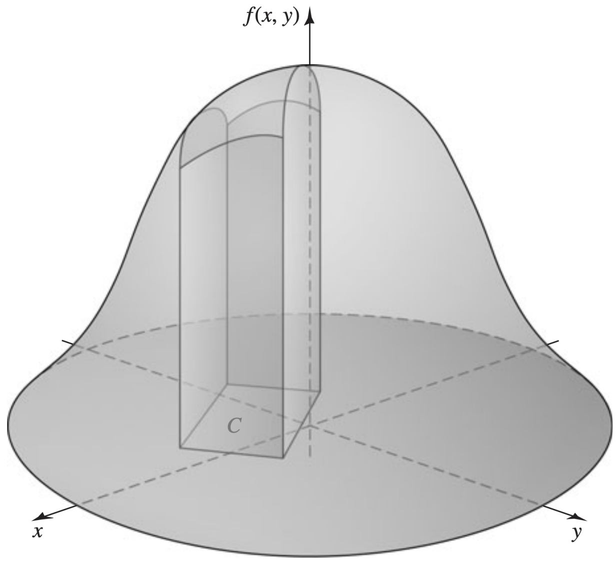

Figure 2.2: A bivariate density is a surface in three dimensional space, and probabilities are volumes beneath the surface. The total volume beneath the surface must be one.

2.2 Jointly absolutely continuous random pairs

Definition 2.6. If (X, Y) are jointly absolutely continuous, then their joint distribution can be summarized with a **joint probability density function** (**joint pdf**). This is a bivariate function $f_{XY} : \mathbb{R}^2 \to \mathbb{R}$ with the following properties:

- $f_{XY}(x, y) \geq 0$;
- $P\left((X, Y) \in \mathbb{R}^2\right) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{XY}(x, y)\,\mathrm{d}x\mathrm{d}y = 1$.

Given the joint pdf, you can compute the joint distribution by integrating:

$$P((X, Y) \in C) = \iint_C f_{XY}(x, y)\,\mathrm{d}x\mathrm{d}y, \quad C \subseteq \mathbb{R}^2.$$

In the univariate case, probability corresponds to area under the density curve. In the bivariate case, probability corresponds to volume under the density *surface*, as in Figure 2.2.

Definition 2.7. If (X, Y) are jointly absolutely continuous, then each random variable X and Y has a **marginal probability density function** (**marginal pdf**) that describes the distribution

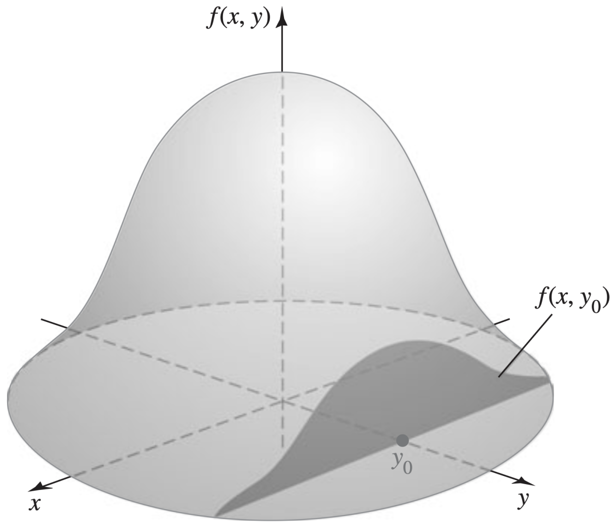

Figure 2.3: The conditional density is a renormalized cross-section of the joint density, where we hold the argument of the conditioning variable constant.

of that random variable on its own:

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y)\,\mathrm{d}y \tag{2.13}$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x, y)\,\mathrm{d}x. \tag{2.14}$$

To compute f_X, for example, we say that we *integrate y out* of the joint density.

