves to the two-dimensional case; in Sect. 2.10, we consider n dimensions.

2.1 Preliminaries

The focus of this section is the notion of a "2-increasing" function—a two-dimensional analog of a nondecreasing function of one variable. But first we need to introduce some notation. We will let $\mathbf{R}$ denote the ordinary real line $(-\infty,\infty)$, $\overline{\mathbf{R}}$ denote the extended real line $[-\infty,\infty]$, and $\overline{\mathbf{R}}^2$ denote the extended real plane $\overline{\mathbf{R}}\times\overline{\mathbf{R}}$. A *rectangle* in $\overline{\mathbf{R}}^2$ is the

Cartesian product B of two closed intervals: $B = [x_1,x_2]\times[y_1,y_2]$. The *vertices* of a rectangle B are the points (x_1,y_1), (x_1,y_2), (x_2,y_1), and (x_2,y_2). The *unit square* $\mathbf{I}^2$ is the product $\mathbf{I}\times\mathbf{I}$ where $\mathbf{I} = [0,1]$. A 2-*place real function* H is a function whose domain, Dom$H$, is a subset of $\overline{\mathbf{R}}^2$ and whose range, Ran$H$, is a subset of $\mathbf{R}$.

Definition 2.1.1. Let S_1 and S_2 be nonempty subsets of $\overline{\mathbf{R}}$, and let H be a two-place real function such that Dom$H$ = $S_1\times S_2$. Let $B = [x_1,x_2]\times[y_1,y_2]$ be a rectangle all of whose vertices are in Dom$H$. Then the *H-volume of B* is given by

$$V_H(B) = H(x_2,y_2) - H(x_2,y_1) - H(x_1,y_2) + H(x_1,y_1). \qquad (2.1.1)$$

Note that if we define the first order differences of H on the rectangle B as

$$\Delta_{x_1}^{x_2} H(x,y) = H(x_2,y) - H(x_1,y) \text{ and } \Delta_{y_1}^{y_2} H(x,y) = H(x,y_2) - H(x,y_1),$$

then the H-volume of a rectangle B is the *second order difference* of H on B,

$$V_H(B) = \Delta_{y_1}^{y_2}\Delta_{x_1}^{x_2} H(x,y).$$

Definition 2.1.2. A 2-place real function H is 2-*increasing* if $V_H(B) \geq 0$ for all rectangles B whose vertices lie in Dom$H$.

When H is 2-increasing, we will occasionally refer to the H-volume of a rectangle B as the *H-measure of B*. Some authors refer to 2-increasing functions as *quasi-monotone*.

We note here that the statement "H is 2-increasing" neither implies nor is implied by the statement "H is nondecreasing in each argument," as the following two examples illustrate. The verifications are elementary, and are left as exercises.

Example 2.1. Let H be the function defined on $\mathbf{I}^2$ by $H(x,y)$ = $\max(x,y)$. Then H is a nondecreasing function of x and of y; however, $V_H(\mathbf{I}^2) = -1$, so that H is not 2-increasing. ■

Example 2.2. Let H be the function defined on $\mathbf{I}^2$ by $H(x,y)$ = $(2x-1)(2y-1)$. Then H is 2-increasing, however it is a decreasing function of x for each y in (0,1/2) and a decreasing function of y for each x in (0,1/2). ■

The following lemmas will be very useful in the next section in establishing the continuity of subcopulas and copulas. The first is a direct consequence of Definitions 2.1.1 and 2.1.2.

Lemma 2.1.3. *Let S_1 and S_2 be nonempty subsets of $\overline{\mathbf{R}}$, and let H be a 2-increasing function with domain $S_1 \times S_2$. Let x_1, x_2 be in S_1 with $x_1 \leq x_2$, and let y_1, y_2 be in S_2 with $y_1 \leq y_2$. Then the function $t \mapsto H(t,y_2) - H(t,y_1)$ is nondecreasing on S_1, and the function $t \mapsto H(x_2,t) - H(x_1,t)$ is nondecreasing on S_2.*

As an immediate application of this lemma, we can show that with an additional hypothesis, a 2-increasing function H is nondecreasing in each argument. Suppose S_1 has a least element a_1 and that S_2 has a least element a_2. We say that a function H from $S_1 \times S_2$ into $\mathbf{R}$ is *grounded* if $H(x,a_2) = 0 = H(a_1,y)$ for all (x,y) in $S_1 \times S_2$. Hence we have

Lemma 2.1.4. *Let S_1 and S_2 be nonempty subsets of $\overline{\mathbf{R}}$, and let H be a grounded 2-increasing function with domain $S_1 \times S_2$. Then H is nondecreasing in each argument.*

Proof. Let a_1, a_2 denote the least elements of S_1, S_2, respectively, and set $x_1 = a_1$, $y_1 = a_2$ in Lemma 2.1.3. □

Now suppose that S_1 has a greatest element b_1 and that S_2 has a greatest element b_2. We then say that a function H from $S_1 \times S_2$ into $\mathbf{R}$ *has margins*, and that the margins of H are the functions F and G given by:

$$\operatorname{Dom} F = S_1, \text{ and } F(x) = H(x, b_2) \text{ for all } x \text{ in } S_1;$$

$$\operatorname{Dom} G = S_2, \text{ and } G(y) = H(b_1, y) \text{ for all } y \text{ in } S_2.$$

Example 2.3. Let H be the function with domain $[-1,1] \times [0,\infty]$ given by

$$H(x,y) = \frac{(x+1)(e^y - 1)}{x + 2e^y - 1}.$$

Then H is grounded because $H(x,0) = 0$ and $H(-1,y) = 0$; and H has margins $F(x)$ and $G(y)$ given by

$$F(x) = H(x,\infty) = (x+1)/2 \text{ and } G(y) = H(1,y) = 1 - e^{-y}. \qquad \blacksquare$$

We close this section with an important lemma concerning grounded 2-increasing functions with margins.

Lemma 2.1.5. *Let S_1 and S_2 be nonempty subsets of $\overline{\mathbf{R}}$, and let H be a grounded 2-increasing function, with margins, whose domain is $S_1 \times S_2$. Let (x_1,y_1) and (x_2,y_2) be any points in $S_1 \times S_2$. Then*

$$\left|H(x_2,y_2) - H(x_1,y_1)\right| \leq \left|F(x_2) - F(x_1)\right| + \left|G(y_2) - G(y_1)\right|.$$

Proof. From the triangle inequality, we have

$$\left|H(x_2,y_2)-H(x_1,y_1)\right| \le \left|H(x_2,y_2)-H(x_1,y_2)\right|+\left|H(x_1,y_2)-H(x_1,y_1)\right|.$$

Now assume $x_1 \le x_2$. Because H is grounded, 2-increasing, and has margins, Lemmas 2.1.3 and 2.1.4 yield $0 \le H(x_2,y_2) - H(x_1,y_2) \le F(x_2)-F(x_1)$. An analogous inequality holds when $x_2 \le x_1$, hence it follows that for any x_1,x_2 in S_1, $\left|H(x_2,y_2)-H(x_1,y_2)\right| \le \left|F(x_2)-F(x_1)\right|$. Similarly for any y_1,y_2 in S_2, $\left|H(x_1,y_2)-H(x_1,y_1)\right| \le \left|G(y_2)-G(y_1)\right|$, which completes the proof. □

2.2 Copulas

We are now in a position to define the functions—copulas—that are the subject of this book. To do so, we first define subcopulas as a certain class of grounded 2-increasing functions with margins; then we define copulas as subcopulas with domain $\mathbf{I}^2$.

Definition 2.2.1. A *two-dimensional subcopula* (or 2-*subcopula*, or briefly, a *subcopula*) is a function C' with the following properties:

1. $\operatorname{Dom} C' = S_1 \times S_2$, where S_1 and S_2 are subsets of $\mathbf{I}$ containing 0 and 1;
2. C' is grounded and 2-increasing;
3. For every u in S_1 and every v in S_2,

$$C'(u,1)=u \text{ and } C'(1,v)=v\,. \tag{2.2.1}$$

Note that for every (u,v) in $\operatorname{Dom} C'$, $0 \le C'(u,v) \le 1$, so that $\operatorname{Ran} C'$ is also a subset of $\mathbf{I}$.

Definition 2.2.2. A *two-dimensional copula* (or 2-*copula*, or briefly, a *copula*) is a 2-subcopula C whose domain is $\mathbf{I}^2$.

Equivalently, a copula is a function C from $\mathbf{I}^2$ to $\mathbf{I}$ with the following properties:

1. For every u, v in $\mathbf{I}$,

$$C(u,0)=0=C(0,v) \tag{2.2.2a}$$

and

$$C(u,1)=u \text{ and } C(1,v)=v\,; \tag{2.2.2b}$$

2. For every u_1,u_2,v_1,v_2 in $\mathbf{I}$ such that $u_1 \le u_2$ and $v_1 \le v_2$,

$$C(u_2,v_2)-C(u_2,v_1)-C(u_1,v_2)+C(u_1,v_1) \ge 0\,. \tag{2.2.3}$$

Because $C(u,v) = V_C([0,u]\times[0,v])$, one can think of $C(u,v)$ as an assignment of a number in **I** to the rectangle $[0,u]\times[0,v]$. Thus (2.2.3) gives an "inclusion-exclusion" type formula for the number assigned by C to each rectangle $[u_1,u_2]\times[v_1,v_2]$ in $\mathbf{I}^2$ and states that the number so assigned must be nonnegative.

The distinction between a subcopula and a copula (the domain) may appear to be a minor one, but it will be rather important in the next section when we discuss Sklar's theorem. In addition, many of the important properties of copulas are actually properties of subcopulas.

Theorem 2.2.3. *Let C' be a subcopula. Then for every (u,v) in* Dom C',

$$\max(u+v-1,0) \le C'(u,v) \le \min(u,v). \tag{2.2.4}$$

Proof. Let (u,v) be an arbitrary point in Dom C'. Now $C'(u,v) \le C'(u,1) = u$ and $C'(u,v) \le C'(1,v) = v$ yield $C'(u,v) \le \min(u,v)$. Furthermore, $V_{C'}([u,1]\times[v,1]) \ge 0$ implies $C'(u,v) \ge u+v-1$, which when combined with $C'(u,v) \ge 0$ yields $C'(u,v) \ge \max(u+v-1,0)$. □

Because every copula is a subcopula, the inequality in the above theorem holds for copulas. Indeed, the bounds in (2.2.4) are themselves copulas (see Exercise 2.2) and are commonly denoted by $M(u,v) = \min(u,v)$ and $W(u,v) = \max(u+v-1,0)$. Thus for every copula C and every (u,v) in $\mathbf{I}^2$,

$$W(u,v) \le C(u,v) \le M(u,v). \tag{2.2.5}$$

Inequality (2.2.5) is the copula version of the *Fréchet-Hoeffding bounds* inequality, which we shall encounter later in terms of distribution functions. We refer to M as the *Fréchet-Hoeffding upper bound* and W as the *Fréchet-Hoeffding lower bound*. A third important copula that we will frequently encounter is the *product copula* $\Pi(u,v) = uv$.

The following theorem, which follows directly from Lemma 2.1.5, establishes the continuity of subcopulas—and hence of copulas—via a Lipschitz condition on $\mathbf{I}^2$.

Theorem 2.2.4. *Let C' be a subcopula. Then for every $(u_1,u_2),(v_1,v_2)$ in* Dom C',

$$\left|C'(u_2,v_2) - C'(u_1,v_1)\right| \le \left|u_2-u_1\right| + \left|v_2-v_1\right|. \tag{2.2.6}$$

Hence C' is uniformly continuous on its domain.

The *sections* of a copula will be employed in the construction of copulas in the next chapter, and will be used in Chapter 5 to provide interpretations of certain dependence properties:

Definition 2.2.5. Let C be a copula, and let a be any number in **I**. The *horizontal section of C at a* is the function from **I** to **I** given by

$t \mapsto C(t,a)$; the *vertical section of* C *at* a is the function from $\mathbf{I}$ to $\mathbf{I}$ given by $t \mapsto C(a,t)$; and the *diagonal section of* C is the function δ_C from $\mathbf{I}$ to $\mathbf{I}$ defined by $\delta_C(t) = C(t,t)$.

The following corollary is an immediate consequence of Lemma 2.1.4 and Theorem 2.2.4.

Corollary 2.2.6. *The horizontal, vertical, and diagonal sections of a copula* C *are all nondecreasing and uniformly continuous on* $\mathbf{I}$.

Various applications of copulas that we will encounter in later chapters involve the shape of the graph of a copula, i.e., the surface $z = C(u,v)$. It follows from Definition 2.2.2 and Theorem 2.2.4 that the graph of any copula is a continuous surface within the unit cube $\mathbf{I}^3$ whose boundary is the skew quadrilateral with vertices (0,0,0), (1,0,0), (1,1,1), and (0,1,0); and from Theorem 2.2.3 that this graph lies between the graphs of the Fréchet-Hoeffding bounds, i.e., the surfaces $z = M(u,v)$ and $z = W(u,v)$. In Fig. 2.1 we present the graphs of the copulas M and W, as well as the graph of Π, a portion of the hyperbolic paraboloid $z = uv$.

