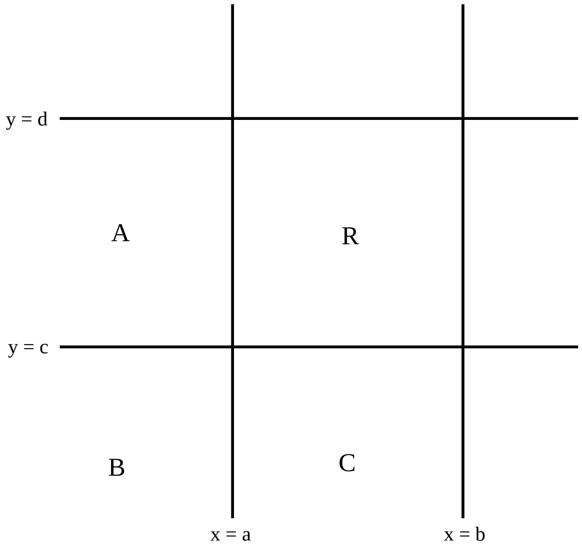
Some Additional Notes on Bivariate Distributions



Define the following sets based on regions depicted above

$$E_R = \{s \in S : a < X(s) \leq b \text{ and } c < Y(s) \leq d\}$$
$$E_A = \{s \in S : X(s) \leq a \text{ and } c < Y(s) \leq d\}$$
$$E_B = \{s \in S : X(s) \leq a \text{ and } Y(s) \leq c\}$$
$$E_C = \{s \in S : a < X(s) \leq b \text{ and } Y(s) \leq c\},$$

and let the joint CDF, $F_{X,Y}$, and joint PDF, $f_{X,Y}$, be defined by

$$P(\{X \leq x\} \cap \{Y \leq y\}) = F_{X,Y}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(u,v)\, dvdu.$$

First, let's show that

$$P(E_R) = F_{X,Y}(b,d) - F_{X,Y}(a,d) + F_{X,Y}(a,c) - F_{X,Y}(b,c). \tag{1}$$

Clearly the sets E_R, E_A, E_B and E_C are disjoint, and their union is

$$\begin{aligned} & E_R \cup E_A \cup E_B \cup E_C = \{X \leq b\} \cap \{Y \leq d\} \\ \Rightarrow \quad & P(E_R) + P(E_A) + P(E_B) + P(E_C) = P(\{X \leq b\} \cap \{Y \leq d\}) = F_{X,Y}(b,d). \end{aligned}$$

Observe that $P(E_B) = F_{X,Y}(a,c)$. Now, the probability $P(E_A)$, (recall from above that $E_A = \{X \leq a\} \cap \{c < Y \leq d\}$), can be determined as follows.

$$\begin{aligned} & (\{X \leq a\} \cap \{c < Y \leq d\}) \cup (\{X \leq a\} \cap \{Y \leq c\}) \\ & = \{X \leq a\} \cap (\{c < Y \leq d\} \cup \{Y \leq c\}) \\ & = \{X \leq a\} \cap \{Y \leq d\}, \end{aligned}$$

and since the left hand side is a disjoint union we find that

$$\begin{aligned} & P(E_A) + P(\{X \le a\} \cap \{Y \le c\}) = P(\{X \le a\} \cap \{Y \le d\}) \\ \Rightarrow \quad & P(E_A) = F_{X,Y}(a,d) - F_{X,Y}(a,c). \end{aligned}$$

A similar argument will show $P(E_C) = F_{X,Y}(b,c) - F_{X,Y}(a,c)$. Putting together these formulae for $P(E_A)$, $P(E_B)$ and $P(E_C)$ will establish the formula given in (1).

- Formula (1) may be used freely on your homework.

Second, let's show that

$$P(E_R) = P((a < X \le b) \cap (c < Y \le d)) = \int_a^b \int_c^d f_{X,Y}(u,v)\, dvdu. \tag{2}$$

This follows from (1) by observing

$$\begin{aligned} P(E_R) =& F_{X,Y}(b,d) - F_{X,Y}(a,d) + F_{X,Y}(a,c) - F_{X,Y}(b,c) \\ =& \left(\int_{-\infty}^{b} \int_{-\infty}^{d} - \int_{-\infty}^{a} \int_{-\infty}^{d} \right) f_{X,Y}(u,v)\, dvdu \\ &+ \left(\int_{-\infty}^{a} \int_{-\infty}^{c} - \int_{-\infty}^{b} \int_{-\infty}^{c} \right) f_{X,Y}(u,v)\, dvdu \\ =& \int_a^b \int_{-\infty}^{d} f_{X,Y}(u,v)\, dvdu - \int_a^b \int_{-\infty}^{c} f_{X,Y}(u,v)\, dvdu \\ =& \int_a^b \left(\int_{-\infty}^{d} - \int_{-\infty}^{c} \right) f_{X,Y}(u,v)\, dvdu = \int_a^b \int_c^d f_{X,Y}(u,v)\, dvdu. \end{aligned}$$

- Formula (2) may also be used freely on your homework.

Finally, this result for a single rectangle can be generalized by using the *additiveity* property of the probability measure and the integral. Let $R_k = (a_k, b_k] \times (c_k, d_k]$ denote a rectangle in $\mathbb{R}^2$, and define an associated event

$$E_k = \{s \in S : a_k < X(s) \le b_k \text{ and } c_k < Y(s) \le d_k\} = \{s \in S : (X(s), Y(s)) \in R_k\}.$$

Now, for a given region $R = \bigcup_{k=1} R_k \subseteq \mathbb{R}^2$ composed of disjoint rectangles R_k, one can show the events $E_1, E_2, \ldots$ are disjoint subsets in S, and

$$E = \bigcup_{k=1} E_k = \{s \in S : (X(s), Y(s)) \in R\}$$

$$\begin{aligned} \Rightarrow \quad & P((X,Y) \in R) = P(E) = \sum_{k=1} P(E_k) \\ & = \sum_{k=1} \iint_{(u,v)\in R_k} f_{X,Y}(u,v)\, dvdu = \iint_{(u,v)\in R} f_{X,Y}(u,v)\, dvdu, \end{aligned} \tag{3}$$

where $P((X,Y) \in R)$ is shorthand notation for $P(\{s \in S : (X(s), Y(s)) \in R\})$. Under reasonable hypotheses (not discuss here), formula (3) can be extended to regions $R \subseteq \mathbb{R}^2$ which are more general than a countable union of such rectangles.