To see where the marginal density formulas come from, consider the marginal CDF of one of the variables:

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x}\int_{-\infty}^{\infty} f_{XY}(t, y)\mathrm{d}y\mathrm{d}t.$$

The marginal PDF is just the derivative, so the fundamental theorem of calculus tells us that

$$f_X(x) = F_X'(x) = \frac{\mathrm{d}}{\mathrm{d}x}\int_{-\infty}^{x}\int_{-\infty}^{\infty} f_{XY}(t, y)\mathrm{d}y\mathrm{d}t = \int_{-\infty}^{\infty} f_{XY}(x, y)\mathrm{d}y.$$

Definition 2.8. The **conditional probability density function** (**conditional pdf**) of X or Y describes its revised behavior *conditional* on the event $Y = y$ or $X = x$, respectively. They are defined to be

$$f_{X|Y}(x \mid y) = \frac{f_{XY}(x, y)}{f_Y(y)} \tag{2.15}$$

$$f_{Y|X}(y \mid x) = \frac{f_{XY}(x, y)}{f_X(x)}. \tag{2.16}$$

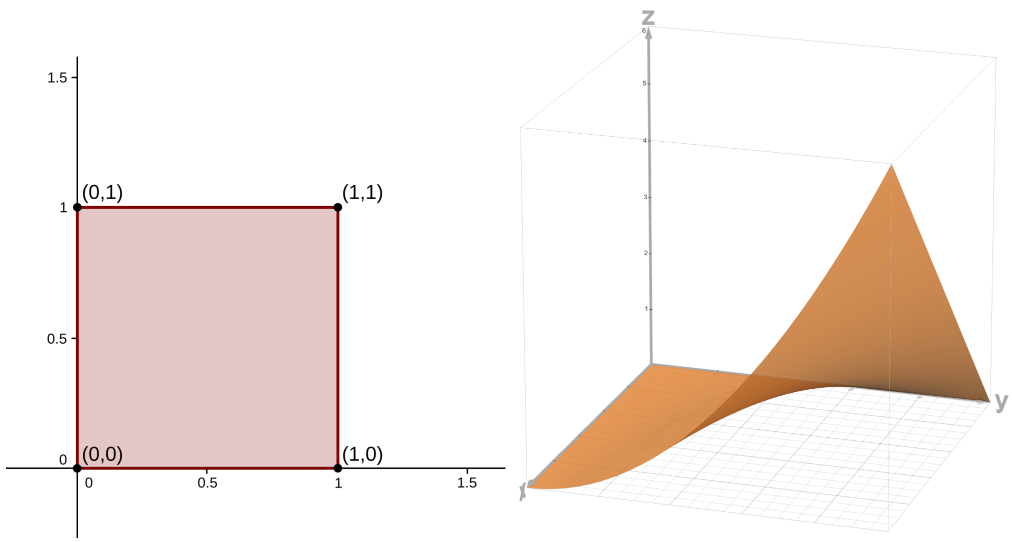

Figure 2.4: In Example 2.4, the joint range is the unit square on the left, and the joint density is the surface on the right.

So $f_{X|Y}(x \mid y)$, for example, is the cross-section of f_{XY} that holds the second argument fixed at y but lets x vary. The denominator renormalizes to ensure that the new *conditional* density integrates to 1 in x. See Figure 2.3 for a cartoon.

Theorem 2.4. (**Marginal-conditional decomposition**) From Definition 2.8, we get "two identities for the price of one:"

$$f_{XY}(x, y) = f_{X|Y}(x \mid y) f_Y(y) = f_{Y|X}(y \mid x) f_X(x). \tag{2.17}$$

Corollary 2.5. (**Hierarchical representation**) In order to specify a joint distribution for (X, Y), you can do so *hierarchically*:

$$\begin{aligned} X &\sim P_X \\ Y \mid X &\sim P_{Y|X}. \end{aligned}$$

Theorem 2.4 guarantees that we can stitch these two pieces together to get a valid joint distribution.