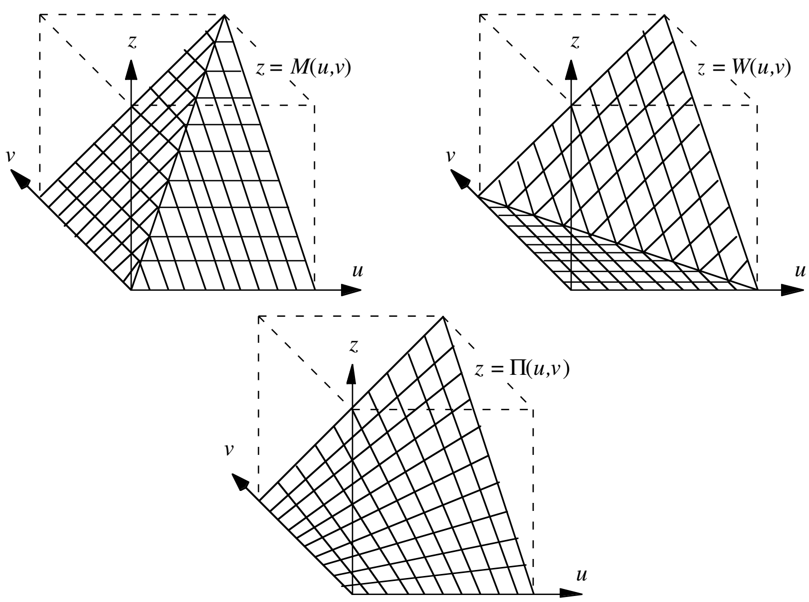

Fig. 2.1. Graphs of the copulas M, Π, and W

A simple but useful way to present the graph of a copula is with a *contour diagram* (Conway 1979), that is, with graphs of its *level sets*—the sets in $\mathbf{I}^2$ given by $C(u,v)$ = a constant, for selected constants in $\mathbf{I}$. In Fig. 2.2 we present the contour diagrams of the copulas M, Π, and W. Note that the points $(t,1)$ and $(1,t)$ are each members of the level set corresponding to the constant t. Hence we do not need to label the level sets in the diagram, as the boundary conditions $C(1,t) = t = C(t,1)$ readily provide the constant for each level set.

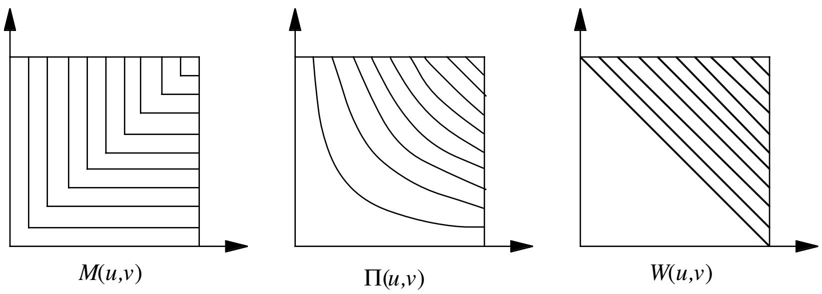

Fig. 2.2. Contour diagrams of the copulas M, Π, and W

Also note that, given any copula C, it follows from (2.2.5) that for a given t in $\mathbf{I}$ the graph of the level set $\{(u,v) \in \mathbf{I}^2 | C(u,v) = t\}$ must lie in the shaded triangle in Fig. 2.3, whose boundaries are the level sets determined by $M(u,v) = t$ and $W(u,v) = t$.

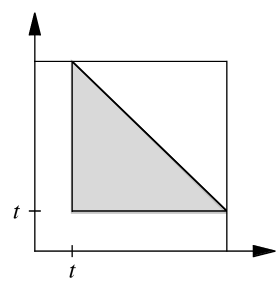

Fig. 2.3. The region that contains the level set $\{(u,v) \in \mathbf{I}^2 | C(u,v) = t\}$

We conclude this section with the two theorems concerning the partial derivatives of copulas. The word "almost" is used in the sense of Lebesgue measure.

Theorem 2.2.7. *Let C be a copula. For any v in $\mathbf{I}$, the partial derivative $\partial C(u,v)/\partial u$ exists for almost all u, and for such v and u,*

$$0 \le \frac{\partial}{\partial u} C(u,v) \le 1. \tag{2.2.7}$$

Similarly, for any u in $\mathbf{I}$, the partial derivative $\partial C(u,v)/\partial v$ exists for almost all v, and for such u and v,

$$0 \le \frac{\partial}{\partial v} C(u,v) \le 1. \tag{2.2.8}$$

Furthermore, the functions $u \mapsto \partial C(u,v)/\partial v$ *and* $v \mapsto \partial C(u,v)/\partial u$ *are defined and nondecreasing almost everywhere on* $\mathbf{I}$.

Proof. The existence of the partial derivatives $\partial C(u,v)/\partial u$ and $\partial C(u,v)/\partial v$ is immediate because monotone functions (here the horizontal and vertical sections of the copula) are differentiable almost everywhere. Inequalities (2.2.7) and (2.2.8) follow from (2.2.6) by setting $v_1 = v_2$ and $u_1 = u_2$, respectively. If $v_1 \le v_2$, then, from Lemma 2.1.3, the function $u \mapsto C(u,v_2) - C(u,v_1)$ is nondecreasing. Hence $\partial\big(C(u,v_2) - C(u,v_1)\big)/\partial u$ is defined and nonnegative almost everywhere on $\mathbf{I}$, from which it follows that $v \mapsto \partial C(u,v)/\partial u$ is defined and nondecreasing almost everywhere on $\mathbf{I}$. A similar result holds for $u \mapsto \partial C(u,v)/\partial v$. □

Theorem 2.2.8. *Let C be a copula. If* $\partial C(u,v)/\partial v$ *and* $\partial^2 C(u,v)/\partial u \partial v$ *are continous on* $\mathbf{I}^2$ *and* $\partial C(u,v)/\partial u$ *exists for all* $u \in (0,1)$ *when* $v = 0$, *then* $\partial C(u,v)/\partial u$ *and* $\partial^2 C(u,v)/\partial v \partial u$ *exist in* $(0,1)^2$ *and* $\partial^2 C(u,v)/\partial u \partial v = \partial^2 C(u,v)/\partial v \partial u$.

Proof. See (Seeley 1961).

Exercises

2.1 Verify the statements in Examples 2.1 and 2.2.

2.2 Show that $M(u,v) = \min(u,v)$, $W(u,v) = \max(u+v-1,0)$, and $\Pi(u,v) = uv$ are indeed copulas.

2.3 (a) Let C_0 and C_1 be copulas, and let θ be any number in $\mathbf{I}$. Show that the weighted arithmetic mean $(1-\theta)C_0 + \theta C_1$ is also a copula. Hence conclude that any convex linear combination of copulas is a copula.
(b) Show that the geometric mean of two copulas may fail to be a copula. [Hint: Let C be the geometric mean of Π and W, and show that the C-volume of the rectangle $[1/2,3/4]\times[1/2,3/4]$ is negative.]

2.4 *The Fréchet and Mardia families of copulas*.
(a) Let α, β be in $\mathbf{I}$ with $\alpha + \beta \le 1$. Set

$$C_{\alpha,\beta}(u,v) = \alpha M(u,v) + (1-\alpha-\beta)\Pi(u,v) + \beta W(u,v).$$

Show that $C_{\alpha,\beta}$ is a copula. A family of copulas that includes M, Π, and W is called *comprehensive*. This two-parameter comprehensive family is due to Fréchet (1958).

(b) Let θ be in [–1,1], and set

$$C_\theta(u,v) = \frac{\theta^2(1+\theta)}{2} M(u,v) + (1-\theta^2)\Pi(u,v) + \frac{\theta^2(1-\theta)}{2} W(u,v). \quad (2.2.9)$$

Show that C_θ is a copula. This one-parameter comprehensive family is due to Mardia (1970).

2.5 *The Cuadras-Augé family of copulas*. Let θ be in $\mathbf{I}$, and set

$$C_\theta(u,v) = [\min(u,v)]^\theta [uv]^{1-\theta} = \begin{cases} uv^{1-\theta}, & u \le v, \\ u^{1-\theta}v, & u \ge v. \end{cases} \quad (2.2.10)$$

Show that C_θ is a copula. Note that $C_0 = \Pi$ and $C_1 = M$. This family (weighted geometric means of M and Π) is due to Cuadras and Augé (1981).

2.6 Let C be a copula, and let (a,b) be any point in $\mathbf{I}^2$. For (u,v) in $\mathbf{I}^2$, define

$$K_{a,b}(u,v) = V_C\big([a(1-u), u+a(1-u)] \times [b(1-v), v+b(1-v)]\big).$$

Show that $K_{a,b}$ is a copula. Note that $K_{0,0}(u,v) = C(u,v)$. Several special cases will be of interest in Sects. 2.4, 2.7, and 6.4, namely:

$$K_{0,1}(u,v) = u - C(u,1-v),$$
$$K_{1,0}(u,v) = v - C(1-u,v), \text{ and}$$
$$K_{1,1}(u,v) = u + v - 1 + C(1-u,1-v).$$

2.7 Let f be a function from $\mathbf{I}^2$ into $\mathbf{I}$ which is nondecreasing in each variable and has margins given by $f(t,1) = t = f(1,t)$ for all t in $\mathbf{I}$. Prove that f is grounded.

2.8 (a) Show that for any copula C, $\max(2t-1,0) \le \delta_C(t) \le t$ for all t in $\mathbf{I}$.

(b) Show that $\delta_C(t) = \delta_M(t)$ for all t in $\mathbf{I}$ implies $C = M$.

(c) Show $\delta_C(t) = \delta_W(t)$ for all t in $\mathbf{I}$ does not imply that $C = W$.

2.9 The *secondary diagonal section* of C is given by $C(t,1-t)$. Show that $C(t,1-t) = 0$ for all t in $\mathbf{I}$ implies $C = W$.

2.10 Let t be in $[0,1)$, and let C_t be the function from $\mathbf{I}^2$ into $\mathbf{I}$ given by

$$C_t(u,v) = \begin{cases} \max(u+v-1,t), & (u,v) \in [t,1]^2, \\ \min(u,v), & \text{otherwise.} \end{cases}$$

(a) Show that C_t is a copula.

(b) Show that the level set $\left\{(u,v) \in \mathbf{I}^2 \middle| C_t(u,v) = t\right\}$ is the set of points in the triangle with vertices $(t,1)$, $(1,t)$, and (t,t), that is, the shaded region in Fig. 2.3. The copula in this exercise illustrates why the term "level set" is preferable to "level curve" for some copulas.

2.11 This exercise shows that the 2-increasing condition (2.2.3) for copulas is not a consequence of simpler properties. Let Q be the function from $\mathbf{I}^2$ into $\mathbf{I}$ given by

$$Q(u,v) = \begin{cases} \min\left(u,v,\dfrac{1}{3},u+v-\dfrac{2}{3}\right), & \dfrac{2}{3} \le u+v \le \dfrac{4}{3}, \\ \max(u+v-1,0), & \text{otherwise;} \end{cases}$$

that is, Q has the values given in Fig. 2.4 in the various parts of $\mathbf{I}^2$.
(a) Show that for every u,v in $\mathbf{I}$, $Q(u,0) = 0 = Q(0,v)$, $Q(u,1) = u$ and $Q(1,v) = v$; $W(u,v) \le Q(u,v) \le M(u,v)$; and that Q is continuous, satisfies the Lipschitz condition (2.2.6), and is nondecreasing in each variable.
(b) Show that Q fails to be 2-increasing, and hence is not a copula. [Hint: consider the Q-volume of the rectangle $[1/3,2/3]^2$.]

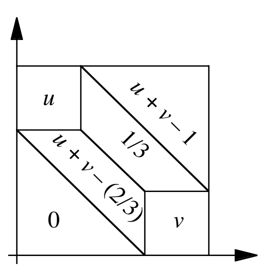

Fig. 2.4. The function Q in Exercise 2.11

2.3 Sklar's Theorem

The theorem in the title of this section is central to the theory of copulas and is the foundation of many, if not most, of the applications of that theory to statistics. Sklar's theorem elucidates the role that copulas play in the relationship between multivariate distribution functions and their univariate margins. Thus we begin this section with a short discussion of distribution functions.

Definition 2.3.1. A *distribution function* is a function F with domain $\overline{\mathbf{R}}$ such that
1. F is nondecreasing,
2. $F(-\infty) = 0$ and $F(\infty) = 1$.

Example 2.4. For any number a in $\mathbf{R}$, the *unit step at* a is the distribution function ε_a given by

$$\varepsilon_a(x) = \begin{cases} 0, & x \in [-\infty, a), \\ 1, & x \in [a, \infty]; \end{cases}$$

and for any numbers a,b in $\mathbf{R}$ with $a < b$, the *uniform distribution on* $[a,b]$ is the distribution function U_{ab} given by

$$U_{ab}(x) = \begin{cases} 0, & x \in [-\infty, a), \\ \dfrac{x-a}{b-a}, & x \in [a, b], \\ 1, & x \in (b, \infty]. \end{cases}$$ ■

Definition 2.3.2. A *joint distribution function* is a function H with domain $\overline{\mathbf{R}}^2$ such that
1. H is 2-increasing,
2. $H(x,-\infty) = H(-\infty,y) = 0$, and $H(\infty,\infty) = 1$.

Thus H is grounded, and because Dom$H$ = $\overline{\mathbf{R}}^2$, H has margins F and G given by $F(x) = H(x,\infty)$ and $G(y) = H(\infty,y)$. By virtue of Corollary 2.2.6, F and G are distribution functions.

Example 2.5. Let H be the function with domain $\overline{\mathbf{R}}^2$ given by

$$H(x,y) = \begin{cases} \dfrac{(x+1)(e^y-1)}{x+2e^y-1}, & (x,y) \in [-1,1] \times [0,\infty], \\ 1 - e^{-y}, & (x,y) \in (1,\infty] \times [0,\infty], \\ 0, & \text{elsewhere}. \end{cases}$$

It is tedious but elementary to verify that H is 2-increasing and grounded, and that $H(\infty,\infty) = 1$. Hence H is a joint distribution function. The margins of H are the distribution functions F and G given by

$$F = U_{-1,1} \quad \text{and} \quad G(y) = \begin{cases} 0, & y \in [-\infty,0), \\ 1 - e^{-y}, & y \in [0,\infty]. \end{cases}$$

[Cf. Examples 2.3 and 2.4.] ■

Note that there is nothing "probabilistic" in these definitions of distribution functions. Random variables are not mentioned, nor is left-continuity or right-continuity. All the distribution functions of one or of two random variables usually encountered in statistics satisfy either the first or the second of the above definitions. Hence any results we derive for such distribution functions will hold when we discuss random variables, regardless of any additional restrictions that may be imposed.

Theorem 2.3.3. Sklar's theorem. *Let H be a joint distribution function with margins F and G. Then there exists a copula C such that for all x,y in $\overline{\mathbf{R}}$,*

$$H(x,y) = C(F(x),G(y)). \tag{2.3.1}$$

If F and G are continuous, then C is unique; otherwise, C is uniquely determined on $\operatorname{Ran}F \times \operatorname{Ran}G$. *Conversely, if C is a copula and F and G are distribution functions, then the function H defined by* (2.3.1) *is a joint distribution function with margins F and G.*

This theorem first appeared in (Sklar 1959). The name "copula" was chosen to emphasize the manner in which a copula "couples" a joint distribution function to its univariate margins. The argument that we give below is essentially the same as in (Schweizer and Sklar 1974). It requires two lemmas.

Lemma 2.3.4. *Let H be a joint distribution function with margins F and G. Then there exists a unique subcopula C' such that*

1. $\operatorname{Dom}C' = \operatorname{Ran}F \times \operatorname{Ran}G$,
2. *For all x,y in $\overline{\mathbf{R}}$, $H(x,y) = C'(F(x),G(y))$.*

Proof. The joint distribution H satisfies the hypotheses of Lemma 2.1.5 with $S_1 = S_2 = \overline{\mathbf{R}}$. Hence for any points (x_1,y_1) and (x_2,y_2) in $\overline{\mathbf{R}}^2$,

$$\left|H(x_2,y_2) - H(x_1,y_1)\right| \le \left|F(x_2) - F(x_1)\right| + \left|G(y_2) - G(y_1)\right|.$$

It follows that if $F(x_1) = F(x_2)$ and $G(y_1) = G(y_2)$, then $H(x_1,y_1) = H(x_2,y_2)$. Thus the set of ordered pairs

$$\left\{\left(\left(F(x),G(y)\right),H(x,y)\right) \middle| x,y \in \overline{\mathbf{R}}\right\}$$

defines a 2-place real function C' whose domain is $\text{Ran}F \times \text{Ran}G$. That this function is a subcopula follows directly from the properties of H. For instance, to show that (2.2.2) holds, we first note that for each u in $\text{Ran}F$, there is an x in $\overline{\mathbf{R}}$ such that $F(x) = u$. Thus $C'(u,1) = C'(F(x),G(\infty)) = H(x,\infty) = F(x) = u$. Verifications of the other conditions in Definition 2.2.1 are similar. □

Lemma 2.3.5. *Let C' be a subcopula. Then there exists a copula C such that $C(u,v) = C'(u,v)$ for all (u,v) in $\text{Dom}\,C'$; i.e., any subcopula can be extended to a copula. The extension is generally non-unique.*

Proof. Let $\text{Dom}\,C' = S_1 \times S_2$. Using Theorem 2.2.4 and the fact that C' is nondecreasing in each place, we can extend C' by continuity to a function C'' with domain $\overline{S}_1 \times \overline{S}_2$, where $\overline{S}_1$ is the closure of S_1 and $\overline{S}_2$ is the closure of S_2. Clearly C'' is also a subcopula. We next extend C'' to a function C with domain $\mathbf{I}^2$. To this end, let (a,b) be any point in $\mathbf{I}^2$, let a_1 and a_2 be, respectively, the greatest and least elements of $\overline{S}_1$ that satisfy $a_1 \le a \le a_2$; and let b_1 and b_2 be, respectively, the greatest and least elements of $\overline{S}_2$ that satisfy $b_1 \le b \le b_2$. Note that if a is in $\overline{S}_1$, then $a_1 = a = a_2$; and if b is in $\overline{S}_2$, then $b_1 = b = b_2$. Now let

$$\lambda_1 = \begin{cases} (a - a_1)/(a_2 - a_1), & \text{if } a_1 < a_2, \\ 1, & \text{if } a_1 = a_2; \end{cases}$$

$$\mu_1 = \begin{cases} (b - b_1)/(b_2 - b_1), & \text{if } b_1 < b_2, \\ 1, & \text{if } b_1 = b_2; \end{cases}$$

and define

$$\begin{aligned} C(a,b) = {} & (1-\lambda_1)(1-\mu_1)C''(a_1,b_1) + (1-\lambda_1)\mu_1 C''(a_1,b_2) \\ & + \lambda_1(1-\mu_1)C''(a_2,b_1) + \lambda_1\mu_1 C''(a_2,b_2). \end{aligned} \tag{2.3.2}$$

Notice that the interpolation defined in (2.3.2) is linear in each place (what we call *bilinear interpolation*) because λ_1 and μ_1 are linear in a and b, respectively.

It is obvious that $\text{Dom}C = \mathbf{I}^2$, that $C(a,b) = C''(a,b)$ for any (a,b) in $\text{Dom}\,C''$; and that C satisfies (2.2.2a) and (2.2.2b). Hence we only must show that C satisfies (2.2.3). To accomplish this, let (c,d) be another point in $\mathbf{I}^2$ such that $c \ge a$ and $d \ge b$, and let c_1, d_1, c_2, d_2, λ_2, μ_2 be related to c and d as a_1, b_1, a_2, b_2, λ_1, μ_1 are related to a and b. In evaluating $V_C(B)$ for the rectangle $B = [a,c] \times [b,d]$, there will be several cases to consider, depending upon whether or not there is a point in $\overline{S}_1$ strictly between a and c, and whether or not there is a point in $\overline{S}_2$

strictly between b and d. In the simplest of these cases, there is no point in $\overline{S}_1$ strictly between a and c, and no point in $\overline{S}_2$ strictly between b and d, so that $c_1 = a_1$, $c_2 = a_2$, $d_1 = b_1$, and $d_2 = b_2$. Substituting (2.3.2) and the corresponding terms for $C(a,d)$, $C(c,b)$ and $C(c,d)$ into the expression given by (2.1.1) for $V_C(B)$ and simplifying yields

$$V_C(B) = V_C([a,c]\times[b,d]) = (\lambda_2 - \lambda_1)(\mu_2 - \mu_1)V_C([a_1,a_2]\times[b_1,b_2]),$$

from which it follows that $V_C(B) \geq 0$ in this case, as $c \geq a$ and $d \geq b$ imply $\lambda_2 \geq \lambda_1$ and $\mu_2 \geq \mu_1$.

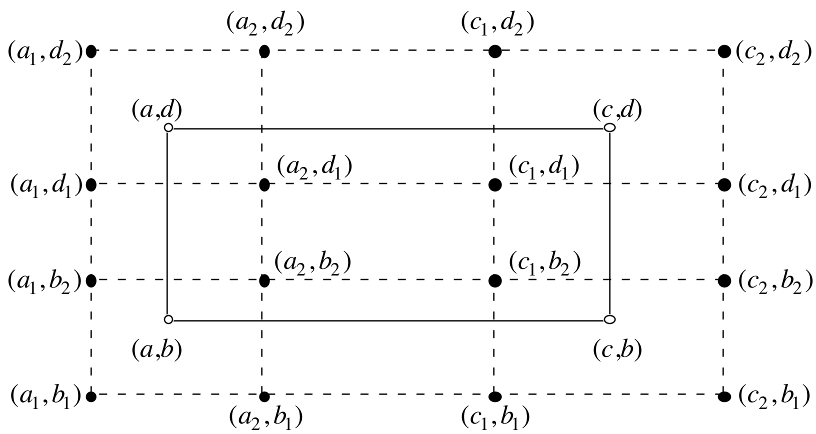

Fig. 2.5. The least simple case in the proof of Lemma 2.3.5

At the other extreme, the least simple case occurs when there is at least one point in $\overline{S}_1$ strictly between a and c, and at least one point in $\overline{S}_2$ strictly between b and d, so that $a < a_2 \leq c_1 < c$ and $b < b_2 \leq d_1 < d$. In this case—which is illustrated in Fig. 2.5—substituting (2.3.2) and the corresponding terms for $C(a,d)$, $C(c,b)$ and $C(c,d)$ into the expression given by (2.1.1) for $V_C(B)$ and rearranging the terms yields

$$\begin{aligned} V_C(B) = {} & (1-\lambda_1)\mu_2 V_C([a_1,a_2]\times[d_1,d_2]) + \mu_2 V_C([a_2,c_1]\times[d_1,d_2]) \\ & + \lambda_2\mu_2 V_C([c_1,c_2]\times[d_1,d_2]) + (1-\lambda_1)V_C([a_1,a_2]\times[b_2,d_1]) \\ & + V_C([a_2,c_1]\times[b_2,d_1]) + \lambda_2 V_C([c_1,c_2]\times[b_2,d_1]) \\ & + (1-\lambda_1)(1-\mu_1)V_C([a_1,a_2]\times[b_1,b_2]) \\ & + (1-\mu_1)V_C([a_2,c_1]\times[b_1,b_2]) + \lambda_2(1-\mu_1)V_C([c_1,c_2]\times[b_1,b_2]). \end{aligned}$$

The right-hand side of the above expression is a combination of nine nonnegative quantities (the C-volumes of the nine rectangles determined by the dashed lines in Fig. 2.5) with nonnegative coefficients, and hence is nonnegative. The remaining cases are similar, which completes the proof. □

Example 2.6. Let (a,b) be any point in $\mathbf{R}^2$, and consider the following distribution function H:

$$H(x,y) = \begin{cases} 0, & x < a \text{ or } y < b, \\ 1, & x \geq a \text{ and } y \geq b. \end{cases}$$

The margins of H are the unit step functions ε_a and ε_b. Applying Lemma 2.3.4 yields the subcopula C' with domain $\{0,1\}\times\{0,1\}$ such that $C'(0,0) = C'(0,1) = C'(1,0) = 0$ and $C'(1,1) = 1$. The extension of C' to a copula C via Lemma 2.3.5 is the copula $C = \Pi$, i.e., $C(u,v) = uv$. Notice however, that *every* copula agrees with C' on its domain, and thus is an extension of this C'. ■

We are now ready to prove Sklar's theorem, which we restate here for convenience.

Theorem 2.3.3. Sklar's theorem. *Let H be a joint distribution function with margins F and G. Then there exists a copula C such that for all x,y in* $\overline{\mathbf{R}}$,

$$H(x,y) = C(F(x),G(y)). \tag{2.3.1}$$

If F and G are continuous, then C is unique; otherwise, C is uniquely determined on $\operatorname{Ran}F\times\operatorname{Ran}G$. *Conversely, if C is a copula and F and G are distribution functions, then the function H defined by* (2.3.1) *is a joint distribution function with margins F and G.*

Proof. The existence of a copula C such that (2.3.1) holds for all x,y in $\overline{\mathbf{R}}$ follows from Lemmas 2.3.4 and 2.3.5. If F and G are continuous, then $\operatorname{Ran}F = \operatorname{Ran}G = \mathbf{I}$, so that the unique subcopula in Lemma 2.3.4 is a copula. The converse is a matter of straightforward verification. □

Equation (2.3.1) gives an expression for joint distribution functions in terms of a copula and two univariate distribution functions. But (2.3.1) can be inverted to express copulas in terms of a joint distribution function and the "inverses" of the two margins. However, if a margin is not strictly increasing, then it does not possess an inverse in the usual sense. Thus we first need to define "quasi-inverses" of distribution functions (recall Definition 2.3.1).

Definition 2.3.6. Let F be a distribution function. Then a *quasi-inverse* of F is any function $F^{(-1)}$ with domain $\mathbf{I}$ such that

1. if t is in $\operatorname{Ran}F$, then $F^{(-1)}(t)$ is any number x in $\overline{\mathbf{R}}$ such that $F(x) = t$, i.e., for all t in $\operatorname{Ran}F$,

$$F(F^{(-1)}(t)) = t;$$

2. if t is not in $\operatorname{Ran}F$, then

$$F^{(-1)}(t)=\inf\{x|F(x)\geq t\}=\sup\{x|F(x)\leq t\}.$$

If F is strictly increasing, then it has but a single quasi-inverse, which is of course the ordinary inverse, for which we use the customary notation F^{-1}.

Example 2.7. The quasi-inverses of ε_a, the unit step at a (see Example 2.4) are the functions given by

$$\varepsilon_a^{(-1)}(t)=\begin{cases}a_0, & t=0,\\ a, & t\in(0,1),\\ a_1, & t=1,\end{cases}$$

where a_0 and a_1 are any numbers in $\overline{\mathbf{R}}$ such that $a_0 < a \leq a_1$. ■

Using quasi-inverses of distribution functions, we now have the following corollary to Lemma 2.3.4.

Corollary 2.3.7. *Let H, F, G, and* C' *be as in Lemma* 2.3.4, *and let* $F^{(-1)}$ *and* $G^{(-1)}$ *be quasi-inverses of F and G, respectively. Then for any* (u,v) *in* Dom C',

$$C'(u,v)=H(F^{(-1)}(u),G^{(-1)}(v)). \tag{2.3.3}$$

When F and G are continuous, the above result holds for copulas as well and provides a method of constructing copulas from joint distribution functions. We will exploit Corollary 2.3.7 in the next chapter to construct families of copulas, but for now the following examples will serve to illustrate the procedure.

Example 2.8. Recall the distribution function H from Example 2.5:

$$H(x,y)=\begin{cases}\dfrac{(x+1)(e^y-1)}{x+2e^y-1}, & (x,y)\in[-1,1]\times[0,\infty],\\ 1-e^{-y}, & (x,y)\in(1,\infty]\times[0,\infty],\\ 0, & \text{elsewhere.}\end{cases}$$

with margins F and G given by

$$F(x)=\begin{cases}0, & x<-1,\\ (x+1)/2, & x\in[-1,1],\\ 1, & x>1,\end{cases}\quad\text{and}\quad G(y)=\begin{cases}0, & y<0,\\ 1-e^{-y}, & y\geq 0.\end{cases}$$

Quasi-inverses of F and G are given by $F^{(-1)}(u) = 2u-1$ and $G^{(-1)}(v) = -\ln(1-v)$ for u,v in $\mathbf{I}$. Because Ran$F$ = Ran$G$ = $\mathbf{I}$, (2.3.3) yields the copula C given by

$$C(u,v) = \frac{uv}{u+v-uv}. \tag{2.3.4}$$

■

Example 2.9. *Gumbel's bivariate exponential distribution* (Gumbel 1960a). Let H_θ be the joint distribution function given by

$$H_\theta(x,y) = \begin{cases} 1-e^{-x}-e^{-y}+e^{-(x+y+\theta xy)}, & x \geq 0, y \geq 0, \\ 0, & \text{otherwise}; \end{cases}$$

where θ is a parameter in [0,1]. Then the marginal distribution functions are exponentials, with quasi-inverses $F^{(-1)}(u) = -\ln(1-u)$ and $G^{(-1)}(v) = -\ln(1-v)$ for u,v in **I**. Hence the corresponding copula is

$$C_\theta(u,v) = u+v-1+(1-u)(1-v)e^{-\theta\ln(1-u)\ln(1-v)}. \tag{2.3.5}$$

■

Example 2.10. It is an exercise in many mathematical statistics texts to find an example of a bivariate distribution with standard normal margins that is not the standard bivariate normal with parameters $\mu_x = \mu_y = 0$, $\sigma_x^2 = \sigma_y^2 = 1$, and Pearson's product-moment correlation coefficient ρ. With Sklar's theorem and Corollary 2.3.7 this becomes trivial—let C be a copula such as one in either of the preceding examples, and use standard normal margins in (2.3.1). Indeed, if Φ denotes the standard (univariate) normal distribution function and N_ρ denotes the standard bivariate normal distribution function (with Pearson's product-moment correlation coefficient ρ), then *any* copula *except* one of the form

$$\begin{aligned} C(u,v) &= N_\rho(\Phi^{-1}(u),\Phi^{-1}(v)) \\ &= \frac{1}{2\pi\sqrt{1-\rho^2}}\int_{-\infty}^{\Phi^{-1}(u)}\int_{-\infty}^{\Phi^{-1}(v)}\exp\left[\frac{-(s^2-2\rho st+t^2)}{2(1-\rho^2)}\right]dsdt \end{aligned} \tag{2.3.6}$$

(with $\rho \neq -1$, 0, or 1) will suffice. Explicit constructions using the copulas in Exercises 2.4, 2.12, and 3.11, Example 3.12, and Sect. 3.3.1 can be found in (Kowalski 1973), and one using the copula $C_{1/2}$ from Exercise 2.10 in (Vitale 1978). ■

We close this section with one final observation. With an appropriate extension of its domain to $\overline{\mathbf{R}}^2$, every copula is a joint distribution function with margins that are uniform on **I**. To be precise, let C be a copula, and define the function H_C on $\overline{\mathbf{R}}^2$ via

$$H_C(x,y)=\begin{cases}0, & x<0 \text{ or } y<0,\\ C(x,y), & (x,y)\in\mathbf{I}^2,\\ x, & y>1, x\in\mathbf{I},\\ y, & x>1, y\in\mathbf{I},\\ 1, & x>1 \text{ and } y>1.\end{cases}$$

Then H_C is a distribution function both of whose margins are readily seen to be U_{01}. Indeed, it is often quite useful to think of copulas as restrictions to $\mathbf{I}^2$ of joint distribution functions whose margins are U_{01}.

2.4 Copulas and Random Variables

In this book, we will use the term "random variable" in the statistical rather than the probabilistic sense; that is, a random variable is a quantity whose values are described by a (known or unknown) probability distribution function. Of course, all of the results to follow remain valid when a random variable is defined in terms of measure theory, i.e., as a measurable function on a given probability space. But for our purposes it suffices to adopt the descriptions of Wald (1947), "a variable x is called a random variable if for any given value c a definite probability can be ascribed to the event that x will take a value less than c"; and of Gnedenko (1962), "a *random variable* is a variable quantity whose values depend on chance and for which there exists a distribution function." For a detailed discussion of this point of view, see (Menger 1956).

In what follows, we will use capital letters, such as X and Y, to represent random variables, and lowercase letters x, y to represent their values. We will say that F is the *distribution function of the random variable X* when for all x in $\overline{\mathbf{R}}$, $F(x) = P[X \le x]$. We are defining distribution functions of random variables to be right-continuous—but that is simply a matter of custom and convenience. Left-continuous distribution functions would serve equally as well. A random variable is continuous if its distribution function is continuous.

When we discuss two or more random variables, we adopt the same convention—two or more random variables are the components of a quantity (now a vector) whose values are described by a joint distribution function. As a consequence, we always assume that the collection of random variables under discussion can be defined on a common probability space.

We are now in a position to restate Sklar's theorem in terms of random variables and their distribution functions:

Theorem 2.4.1. *Let X and Y be random variables with distribution functions F and G, respectively, and joint distribution function H. Then* *there exists a copula C such that* (2.3.1) *holds. If F and G are continuous, C is unique. Otherwise, C is uniquely determined on* Ran$F\times$Ran$G$.

The copula C in Theorem 2.4.1 will be called the *copula of X and Y*, and denoted C_{XY} when its identification with the random variables X and Y is advantageous.

The following theorem shows that the product copula $\Pi(u,v) = uv$ characterizes independent random variables when the distribution functions are continuous. Its proof follows from Theorem 2.4.1 and the observation that X and Y are independent if and only if $H(x,y) = F(x)G(y)$ for all x,y in $\overline{\mathbf{R}}^2$.

Theorem 2.4.2. *Let X and Y be continuous random variables. Then X and Y are independent if and only if* $C_{XY} = \Pi$.

Much of the usefulness of copulas in the study of nonparametric statistics derives from the fact that for strictly monotone transformations of the random variables, copulas are either invariant or change in predictable ways. Recall that if the distribution function of a random variable X is continuous, and if α is a strictly monotone function whose domain contains Ran$X$, then the distribution function of the random variable $\alpha(X)$ is also continuous. We treat the case of strictly increasing transformations first.

Theorem 2.4.3. *Let X and Y be continuous random variables with copula* C_{XY}. *If* α *and* β *are strictly increasing on* Ran$X$ *and* Ran$Y$, *respectively, then* $C_{\alpha(X)\beta(Y)} = C_{XY}$. *Thus* C_{XY} *is invariant under strictly increasing transformations of X and Y.*

Proof. Let F_1, G_1, F_2, and G_2 denote the distribution functions of X, Y, $\alpha(X)$, and $\beta(Y)$, respectively. Because α and β are strictly increasing, $F_2(x) = P[\alpha(X) \le x] = P[X \le \alpha^{-1}(x)] = F_1(\alpha^{-1}(x))$, and likewise $G_2(y) = G_1(\beta^{-1}(y))$. Thus, for any x,y in $\overline{\mathbf{R}}$,

$$\begin{aligned} C_{\alpha(X)\beta(Y)}(F_2(x),G_2(y)) &= P[\alpha(X) \le x,\beta(Y) \le y] \\ &= P[X \le \alpha^{-1}(x),Y \le \beta^{-1}(y)] \\ &= C_{XY}(F_1(\alpha^{-1}(x)),G_1(\beta^{-1}(y))) \\ &= C_{XY}(F_2(x),G_2(y)). \end{aligned}$$

Because X and Y are continuous, Ran F_2 = Ran G_2 = $\mathbf{I}$, whence it follows that $C_{\alpha(X)\beta(Y)} = C_{XY}$ on $\mathbf{I}^2$. □

When at least one of α and β is strictly decreasing, we obtain results in which the copula of the random variables $\alpha(X)$ and $\beta(Y)$ is a simple transformation of C_{XY}. Specifically, we have:

Theorem 2.4.4. *Let X and Y be continuous random variables with copula* C_{XY}. *Let* α *and* β *be strictly monotone on* Ran$X$ *and* Ran$Y$, *respectively.*

1. *If* α *is strictly increasing and* β *is strictly decreasing, then*

$$C_{\alpha(X)\beta(Y)}(u,v) = u - C_{XY}(u,1-v).$$

2. *If* α *is strictly decreasing and* β *is strictly increasing, then*

$$C_{\alpha(X)\beta(Y)}(u,v) = v - C_{XY}(1-u,v).$$

3. *If* α *and* β *are both strictly decreasing, then*

$$C_{\alpha(X)\beta(Y)}(u,v) = u + v - 1 + C_{XY}(1-u,1-v).$$

The proof of Theorem 2.4.4 is left as an exercise. Note that in each case the form of the copula is independent of the particular choices of α and β, and note further that the three forms for $C_{\alpha(X)\beta(Y)}$ that appear in this theorem were first encountered in Exercise 2.6. [Remark: We could be somewhat more general in the preceding two theorems by replacing phrases such as "strictly increasing" by "almost surely strictly increasing"—to allow for subsets of Lebesgue measure zero where the property may fail to hold.]

Although we have chosen to avoid measure theory in our definition of random variables, we will nevertheless need some terminology and results from measure theory in the remaining sections of this chapter and in chapters to come. Each joint distribution function H induces a probability measure on $\mathbf{R}^2$ via $V_H\big((-\infty,x]\times(-\infty,y]\big) = H(x,y)$ and a standard extension to Borel subsets of $\mathbf{R}^2$ using measure-theoretic techniques. Because copulas are joint distribution functions (with uniform (0,1) margins), each copula C induces a probability measure on $\mathbf{I}^2$ via $V_C([0,u]\times[0,v]) = C(u,v)$ in a similar fashion—that is, the C-measure of a set is its C-volume V_C. Hence, at an intuitive level, the C-measure of a subset of $\mathbf{I}^2$ is the probability that two uniform (0,1) random variables U and V with joint distribution function C assume values in that subset. C-measures are often called *doubly stochastic measures*, as for any measurable subset S of $\mathbf{I}$, $V_C(S\times\mathbf{I}) = V_C(\mathbf{I}\times S) = \lambda(S)$, where λ denotes ordinary Lebesgue measure on $\mathbf{I}$. The term "doubly stochastic" is taken from matrix theory, where doubly stochastic matrices have nonnegative entries and all row sums and column sums are 1.

For any copula C, let

$$C(u,v) = A_C(u,v) + S_C(u,v),$$

where

$$A_C(u,v) = \int_0^u \int_0^v \frac{\partial^2}{\partial s \partial t} C(s,t)\, dtds \text{ and } S_C(u,v) = C(u,v) - A_C(u,v). \quad (2.4.1)$$

Unlike bivariate distributions in general, the margins of a copula are continuous, hence a copula has no "atoms" (individual points in $\mathbf{I}^2$ whose C-measure is positive).

If $C \equiv A_C$ on $\mathbf{I}^2$—that is, if considered as a joint distribution function, C has a joint density given by $\partial^2 C(u,v)/\partial u \partial v$—then C is *absolutely continuous*, whereas if $C \equiv S_C$ on $\mathbf{I}^2$—that is, if $\partial^2 C(u,v)/\partial u \partial v = 0$ almost everywhere in $\mathbf{I}^2$—then C is *singular*. Otherwise, C has an *absolutely continuous component* A_C and a *singular component* S_C. In this case neither A_C nor S_C is a copula, because neither has uniform (0,1) margins. In addition, the C-measure of the absolutely continuous component is $A_C(1,1)$, and the C-measure of the singular component is $S_C(1,1)$.

Just as the support of a joint distribution function H is the complement of the union of all open subsets of $\mathbf{R}^2$ with H-measure zero, the *support of a copula* is the complement of the union of all open subsets of $\mathbf{I}^2$ with C-measure zero. When the support of C is $\mathbf{I}^2$, we say C has "full support." When C is singular, its support has Lebesgue measure zero (and conversely). However, many copulas that have full support have both an absolutely continuous and a singular component.

Example 2.11. The support of the Fréchet-Hoeffding upper bound M is the main diagonal of $\mathbf{I}^2$, i.e., the graph of $v = u$ for u in $\mathbf{I}$, so that M is singular. This follows from the fact that the M-measure of any open rectangle that lies entirely above or below the main diagonal is zero. Also note that $\partial^2 M/\partial u \partial v = 0$ everywhere in $\mathbf{I}^2$ except on the main diagonal. Similarly, the support of the Fréchet-Hoeffding lower bound W is the secondary diagonal of $\mathbf{I}^2$, i.e., the graph of $v = 1 - u$ for u in $\mathbf{I}$, and thus W is singular as well. ■

Example 2.12. The product copula $\Pi(u,v) = uv$ is absolutely continuous, because for all (u,v) in $\mathbf{I}^2$,

$$A_\Pi(u,v) = \int_0^u \int_0^v \frac{\partial^2}{\partial s \partial t} \Pi(s,t)\, dtds = \int_0^u \int_0^v 1\, dtds = uv = \Pi(u,v). \quad ■$$

In Sect. 3.1.1 we will illustrate a general procedure for decomposing a copula into the sum of its absolutely continuous and singular components and for finding the probability mass (i.e., C-measure) of each component.

Exercises

2.12 *Gumbel's bivariate logistic distribution* (Gumbel 1961). Let X and Y be random variables with a joint distribution function given by

$$H(x,y) = (1 + e^{-x} + e^{-y})^{-1}$$

for all x,y in $\overline{\mathbf{R}}$.
(a) Show that X and Y have standard (univariate) logistic distributions, i.e.,

$$F(x) = (1 + e^{-x})^{-1} \text{ and } G(y) = (1 + e^{-y})^{-1}.$$

(b) Show that the copula of X and Y is the copula given by (2.3.4) in Example 2.8.

2.13 *Type B bivariate extreme value distributions* (Johnson and Kotz 1972). Let X and Y be random variables with a joint distribution function given by

$$H_\theta(x,y) = \exp[-(e^{-\theta x} + e^{-\theta y})^{1/\theta}]$$

for all x,y in $\overline{\mathbf{R}}$, where $\theta \geq 1$. Show that the copula of X and Y is given by

$$C_\theta(u,v) = \exp\left(-\left[(-\ln u)^\theta + (-\ln v)^\theta\right]^{1/\theta}\right). \qquad (2.4.2)$$

This parametric family of copulas is known as the *Gumbel-Hougaard* family (Hutchinson and Lai 1990), which we shall see again in Chapter 4.

2.14 Conway (1979) and Hutchinson and Lai (1990) note that Gumbel's bivariate logistic distribution (Exercise 2.12) suffers from the defect that it lacks a parameter, which limits its usefulness in applications. This can be corrected in a number of ways, one of which (Ali et al. 1978) is to define H_θ as

$$H_\theta(x,y) = \left(1 + e^{-x} + e^{-y} + (1-\theta)e^{-x-y}\right)^{-1}$$

for all x,y in $\overline{\mathbf{R}}$, where θ lies in [–1,1]. Show that
(a) the margins are standard logistic distributions;

(b) when $\theta = 1$, we have Gumbel's bivariate logistic distribution;

(c) when $\theta = 0$, X and Y are independent; and

(d) the copula of X and Y is given by

$$C_\theta(u,v) = \frac{uv}{1-\theta(1-u)(1-v)}. \tag{2.4.3}$$

This is the *Ali-Mikhail-Haq* family of copulas (Hutchinson and Lai 1990), which we will encounter again in Chapters 3 and 4.

2.15 Let X_1 and Y_1 be random variables with continuous distribution functions F_1 and G_1, respectively, and copula C. Let F_2 and G_2 be another pair of continuous distribution functions, and set $X_2 = F_2^{(-1)}(F_1(X_1))$ and $Y_2 = G_2^{(-1)}(G_1(Y_1))$. Prove that

(a) the distribution functions of X_2 and Y_2 are F_2 and G_2, respectively; and

(b) the copula of X_2 and Y_2 is C.

2.16 (a) Let X and Y be continuous random variables with copula C and univariate distribution functions F and G, respectively. The random variables max(X,Y) and min(X,Y) are the *order statistics* for X and Y. Prove that the distribution functions of the order statistics are given by

$$P[\max(X,Y) \le t] = C(F(t),G(t))$$

and

$$P[\min(X,Y) \le t] = F(t) + G(t) \pm C(F(t),G(t)),$$

so that when $F = G$,

$$P[\max(X,Y) \le t] = \delta_C(F(t)) \text{ and}$$
$$P[\min(X,Y) \le t] = 2F(t) \pm \delta_C(F(t)).$$

(b) Show that bounds on the distribution functions of the order statistics are given by

$$\max(F(t)+G(t)-1,0) \le P[\max(X,Y) \le t] \le \min(F(t),G(t))$$

and

$$\max(F(t),G(t)) \le P[\min(X,Y) \le t] \le \min(F(t)+G(t),1).$$

2.17 Prove Theorem 2.4.4.

2.5 The Fréchet-Hoeffding Bounds for Joint Distribution Functions

In Sect. 2.2 we encountered the Fréchet-Hoeffding bounds as universal bounds for copulas, i.e., for any copula C and for all u,v in $\mathbf{I}$,

$$W(u,v) = \max(u+v-1,0) \le C(u,v) \le \min(u,v) = M(u,v).$$

As a consequence of Sklar's theorem, if X and Y are random variables with a joint distribution function H and margins F and G, respectively, then for all x,y in $\overline{\mathbf{R}}$,

$$\max(F(x)+G(y)-1,0) \le H(x,y) \le \min(F(x),G(y)) \tag{2.5.1}$$

Because M and W are copulas, the above bounds are joint distribution functions and are called the *Fréchet-Hoeffding bounds* for joint distribution functions H with margins F and G. Of interest in this section is the following question: What can we say about the random variables X and Y when their joint distribution function H is equal to one of its Fréchet-Hoeffding bounds?

To answer this question, we first need to introduce the notions of nondecreasing and nonincreasing sets in $\overline{\mathbf{R}}^2$.

Definition 2.5.1. A subset S of $\overline{\mathbf{R}}^2$ is *nondecreasing* if for any (x,y) and (u,v) in S, $x < u$ implies $y \le v$. Similarly, a subset S of $\overline{\mathbf{R}}^2$ is *nonincreasing* if for any (x,y) and (u,v) in S, $x < u$ implies $y \ge v$.

Fig. 2.6 illustrates a simple nondecreasing set.

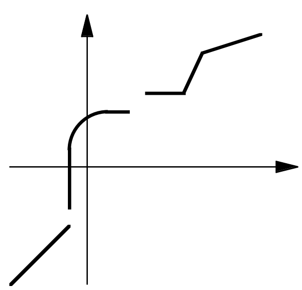

Fig. 2.6. The graph of a nondecreasing set

We will now prove that the joint distribution function H for a pair (X,Y) of random variables is the Fréchet-Hoeffding upper bound (i.e., the copula is M) if and only if the support of H lies in a nondecreasing set. The following proof is based on the one that appears in (Mikusiński, Sherwood and Taylor 1991-1992). But first, we need two lemmas:

Lemma 2.5.2. *Let S be a subset of* $\overline{\mathbf{R}}^2$. *Then S is nondecreasing if and only if for each* (x,y) *in* $\overline{\mathbf{R}}^2$, *either*

1. *for all* (u,v) *in S,* $u \le x$ *implies* $v \le y$; *or* (2.5.2)
2. *for all* (u,v) *in S,* $v \le y$ *implies* $u \le x$. (2.5.3)

Proof. First assume that S is nondecreasing, and that neither (2.5.2) nor (2.5.3) holds. Then there exist points (a,b) and (c,d) in S such that $a \le x$, $b > y$, $d \le y$, and $c > x$. Hence $a < c$ and $b > d$; a contradiction. In the opposite direction, assume that S is not nondecreasing. Then there exist points (a,b) and (c,d) in S with $a < c$ and $b > d$. For $(x,y) = \big((a+c)/2,(b+d)/2\big)$, neither (2.5.2) nor (2.5.3) holds. □

Lemma 2.5.3. *Let X and Y be random variables with joint distribution function H. Then H is equal to its Fréchet-Hoeffding upper bound if and only if for every* (x,y) *in* $\overline{\mathbf{R}}^2$, *either* $P[X > x, Y \le y] = 0$ *or* $P[X \le x, Y > y] = 0$.

Proof: As usual, let F and G denote the margins of H. Then

$$\begin{aligned} F(x) = P[X \le x] &= P[X \le x, Y \le y] + P[X \le x, Y > y] \\ &= H(x,y) + P[X \le x, Y > y], \end{aligned}$$

and

$$\begin{aligned} G(y) = P[Y \le y] &= P[X \le x, Y \le y] + P[X > x, Y \le y] \\ &= H(x,y) + P[X > x, Y \le y]. \end{aligned}$$

Hence $H(x,y) = M(F(x),G(y))$ if and only if $\min(P[X \le x, Y > y], P[X > x, Y \le y]) = 0$, from which the desired conclusion follows. □

We are now ready to prove

Theorem 2.5.4. *Let X and Y be random variables with joint distribution function H. Then H is identically equal to its Fréchet-Hoeffding upper bound if and only if the support of H is a nondecreasing subset of* $\overline{\mathbf{R}}^2$.

Proof. Let S denote the support of H, and let (x,y) be any point in $\overline{\mathbf{R}}^2$. Then (2.5.2) holds if and only if $\{(u,v) \mid u \le x \text{ and } v > y\} \cap S = \varnothing$; or equivalently, if and only if $P[X \le x, Y > y] = 0$. Similarly, (2.5.3) holds if and only if $\{(u,v) \mid u > x \text{ and } v \le y\} \cap S = \varnothing$; or equivalently, if and only if $P[X > x, Y \le y] = 0$. The theorem now follows from Lemmas 2.5.2 and 2.5.3. □

Of course, there is an analogous result for the Fréchet-Hoeffding lower bound—its proof is outlined in Exercises 2.18 through 2.20:

Theorem 2.5.5. *Let X and Y be random variables with joint distribution function H. Then H is identically equal to its Fréchet-Hoeffding lower bound if and only if the support of H is a nonincreasing subset of* $\overline{\mathbf{R}}^2$.

When X and Y are continuous, the support of H can have no horizontal or vertical line segments, and in this case it is common to say that "Y is almost surely an increasing function of X" if and only if the copula of X and Y is M; and "Y is almost surely a decreasing function of X" if and only if the copula of X and Y is W. If U and V are uniform (0,1) random variables whose joint distribution function is the copula M, then $P[U = V] = 1$; and if the copula is W, then $P[U + V = 1] = 1$.

Random variables with copula M are often called *comonotonic*, and random variables with copula W are often called *countermonotonic*.

2.6 Survival Copulas

In many applications, the random variables of interest represent the lifetimes of individuals or objects in some population. The probability of an individual living or surviving beyond time x is given by the *survival function* (or *survivor function*, or *reliability function*) $\overline{F}(x) = P[X > x] = 1 - F(x)$, where, as before, F denotes the distribution function of X. When dealing with lifetimes, the natural range of a random variable is often $[0,\infty)$; however, we will use the term "survival function" for $P[X > x]$ even when the range is $\overline{\mathbf{R}}$.

For a pair (X,Y) of random variables with joint distribution function H, the *joint survival function* is given by $\overline{H}(x,y) = P[X > x, Y > y]$. The margins of $\overline{H}$ are the functions $\overline{H}(x,-\infty)$ and $\overline{H}(-\infty,y)$, which are the univariate survival functions $\overline{F}$ and $\overline{G}$, respectively. A natural question is the following: Is there a relationship between univariate and joint survival functions analogous to the one between univariate and joint distribution functions, as embodied in Sklar's theorem? To answer this question, suppose that the copula of X and Y is C. Then we have

$$\begin{aligned}\overline{H}(x,y) &= 1 - F(x) - G(y) + H(x,y)\\ &= \overline{F}(x) + \overline{G}(y) - 1 + C(F(x),G(y))\\ &= \overline{F}(x) + \overline{G}(y) - 1 + C(1-\overline{F}(x),1-\overline{G}(y)),\end{aligned}$$

so that if we define a function $\hat{C}$ from $\mathbf{I}^2$ into $\mathbf{I}$ by

$$\hat{C}(u,v) = u + v - 1 + C(1-u,1-v), \tag{2.6.1}$$

we have

$$\overline{H}(x,y) = \hat{C}(\overline{F}(x),\overline{G}(y)). \tag{2.6.2}$$

First note that, as a consequence of Exercise 2.6, the function $\hat{C}$ in (2.6.1) is a copula (see also part 3 of Theorem 2.4.4). We refer to $\hat{C}$ as the *survival copula* of X and Y. Secondly, notice that $\hat{C}$ "couples" the joint survival function to its univariate margins in a manner completely analogous to the way in which a copula connects the joint distribution function to its margins.

Care should be taken not to confuse the survival copula $\hat{C}$ with the joint survival function $\overline{C}$ for two uniform (0,1) random variables whose joint distribution function is the copula C. Note that $\overline{C}(u,v) = P[U > u, V > v] = 1 - u - v + C(u,v) = \hat{C}(1-u, 1-v)$.

Example 2.13. In Example 2.9, we obtained the copula C_θ in (2.3.5) for Gumbel's bivariate exponential distribution: for θ in [0,1],

$$C_\theta(u,v) = u + v - 1 + (1-u)(1-v)e^{-\theta \ln(1-u)\ln(1-v)}.$$

Just as the survival function for univariate exponentially distributed random variables is functionally simpler than the distribution function, the same is often true in the bivariate case. Employing (2.6.1), we have

$$\hat{C}_\theta(u,v) = uve^{-\theta \ln u \ln v}. \qquad \blacksquare$$

Example 2.14. *A bivariate Pareto distribution* (Hutchinson and Lai 1990). Let X and Y be random variables whose joint survival function is given by

$$\overline{H}_\theta(x,y) = \begin{cases} (1+x+y)^{-\theta}, & x \ge 0, y \ge 0, \\ (1+x)^{-\theta}, & x \ge 0, y < 0, \\ (1+y)^{-\theta}, & x < 0, y \ge 0, \\ 1, & x < 0, y < 0; \end{cases}$$

where $\theta > 0$. Then the marginal survival functions $\overline{F}$ and $\overline{G}$ are

$$\overline{F}(x) = \begin{cases} (1+x)^{-\theta}, & x \ge 0 \\ 1, & x < 0, \end{cases} \quad \text{and} \quad \overline{G}(y) = \begin{cases} (1+y)^{-\theta}, & y \ge 0 \\ 1, & y < 0, \end{cases}$$

so that X and Y have identical Pareto distributions. Inverting the survival functions and employing the survival version of Corollary 2.3.7 (see Exercise 2.25) yields the survival copula

$$\hat{C}_\theta(u,v) = (u^{-1/\theta} + v^{-1/\theta} - 1)^{-\theta}. \qquad (2.6.3)$$

We shall encounter this family again in Chapter 4. $\blacksquare$

Two other functions closely related to copulas—and survival copulas—are the *dual of a copula* and the *co-copula* (Schweizer and Sklar 1983). The dual of a copula C is the function $\tilde{C}$ defined by $\tilde{C}(u,v) = u + v - C(u,v)$; and the co-copula is the function C^* defined by $C^*(u,v) = 1 - C(1-u, 1-v)$. Neither of these is a copula, but when C

is the copula of a pair of random variables X and Y, the dual of the copula and the co-copula each express a probability of an event involving X and Y. Just as

$$P[X \le x, Y \le y] = C(F(x),G(y)) \text{ and } P[X > x, Y > y] = \hat{C}(\overline{F}(x),\overline{G}(y)),$$

we have

$$P[X \le x \text{ or } Y \le y] = \tilde{C}(F(x),G(y)), \qquad (2.6.4)$$

and

$$P[X > x \text{ or } Y > y] = C^*(\overline{F}(x),\overline{G}(y)). \qquad (2.6.5)$$

Other relationships among C, $\hat{C}$, $\tilde{C}$, and C^* are explored in Exercises 2.24 and 2.25.

Exercises

2.18 Prove the “Fréchet-Hoeffding lower bound” version of Lemma 2.5.2: Let S be a subset of $\overline{\mathbf{R}}^2$. Then S is nonincreasing if and only if for each (x,y) in $\overline{\mathbf{R}}^2$, either
   1. for all (u,v) in S, $u \le x$ implies $v > y$; or
   2. for all (u,v) in S, $v > y$ implies $u \le x$.

2.19 Prove the “Fréchet-Hoeffding lower bound” version of Lemma 2.5.3: Let X and Y be random variables whose joint distribution function H is equal to its Fréchet-Hoeffding lower bound. Then for every (x,y) in $\overline{\mathbf{R}}^2$, either $P[X > x, Y > y] = 0$ or $P[X \le x, Y \le y] = 0$

2.20 Prove Theorem 2.5.5.

2.21 Let X and Y be nonnegative random variables whose survival function is $\overline{H}(x,y) = (e^x + e^y - 1)^{-1}$ for $x,y \ge 0$.
(a) Show that X and Y are standard exponential random variables.
(b) Show that the copula of X and Y is the copula given by (2.3.4) in Example 2.8 [cf. Exercise 2.12].

2.22 Let X and Y be continuous random variables whose joint distribution function is given by $C(F(x),G(y))$, where C is the copula of X and Y, and F and G are the distribution functions of X and Y respectively. Verify that (2.6.4) and (2.6.5) hold.

2.23 Let X_1, Y_1, F_1, G_1, F_2, G_2, and C be as in Exercise 2.15. Set $X_2 = F_2^{(-1)}(1 - F_1(X_1))$ and $Y_2 = G_2^{(-1)}(1 - G_1(Y_1))$. Prove that
(a) The distribution functions of X_2 and Y_2 are F_2 and G_2, respectively; and
(b) The copula of X_2 and Y_2 is $\hat{C}$.

2.24 Let X and Y be continuous random variables with copula C and a common univariate distribution function F. Show that the distribution and survival functions of the order statistics (see Exercise 2.16) are given by

Order statistic	Distribution function	Survival function
$\max(X,Y)$	$\delta(F(t))$	$\delta^*(\overline{F}(t))$
$\min(X,Y)$	$\tilde{\delta}(F(t))$	$\hat{\delta}(\overline{F}(t))$

where δ, $\hat{\delta}$, $\tilde{\delta}$, and δ^* denote the diagonal sections of C, $\hat{C}$, $\tilde{C}$, and C^*, respectively.

2.25 Show that under composition $\circ$, the set of operations of forming the survival copula, the dual of a copula, and the co-copula of a given copula, along with the identity (i.e., "$\wedge$", "$\sim$", "$*$", and "i") yields the dihedral group (e.g., $C^{**} = C$, so $*\circ* = i$; $\hat{C}^* = \tilde{C}$, so $\wedge\circ* = \sim$, etc.):

$\circ$	i	$\wedge$	$\sim$	$*$
i	i	$\wedge$	$\sim$	$*$
$\wedge$	$\wedge$	i	$*$	$\sim$
$\sim$	$\sim$	$*$	i	$\wedge$
$*$	$*$	$\sim$	$\wedge$	i

2.26 Prove the following "survival" version of Corollary 2.3.7: Let $\overline{H}$, $\overline{F}$, $\overline{G}$, and $\hat{C}$ be as in (2.6.2), and let $\overline{F}^{(-1)}$ and $\overline{G}^{(-1)}$ be quasi-inverses of $\overline{F}$ and $\overline{G}$, respectively. Then for any (u,v) in $\mathbf{I}^2$,

$$\hat{C}(u,v) = \overline{H}(\overline{F}^{(-1)}(u), \overline{G}^{(-1)}(v)).$$

2.7 Symmetry

If X is a random variable and a is a real number, we say that X is *symmetric about* a if the distribution functions of the random variables $X-a$ and $a-X$ are the same, that is, if for any x in $\mathbf{R}$, $P[X-a \le x] = P[a-X \le x]$. When X is continuous with distribution function F, this is equivalent to

$$F(a+x) = \overline{F}(a-x) \tag{2.7.1}$$

[when F is discontinuous, (2.7.1) holds only at the points of continuity of F].

Now consider the bivariate situation. What does it mean to say that a pair (X,Y) of random variables is "symmetric" about a point (a,b)? There are a number of ways to answer this question, and each answer leads to a different type of bivariate symmetry.

Definition 2.7.1. Let X and Y be random variables and let (a,b) be a point in $\mathbf{R}^2$.

1. (X,Y) is *marginally symmetric* about (a,b) if X and Y are symmetric about a and b, respectively.
2. (X,Y) is *radially symmetric* about (a,b) if the joint distribution function of $X-a$ and $Y-b$ is the same as the joint distribution function of $a-X$ and $b-Y$.
3. (X,Y) is *jointly symmetric* about (a,b) if the following four pairs of random variables have a common joint distribution: $(X-a, Y-b)$, $(X-a, b-Y)$, $(a-X, Y-b)$, and $(a-X, b-Y)$.

When X and Y are continuous, we can express the condition for radial symmetry in terms of the joint distribution and survival functions of X and Y in a manner analogous to the relationship in (2.7.1) between univariate distribution and survival functions:

Theorem 2.7.2. *Let X and Y be continuous random variables with joint distribution function H and margins F and G, respectively. Let (a,b) be a point in $\mathbf{R}^2$. Then (X,Y) is radially symmetric about (a,b) if and only if*

$$H(a+x,b+y) = \overline{H}(a-x,b-y) \text{ for all } (x,y) \text{ in } \mathbf{R}^2. \tag{2.7.2}$$

The term "radial" comes from the fact that the points $(a+x,b+y)$ and $(a-x,b-y)$ that appear in (2.7.2) lie on rays emanating in opposite directions from (a,b). Graphically, Theorem 2.7.2 states that regions such as those shaded in Fig. 2.7(a) always have equal H-volume.

Example 2.15. The bivariate normal distribution with parameters μ_x, μ_y, σ_x^2, σ_y^2, and ρ is radially symmetric about the point (μ_x,μ_y). The proof is straightforward (but tedious)—evaluate double integrals of the joint density over the shaded regions in Fig. 2.7(a). ■

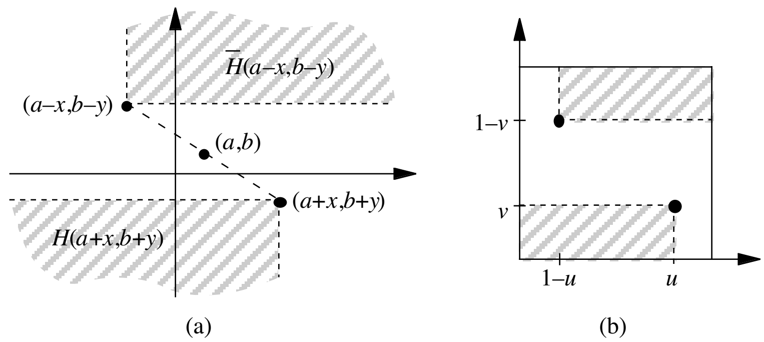

Fig. 2.7. Regions of equal probability for radially symmetric random variables

Example 2.16. The bivariate normal is a member of the family of elliptically contoured distributions. The densities for such distributions have contours that are concentric ellipses with constant eccentricity. Well-known members of this family, in addition to the bivariate normal, are bivariate Pearson type II and type VII distributions (the latter including bivariate t and Cauchy distributions as special cases). Like the bivariate normal, elliptically contoured distributions are radially symmetric. ■

It is immediate that joint symmetry implies radial symmetry and easy to see that radial symmetry implies marginal symmetry (setting $x = \infty$ in (2.7.2) yields (2.7.1); similarly for $y = \infty$). Indeed, joint symmetry is a very strong condition—it is easy to show that jointly symmetric random variables must be uncorrelated when the requisite second-order moments exist (Randles and Wolfe 1979). Consequently, we will focus on radial symmetry, rather than joint symmetry, for bivariate distributions.

Because the condition for radial symmetry in (2.7.2) involves both the joint distribution and survival functions, it is natural to ask if copulas and survival copulas play a role in radial symmetry. The answer is provided by the next theorem.

Theorem 2.7.3. *Let X and Y be continuous random variables with joint distribution function H, marginal distribution functions F and G, respectively, and copula C. Further suppose that X and Y are symmetric about a and b, respectively. Then* (X,Y) *is radially symmetric about* (a,b), *i.e., H satisfies* (2.7.2), *if and only if* $C = \hat{C}$, *i.e., if and only if C satisfies the functional equation*

$$C(u,v) = u + v - 1 + C(1-u,1-v) \text{ for all } (u,v) \text{ in } \mathbf{I}^2. \tag{2.7.3}$$

Proof. Employing (2.6.2) and (2.7.1), the theorem follows from the following chain of equivalent statements:

$$\begin{aligned}
&H(a+x,b+y)=\overline{H}(a-x,b-y) \text{ for all } (x,y) \text{ in } \mathbf{R}^2\\
&\Leftrightarrow C(F(a+x),G(b+y))=\hat{C}(\overline{F}(a-x),\overline{G}(b-y)) \text{ for all } (x,y) \text{ in } \mathbf{R}^2,\\
&\Leftrightarrow C(F(a+x),G(b+y))=\hat{C}(F(a+x),G(b+y)) \text{ for all } (x,y) \text{ in } \mathbf{R}^2,\\
&\Leftrightarrow C(u,v)=\hat{C}(u,v) \text{ for all } (u,v) \text{ in } \mathbf{I}^2. \qquad \square
\end{aligned}$$

Geometrically, (2.7.3) states that for any (u,v) in $\mathbf{I}^2$, the rectangles $[0,u]\times[0,v]$ and $[1-u,1]\times[1-v,1]$ have equal C-volume, as illustrated in Fig. 2.7(b).

Another form of symmetry is exchangeability—random variables X and Y are *exchangeable* if the vectors (X,Y) and (Y,X) are identically distributed. Hence if the joint distribution function of X and Y is H, then $H(x,y) = H(y,x)$ for all x,y in $\overline{\mathbf{R}}^2$. Clearly exchangeable random variables must be identically distributed, i.e., have a common univariate distribution function. For identically distributed random variables, exchangeability is equivalent to the symmetry of their copula as expressed in the following theorem, whose proof is straightforward.

Theorem 2.7.4. *Let X and Y be continuous random variables with joint distribution function H, margins F and G, respectively, and copula C. Then X and Y are exchangeable if and only if $F = G$ and $C(u,v) = C(v,u)$ for all (u,v) in $\mathbf{I}^2$.*

When $C(u,v) = C(v,u)$ for all (u,v) in $\mathbf{I}^2$, we will say simply that C is *symmetric*.

Example 2.17. Although identically distributed independent random variables must be exchangeable (because the copula Π is symmetric), the converse is of course not true—identically distributed exchangeable random variables need not be independent. To show this, simply choose for the copula of X and Y *any* symmetric copula except Π, such as one from Example 2.8, 2.9 (or 2.13), or from one of the families in Exercises 2.4 and 2.5. ■

There are other bivariate symmetry concepts. See (Nelsen 1993) for details.

2.8 Order

The Fréchet-Hoeffding bounds inequality—$W(u,v) \le C(u,v) \le M(u,v)$ for every copula C and all u,v in $\mathbf{I}$—suggests a partial order on the set of copulas:

Definition 2.8.1. If C_1 and C_2 are copulas, we say that C_1 *is smaller than* C_2 (or C_2 *is larger than* C_1), and write $C_1 \prec C_2$ (or $C_2 \succ C_1$) if $C_1(u,v) \le C_2(u,v)$ for all u,v in $\mathbf{I}$.

In other words, the Fréchet-Hoeffding lower bound copula W is smaller than every copula, and the Fréchet-Hoeffding upper bound copula M is larger than every copula. This point-wise partial ordering of the set of copulas is called the *concordance ordering* and will be important in Chapter 5 when we discuss the relationship between copulas and dependence properties for random variables (at which time the reason for the name of the ordering will become apparent). It is a partial order rather than a total order because not every pair of copulas is comparable.

Example 2.18. The product copula Π and the copula obtained by averaging the Fréchet-Hoeffding bounds are not comparable. If we let $C(u,v) = [W(u,v)+M(u,v)]/2$, then $C(1/4,1/4) > \Pi(1/4,1/4)$ and $C(1/4,3/4) < \Pi(1/4,3/4)$, so that neither $C \prec \Pi$ nor $\Pi \prec C$ holds. ■

However, there are families of copulas that are totally ordered. We will call a totally ordered parametric family $\{C_\theta\}$ of copulas *positively ordered* if $C_\alpha \prec C_\beta$ whenever $\alpha \le \beta$; and *negatively ordered* if $C_\alpha \succ C_\beta$ whenever $\alpha \le \beta$.

Example 2.19. The Cuadras-Augé family of copulas (2.2.10), introduced in Exercise 2.5, is positively ordered, as for $0 \le \alpha \le \beta \le 1$ and u,v in (0,1),

$$\frac{C_\alpha(u,v)}{C_\beta(u,v)} = \left(\frac{uv}{\min(u,v)}\right)^{\beta-\alpha} \le 1$$

and hence $C_\alpha \prec C_\beta$. ■

Exercises

2.27 Let X and Y be continuous random variables symmetric about a and b with marginal distribution functions F and G, respectively, and with copula C. Is (X,Y) is radially symmetric (or jointly symmetric) about (a,b) if C is
(a) a member of the Fréchet family in Exercise 2.4?
(b) a member of the Cuadras-Augé family in Exercise 2.5?

2.28. Suppose X and Y are identically distributed continuous random variables, each symmetric about a. Show that "exchangeability" does not imply "radial symmetry," nor does "radial symmetry" imply "exchangeability."

2.29 Prove the following analog of Theorem 2.7.2 for jointly symmetric random variables: Let X and Y be continuous random variables with joint distribution function H and margins F and G, respectively. Let (a,b) be a point in $\mathbf{R}^2$. Then (X,Y) is jointly symmetric about (a,b) if and only if

$$H(a+x,b+y) = F(a+x) - H(a+x,b-y) \text{ for all } (x,y) \text{ in } \overline{\mathbf{R}}^2$$

and

$$H(a+x,b+y) = G(b+y) - H(a-x,b+y) \text{ for all } (x,y) \text{ in } \overline{\mathbf{R}}^2.$$

2.30 Prove the following analog of Theorem 2.7.3 for jointly symmetric random variables: Let X and Y be continuous random variables with joint distribution function H, marginal distribution functions F and G, respectively, and copula C. Further suppose that X and Y are symmetric about a and b, respectively. Then (X,Y) is jointly symmetric about (a,b), i.e., H satisfies the equations in Exercise 2.28, if and only if C satisfies

$$C(u,v) = u - C(u,1-v) \text{ and } C(u,v) = v - C(1-u,v) \qquad (2.8.1)$$

for all (u,v) in $\mathbf{I}^2$. [Cf. Exercise 2.6 and Theorem 2.4.4].

2.31 (a) Show that $C_1 \prec C_2$ if and only if $\overline{C}_1 \prec \overline{C}_2$.
(b) Show that $C_1 \prec C_2$ if and only if $\hat{C}_1 \prec \hat{C}_2$.

2.32 Show that the Ali-Mikhail-Haq family of copulas (2.4.3) from Exercise 2.14 is positively ordered.

2.33 Show that the Mardia family of copulas (2.2.9) from Exercise 2.4 is neither positively nor negatively ordered. [Hint: evaluate C_0, $C_{1/4}$, and $C_{1/2}$ at $(u,v) = (3/4,1/4)$.]

2.9 Random Variate Generation

One of the primary applications of copulas is in simulation and Monte Carlo studies. In this section, we will address the problem of generating a sample from a specified joint distribution. Such samples can then be used to study mathematical models of real-world systems, or for statistical studies, such as the comparison of a new statistical method with competitors, robustness properties, or the agreement of asymptotic with small sample results.

We assume that the reader is familiar with various procedures used to generate independent uniform variates and with algorithms for using those variates to obtain samples from a given univariate distribution. One such method is the inverse distribution function method. To obtain an observation x of a random variable X with distribution function F:

1. Generate a variate u that is uniform on (0,1);
2. Set $x = F^{(-1)}(u)$, where $F^{(-1)}$ is any quasi-inverse of F (see Definition 2.3.6).

For a discussion and for alternative methods, see (Johnson 1987) or (Devroye 1986).

There are a variety of procedures used to generate observations (x,y) of a pair or random variables (X,Y) with a joint distribution function H. In this section, we will focus on using the copula as a tool. By virtue of Sklar's theorem, we need only generate a pair (u,v) of observations of uniform (0,1) random variables (U,V) whose joint distribution function is C, the copula of X and Y, and then transform those uniform variates via the algorithm such as the one in the preceding paragraph. One procedure for generating such of a pair (u,v) of uniform (0,1) variates is the conditional distribution method. For this method, we need the conditional distribution function for V given $U = u$, which we denote $c_u(v)$:

$$c_u(v) = P[V \le v | U = u] = \lim_{\Delta u \to 0} \frac{C(u+\Delta u, v) - C(u,v)}{\Delta u} = \frac{\partial C(u,v)}{\partial u} \qquad (2.9.1)$$

[Recall from Theorem 2.2.7 that the function $v \mapsto \partial C(u,v)/\partial u$, which we are now denoting $c_u(v)$, exists and is nondecreasing almost everywhere in **I**].

1. Generate two independent uniform (0,1) variates u and t;
2. Set $v = c_u^{(-1)}(t)$, where $c_u^{(-1)}$ denotes a quasi-inverse of c_u.
3. The desired pair is (u,v).

As with univariate distributions, there are many other algorithms—see (Johnson 1987) or (Devroye 1986) for details.

Example 2.20. Let X and Y be random variables whose joint distribution function H is

$$H(x,y) = \begin{cases} \dfrac{(x+1)(e^y - 1)}{x + 2e^y - 1}, & (x,y) \in [-1,1] \times [0,\infty], \\ 1 - e^{-y}, & (x,y) \in (1,\infty] \times [0,\infty], \\ 0, & \text{elsewhere.} \end{cases}$$

[recall Examples 2.5 and 2.8]. The copula C of X and Y is

$$C(u,v) = \frac{uv}{u + v - uv},$$

and so the conditional distribution function c_u and its inverse $c_u^{(-1)}$ are given by

$$c_u(v) = \frac{\partial}{\partial u}C(u,v) = \left(\frac{v}{u+v-uv}\right)^2 \quad \text{and} \quad c_u^{-1}(t) = \frac{u\sqrt{t}}{1-(1-u)\sqrt{t}}.$$

Thus an algorithm to generate random variates (x,y) is:

1. Generate two independent uniform (0,1) variates u and t;
2. Set $v = \dfrac{u\sqrt{t}}{1-(1-u)\sqrt{t}}$,
3. Set $x = 2u - 1$ and $y = -\ln(1 - v)$ [See Example 2.8 for the inverses of the marginal distribution functions.]
4. The desired pair is (x,y). ■

Survival copulas can also be used in the conditional distribution function method to generate random variates from a distribution with a given survival function. Recall [see part 3 of Theorem 2.4.4 and (2.6.1)] that if the copula C is the distribution function of a pair (U,V), then the corresponding survival copula $\hat{C}(u,v) = u+v-1+C(1-u,1-v)$ is the distribution function of the pair $(1-U,1-V)$. Also note that if U is uniform on (0,1), so is the random variable $1-U$. Hence we have the following algorithm to generate a pair (U,V) whose distribution function is the copula C, given $\hat{C}$:

1. Generate two independent uniform (0,1) variates u and t;
2. Set $v = \hat{c}_u^{(-1)}(t)$, where $\hat{c}_u^{(-1)}$ denotes a quasi-inverse of $\hat{c}_u(v) = \partial\hat{C}(u,v)/\partial u$.
3. The desired pair is (u,v).

In the next chapter we will be presenting methods that can be used to construct families of copulas. For many of those families, we will also indicate methods for generating random samples from the distributions that correspond to those copulas.

2.10 Multivariate Copulas

In this section, we extend the results of the preceding sections to the multivariate case. Although many of the definitions and theorems have analogous multivariate versions, not all do, so one must proceed with care. In the interest of clarity, we will restate most of the definitions and theorems in their multivariate versions. We will omit the proofs of theorems for which the proof is similar to that in the bivariate case. Many of the theorems in this section (with proofs) may be found in (Schweizer and Sklar 1983) or the references contained therein.

Some new notation will be advantageous here. For any positive integer n, we let $\overline{\mathbf{R}}^n$ denote the extended n-space $\overline{\mathbf{R}} \times \overline{\mathbf{R}} \times \cdots \times \overline{\mathbf{R}}$. We will use vector notation for points in $\overline{\mathbf{R}}^n$, e.g., $\mathbf{a} = (a_1, a_2, \cdots, a_n)$, and we will write $\mathbf{a} \le \mathbf{b}$ when $a_k \le b_k$ for all k; and $\mathbf{a} < \mathbf{b}$ when $a_k < b_k$ for all k. For $\mathbf{a} \le \mathbf{b}$, we will let $[\mathbf{a},\mathbf{b}]$ denote the *n-box* $B = [a_1,b_1]\times[a_2,b_2]\times\cdots\times[a_n,b_n]$, the Cartesian product of n closed intervals. The *vertices* of an n-box B are the points $\mathbf{c} = (c_1, c_2, \cdots, c_n)$ where each c_k is equal to either a_k or b_k. The unit n-cube $\mathbf{I}^n$ is the product $\mathbf{I} \times \mathbf{I} \times \cdots \times \mathbf{I}$. An *n-place real function H* is a function whose domain, Dom$H$, is a subset of $\overline{\mathbf{R}}^n$ and whose range, Ran$H$, is a subset of $\mathbf{R}$. Note that the unit "2-cube" is the unit square $\mathbf{I}^2$, and a "2-box" is a rectangle $[x_1,x_2]\times[y_1,y_2]$ in $\overline{\mathbf{R}}^2$.

Definition 2.10.1. Let S_1, $S_2, \cdots S_n$ be nonempty subsets of $\overline{\mathbf{R}}$, and let H be an n-place real function such that Dom$H$ = $S_1 \times S_2 \times \cdots \times S_n$. Let B = $[\mathbf{a},\mathbf{b}]$ be an n-box all of whose vertices are in Dom$H$. Then the *H-volume of B* is given by

$$V_H(B) = \sum \operatorname{sgn}(\mathbf{c}) H(\mathbf{c}), \tag{2.10.1}$$

where the sum is taken over all vertices $\mathbf{c}$ of B, and sgn($\mathbf{c}$) is given by

$$\operatorname{sgn}(\mathbf{c}) = \begin{cases} 1, & \text{if } c_k = a_k \text{ for an even number of } k\text{'s}, \\ -1, & \text{if } c_k = a_k \text{ for an odd number of } k\text{'s}. \end{cases}$$

Equivalently, the H-volume of an n-box B = $[\mathbf{a},\mathbf{b}]$ is the nth order difference of H on B

$$V_H(B) = \Delta_{\mathbf{a}}^{\mathbf{b}} H(\mathbf{t}) = \Delta_{a_n}^{b_n} \Delta_{a_{n-1}}^{b_{n-1}} \cdots \Delta_{a_2}^{b_2} \Delta_{a_1}^{b_1} H(\mathbf{t}),$$

where we define the n first order differences of an n-place function (such as H) as

$$\Delta_{a_k}^{b_k} H(\mathbf{t}) = H(t_1, \cdots, t_{k-1}, b_k, t_{k+1}, \cdots, t_n) - H(t_1, \cdots, t_{k-1}, a_k, t_{k+1}, \cdots, t_n).$$

Example 2.21. Let H be a 3-place real function with domain $\overline{\mathbf{R}}^3$, and let B be the 3-box $[x_1,x_2]\times[y_1,y_2]\times[z_1,z_2]$. The H-volume of B is

$$\begin{aligned} V_H(B) = {} & H(x_2,y_2,z_2) - H(x_2,y_2,z_1) - H(x_2,y_1,z_2) - H(x_1,y_2,z_2) \\ & + H(x_2,y_1,z_1) + H(x_1,y_2,z_1) + H(x_1,y_1,z_2) - H(x_1,y_1,z_1). \end{aligned}$$ ■

Definition 2.10.2. An n-place real function H is *n-increasing* if $V_H(B) \ge 0$ for all n-boxes B whose vertices lie in Dom H.

Suppose that the domain of an n-place real function H is given by Dom$H$ = $S_1 \times S_2 \times \cdots \times S_n$ where each S_k has a least element a_k. We say that H is *grounded* if $H(\mathbf{t}) = 0$ for all $\mathbf{t}$ in Dom$H$ such that $t_k = a_k$ for at least one k. If each S_k is nonempty and has a greatest element b_k, then we say that H *has margins*, and the *one-dimensional margins* of H are the functions H_k given by Dom$H_k = S_k$ and

$$H_k(x) = H(b_1,\cdots,b_{k-1},x,b_{k+1},\cdots,b_n) \text{ for all } x \text{ in } S_k. \quad (2.10.2)$$

Higher dimensional margins are defined by fixing fewer places in H.

Example 2.22. Let H be the function with domain $[-1,1]\times[0,\infty]\times[0,\pi/2]$ given by

$$H(x,y,z) = \frac{(x+1)(e^y-1)\sin z}{x+2e^y-1}.$$

Then H is grounded because $H(x,y,0) = 0$, $H(x,0,z) = 0$, and $H(-1,y,z) = 0$; H has one-dimensional margins $H_1(x)$, $H_2(y)$, and $H_3(z)$ given by

$$H_1(x) = H(x,\infty,\pi/2) = (x+1)/2,\ H_2(y) = H(1,y,\pi/2) = 1-e^{-y},$$
$$\text{and } H_3(z) = H(1,\infty,z) = \sin z;$$

and H has two-dimensional margins $H_{1,2}(x,y)$, $H_{2,3}(y,z)$, and $H_{1,3}(x,z)$ given by

$$H_{1,2}(x,y) = H(x,y,\pi/2) = \frac{(x+1)(e^{y)}-1)}{x+2e^y-1},$$
$$H_{2,3}(y,z) = H(1,y,z) = (1-e^{-y})\sin z, \text{ and}$$
$$H_{1,3}(x,z) = H(x,\infty,z,) = \frac{(x+1)\sin z}{2}. \quad \blacksquare$$

In the sequel, one-dimensional margins will be simply “margins,” and for $k \ge 2$, we will write “k-margins” for k-dimensional margins.

Lemma 2.10.3. *Let $S_1, S_2,\cdots S_n$ be nonempty subsets of $\overline{\mathbf{R}}$, and let H be a grounded n-increasing function with domain $S_1\times S_2\times\cdots\times S_n$. Then H is nondecreasing in each argument, that is, if $(t_1,\cdots,t_{k-1},x,t_{k+1},\cdots,t_n)$ and $(t_1,\cdots,t_{k-1},y,t_{k+1},\cdots,t_n)$ are in* Dom$H$ *and $x < y$, then $H(t_1,\cdots,t_{k-1},x,t_{k+1},\cdots,t_n) \le H(t_1,\cdots,t_{k-1},y,t_{k+1},\cdots,t_n)$.*

The following lemma, which is the n-dimensional analog of Lemma 2.1.5, is needed to show that n-copulas are uniformly continuous, and in the proof of the n-dimensional version of Sklar’s theorem. Its proof, however, is somewhat more complicated than that of Lemma 2.1.5; see (Schweizer and Sklar 1983) for details.

Lemma 2.10.4. *Let $S_1, S_2,\cdots S_n$ be nonempty subsets of $\overline{\mathbf{R}}$, and let H be a grounded n-increasing function with margins whose domain is*

$S_1 \times S_2 \times \cdots \times S_n$. *Let* $\mathbf{x} = (x_1, x_2, \cdots, x_n)$ *and* $\mathbf{y} = (y_1, y_2, \cdots, y_n)$ *be any points in* $S_1 \times S_2 \times \cdots \times S_n$. *Then*

$$\left|H(\mathbf{x}) - H(\mathbf{y})\right| \le \sum_{k=1}^{n} \left|H_k(x_k) - H_k(y_k)\right|.$$

We are now in a position to define n-dimensional subcopulas and copulas. The definitions are analogous to Definitions 2.2.1 and 2.2.2.

Definition 2.10.5. An *n-dimensional subcopula* (or *n-subcopula*) is a function C' with the following properties:

1. $\operatorname{Dom} C' = S_1 \times S_2 \times \cdots \times S_n$, where each S_k is a subset of $\mathbf{I}$ containing 0 and 1;
2. C' is grounded and n-increasing;
3. C' has (one-dimensional) margins C'_k, $k = 1, 2, \cdots, n$, which satisfy

$$C'_k(u) = u \text{ for all } u \text{ in } S_k. \tag{2.10.3}$$

Note that for every $\mathbf{u}$ in $\operatorname{Dom} C'$, $0 \le C'(\mathbf{u}) \le 1$, so that $\operatorname{Ran} C'$ is also a subset of $\mathbf{I}$.

Definition 2.10.6. An *n-dimensional copula* (or *n-copula*) is an n-subcopula C whose domain is $\mathbf{I}^n$.

Equivalently, an n-copula is a function C from $\mathbf{I}^n$ to $\mathbf{I}$ with the following properties:

1. For every $\mathbf{u}$ in $\mathbf{I}^n$,

$$C(\mathbf{u}) = 0 \text{ if at least one coordinate of } \mathbf{u} \text{ is } 0, \tag{2.10.4a}$$

and

$$\text{if all coordinates of } \mathbf{u} \text{ are 1 except } u_k, \text{ then } C(\mathbf{u}) = u_k; \tag{2.10.4b}$$

2. For every $\mathbf{a}$ and $\mathbf{b}$ in $\mathbf{I}^n$ such that $\mathbf{a} \le \mathbf{b}$,

$$V_C([\mathbf{a}, \mathbf{b}]) \ge 0. \tag{2.10.4c}$$

It is easy to show (see Exercise 2.34) that for any n-copula C, $n \ge 3$, each k-margin of C is a k-copula, $2 \le k < n$.

Example 2.23. (a) Let $C(u,v,w) = w \cdot \min(u,v)$. Then C is a 3-copula, as it is easily seen that C satisfies (2.10.4a) and (2.10.4b), and the C-volume of the 3-box $B = [a_1,b_1] \times [a_2,b_2] \times [a_3,b_3]$ (where $a_k \le b_k$) is

$$V_C(B) = \Delta_{a_3}^{b_3} \Delta_{a_2}^{b_2} \Delta_{a_1}^{b_1} C(u,v,w) = (b_3 - a_3) \Delta_{a_2}^{b_2} \Delta_{a_1}^{b_1} \min(u,v) \ge 0.$$

The 2-margins of C are the 2-copulas $C_{1,2}(u,v) = C(u,v,1) = 1\cdot\min(u,v) = M(u,v)$, $C_{1,3}(u,w) = C(u,1,w) = w\cdot\min(u,1) = \Pi(u,w)$, and $C_{2,3}(v,w) = C(1,v,w) = w\cdot\min(1,v) = \Pi(v,w)$.

(b) Let $C(u,v,w) = \min(u,v) - \min(u,v,1-w)$. The verification that C is a 3-copula is somewhat tedious. Here the 2-margins are $C_{1,2}(u,v) = M(u,v)$, $C_{1,3}(u,w) = u - \min(u,1-w) = W(u,w)$, and $C_{2,3}(v,w) = v - \min(v,1-w) = W(v,w)$. ■

A consequence of Lemma 2.10.4 is the uniform continuity of n-subcopulas (and hence n-copulas):

Theorem 2.10.7. *Let C' be an n-subcopula. Then for every* $\mathbf{u}$ *and* $\mathbf{v}$ *in* $\operatorname{Dom} C'$,

$$\left|C'(\mathbf{v}) - C'(\mathbf{u})\right| \le \sum_{k=1}^{n} \left|v_k - u_k\right|. \tag{2.10.5}$$

Hence C' is uniformly continuous on its domain.

We are now in a position to state the n-dimensional version of Sklar's theorem. To do so, we first define n-dimensional distribution functions:

Definition 2.10.8. An *n-dimensional distribution function* is a function H with domain $\overline{\mathbf{R}}^n$ such that

1. H is n-increasing,
2. $H(\mathbf{t}) = 0$ for all $\mathbf{t}$ in $\overline{\mathbf{R}}^n$ such that $t_k = -\infty$ for at least one k, and $H(\infty,\infty,\cdots,\infty) = 1$.

Thus H is grounded, and because $\operatorname{Dom} H = \overline{\mathbf{R}}^n$, it follows from Lemma 2.10.3 that the one-dimensional margins, given by (2.10.2), of an n-dimensional distribution function are distribution functions, which for $n \ge 3$ we will denote by $F_1, F_2, \cdots, F_n$.

Theorem 2.10.9. Sklar's theorem in *n*-dimensions. *Let H be an n-dimensional distribution function with margins $F_1, F_2, \cdots, F_n$. Then there exists an n-copula C such that for all* $\mathbf{x}$ *in* $\overline{\mathbf{R}}^n$,

$$H(x_1,x_2,\cdots,x_n) = C\big(F_1(x_1),F_2(x_2),\cdots,F_n(x_n)\big). \tag{2.10.6}$$

If $F_1, F_2, \cdots, F_n$ are all continuous, then C is unique; otherwise, C is uniquely determined on $\operatorname{Ran} F_1 \times \operatorname{Ran} F_2 \times \cdots \times \operatorname{Ran} F_n$. *Conversely, if C is an n-copula and $F_1, F_2, \cdots, F_n$ are distribution functions, then the function H defined by* (2.10.6) *is an n-dimensional distribution function with margins $F_1, F_2, \cdots, F_n$.*

The proof of Theorem 2.10.9 proceeds as in the case of two dimensions—one first proves the n-dimensional versions of Lemma 2.3.4 (which is straightforward) and then Lemma 2.3.5, the "extension lemma." The proof of the n-dimensional extension lemma, in which one shows that every n-subcopula can be extended to an n-copula, proceeds via a "multilinear interpolation" of the subcopula to a copula similar to two-dimensional version in (2.3.2). The proof in the n-dimensional case, however, is somewhat more involved (Moore and Spruill 1975; Deheuvels 1978; Sklar 1996).

Corollary 2.10.10. *Let* $H, C, F_1, F_2, \cdots, F_n$ *be as in Theorem* 2.10.9, *and let* $F_1^{(-1)}, F_2^{(-1)}, \cdots, F_n^{(-1)}$ *be quasi-inverses of* $F_1, F_2, \cdots, F_n$, *respectively. Then for any* $\mathbf{u}$ *in* $\mathbf{I}^n$,

$$C(u_1, u_2, \cdots, u_n) = H\left(F_1^{(-1)}(u_1), F_2^{(-1)}(u_2), \cdots, F_n^{(-1)}(u_n)\right). \tag{2.10.7}$$

Of course, the n-dimensional version of Sklar's theorem for random variables (again defined on a common probability space) is similar to Theorem 2.4.1:

Theorem 2.10.11. *Let* $X_1, X_2, \cdots, X_n$ *be random variables with distribution functions* $F_1, F_2, \cdots, F_n$, *respectively, and joint distribution function* H. *Then there exists an* n*-copula* C *such that* (2.10.6) *holds. If* $F_1, F_2, \cdots, F_n$ *are all continuous,* C *is unique. Otherwise,* C *is uniquely determined on* $\operatorname{Ran} F_1 \times \operatorname{Ran} F_2 \times \cdots \times \operatorname{Ran} F_n$.

The extensions of the 2-copulas M, Π, and W to n dimensions are denoted M^n, Π^n, and W^n (a superscript on the name of a copula will denote dimension rather than exponentiation), and are given by:

$$\begin{aligned} M^n(\mathbf{u}) &= \min(u_1, u_2, \cdots, u_n); \\ \Pi^n(\mathbf{u}) &= u_1 u_2 \cdots u_n; \\ W^n(\mathbf{u}) &= \max(u_1 + u_2 + \cdots + u_n - n + 1, 0). \end{aligned} \tag{2.10.8}$$

The functions M^n and Π^n are n-copulas for all $n \geq 2$ (Exercise 2.34), whereas the function W^n fails to be an n-copula for any $n > 2$ (Exercise 2.36). However, we do have the following n-dimensional version of the Fréchet-Hoeffding bounds inequality first encountered in (2.2.5). The proof follows directly from Lemmas 2.10.3 and 2.10.4.

Theorem 2.10.12. *If* C' *is any* n*-subcopula, then for every* $\mathbf{u}$ *in* $\operatorname{Dom} C'$,

$$W^n(\mathbf{u}) \leq C'(\mathbf{u}) \leq M^n(\mathbf{u}). \tag{2.10.9}$$

Although the Fréchet-Hoeffding lower bound W^n is never a copula for $n > 2$, the left-hand inequality in (2.10.9) is "best-possible," in the sense that for any $n \geq 3$ and any $\mathbf{u}$ in $\mathbf{I}^n$, there is an n-copula C such that $C(\mathbf{u}) = W^n(\mathbf{u})$:

Theorem 2.10.13. *For any $n \geq 3$ and any $\mathbf{u}$ in $\mathbf{I}^n$, there exists an n-copula C (which depends on $\mathbf{u}$) such that*

$$C(\mathbf{u}) = W^n(\mathbf{u}).$$

Proof (Sklar 1998). Let $\mathbf{u} = (u_1, u_2, \cdots, u_n)$ be a (fixed) point in $\mathbf{I}^n$ other than $\mathbf{0} = (0,0,\cdots,0)$ or $\mathbf{1} = (1,1,\cdots,1)$. There are two cases to consider.

1. Suppose $0 < u_1 + u_2 + \cdots + u_n \leq n - 1$. Consider the set of points $\mathbf{v} = (v_1, v_2, \cdots, v_n)$ where each v_k is 0, 1, or $t_k = \min\{(n-1)u_k/(u_1 + u_2 + \cdots + u_n), 1\}$. Define an n-place function C' on these points by $C'(\mathbf{v}) = W^n(\mathbf{v})$. It is straightforward to verify that C' satisfies the conditions in Definition 2.10.5 and hence is an n-subcopula. Now extend C' to an n-copula C via a "multilinear interpolation" similar to (2.3.2). Then for each $\mathbf{x}$ in the n-box $[\mathbf{0},\mathbf{t}]$, $\mathbf{t} = (t_1, t_2, \cdots, t_n)$ (which includes $\mathbf{u}$), $C(\mathbf{x}) = W^n(\mathbf{x}) = 0$.

2. Suppose $n - 1 < u_1 + u_2 + \cdots + u_n < n$, and consider the set of points $\mathbf{v} = (v_1, v_2, \cdots, v_n)$ where now each v_k is 0, 1, or $s_k = 1 - (1 - u_k)/[n - (u_i + u_2 + \cdots + u_n)]$. Define an n-place function C' on these points by $C'(\mathbf{v}) = W^n(\mathbf{v})$, and extend to an n-copula C as before. Let $\mathbf{s} = (s_1, s_2, \cdots, s_n)$, then for each $\mathbf{x}$ in the n-box $[\mathbf{s},\mathbf{1}]$ (which includes $\mathbf{u}$), we have $C(\mathbf{x}) = W^n(\mathbf{x}) = x_1 + x_2 + \cdots + x_n - n + 1$. □

The n-copulas M^n and Π^n have characterizations similar to the characterizations of M and Π given in Theorems 2.4.2 and 2.5.4.

Theorem 2.10.14. *For $n \geq 2$, let $X_1, X_2, \cdots, X_n$ be continuous random variables. Then*

1. *$X_1, X_2, \cdots, X_n$ are independent if and only if the n-copula of $X_1, X_2, \cdots, X_n$ is Π^n, and*

2. *each of the random variables $X_1, X_2, \cdots, X_n$ is almost surely a strictly increasing function of any of the others if and only if the n-copula of $X_1, X_2, \cdots, X_n$ is M^n.*

Exercises

2.34 (a) Show that the $(n-1)$-margins of an n-copula are $(n-1)$-copulas. [Hint: consider n-boxes of the form $[a_1,b_1]\times\cdots\times[a_{k-1},b_{k-1}]\times[0,1]\times[a_{k+1},b_{k+1}]\times\cdots\times[a_n,b_n]$.]
(b) Show that if C is an n-copula, $n \geq 3$, then for any k, $2 \leq k < n$, all $\binom{n}{k}$ k-margins of C are k-copulas.

2.35 Let M^n and Π^n be the functions defined in (2.10.4), and let $[\mathbf{a},\mathbf{b}]$ be an n-box in $\mathbf{I}^n$. Prove that

$$V_{M^n}([\mathbf{a},\mathbf{b}]) = \max\left(\min(b_1,b_2,\cdots,b_n) - \max(a_1,a_2,\cdots,a_n),0\right)$$

and

$$V_{\Pi^n}([\mathbf{a},\mathbf{b}]) = (b_1-a_1)(b_2-a_2)\cdots(b_n-a_n),$$

and hence conclude that M^n and Π^n are n-copulas for all $n \geq 2$.

2.36 Show that

$$V_{W^n}([\mathbf{1/2},\mathbf{1}]) = 1-(n/2),$$

where $\mathbf{1}$ = (1,1, $\cdots$,1) and $\mathbf{1/2}$ = (1/2,1/2, $\cdots$,1/2), and hence W^n fails to be an n-copula whenever $n > 2$.

2.37 Let $X_1,X_2,\cdots,X_n$ be continuous random variables with copula C and distribution functions $F_1,F_2,\cdots,F_n$, respectively. Let $X_{(1)}$ and $X_{(n)}$ denote the extreme *order statistics* for $X_1,X_2,\cdots,X_n$ (i.e., $X_{(1)}$ = $\min(X_1,X_2,\cdots,X_n)$ and $X_{(n)}$ = $\max(X_1,X_2,\cdots,X_n)$) [cf. Exercise 2.16]. Prove that the distribution functions $F_{(1)}$ and $F_{(n)}$, respectively, of $X_{(1)}$ and $X_{(n)}$ satisfy

$$\max\left(F_1(t),F_2(t),\cdots,F_n(t)\right) \leq F_{(1)}(t) \leq \min\left(\textstyle\sum_{k=1}^n F_k(t),1\right)$$

and

$$\max\left(\textstyle\sum_{k=1}^n F_k(t) - n + 1,0\right) \leq F_{(n)}(t) \leq \min\left(F_1(t),F_2(t),\cdots,F_n(t)\right).$$

http://www.springer.com/978-0-387-28659-4

An Introduction to Copulas
Nelsen, R.B.
2006, XIV, 272 p., Hardcover
ISBN: 978-0-387-28659-4