Theorem 2.6. (**Bayes' theorem, again**) Combining the definition of conditional density with the marginal-conditional decomposition gives a version of Bayes' theorem for the conditional PDFs:

$$f_{X|Y}(x \mid y) = \frac{f_{XY}(x, y)}{f_Y(y)} = \frac{f_{Y|X}(y \mid x) f_X(x)}{f_Y(y)} \tag{2.18}$$

$$f_{Y|X}(y \mid x) = \frac{f_{XY}(x, y)}{f_X(x)} = \frac{f_{X|Y}(x \mid y) f_Y(y)}{f_X(x)}. \tag{2.19}$$

Example 2.4. Let X and Y have joint density

$$f_{XY}(x, y) = \begin{cases} 6xy^2, & x, y \in (0, 1) \\ 0 & \text{else.} \end{cases}$$

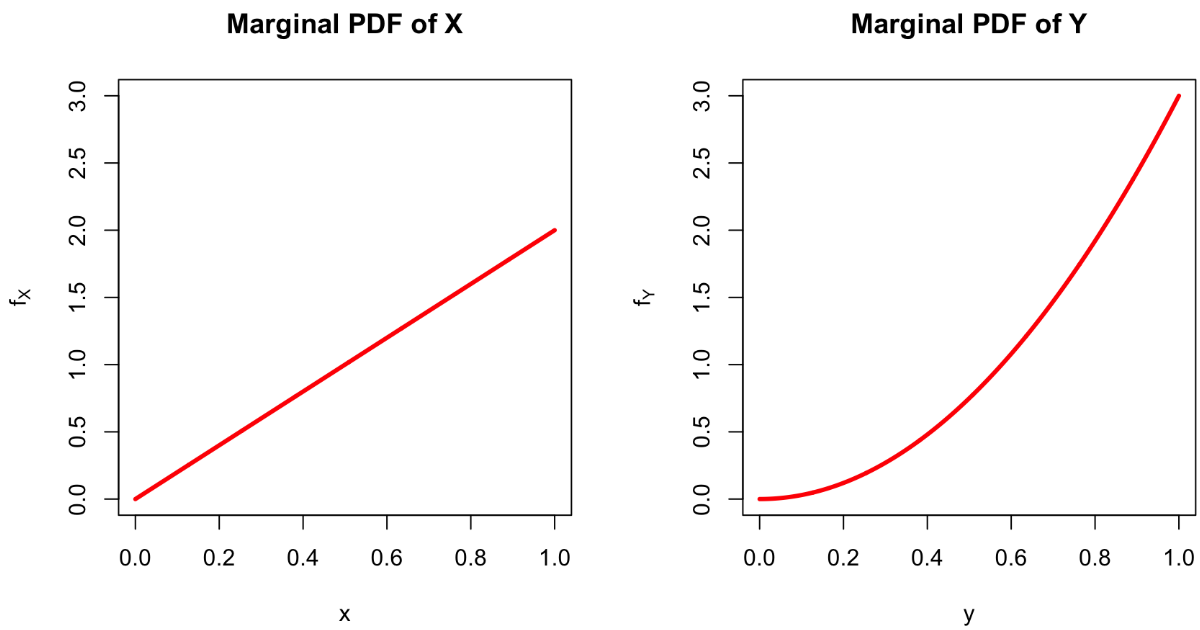

Figure 2.5: Marginal densities in Example 2.4

The joint range is $\text{Range}(X, Y) = \text{supp}(f_{XY}) = (0, 1) \times (0, 1)$, the unit square in the plane (Figure 2.4). The marginal densities are

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y)\,\mathrm{d}y = \int_0^1 6xy^2\,\mathrm{d}y = [2xy^3]_0^1 = 2x, \quad x \in (0, 1).$$
$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x, y)\,\mathrm{d}x = \int_0^1 6xy^2\,\mathrm{d}x = [3x^2y^2]_0^1 = 3y^2, \quad y \in (0, 1).$$

The conditional densities are

$$f_{X|Y}(x \mid y) = \frac{f_{XY}(x, y)}{f_Y(y)} = \frac{6xy^2}{3y^2} = 2x, \quad x \in (0, 1)$$
$$f_{Y|X}(y \mid x) = \frac{f_{XY}(x, y)}{f_X(x)} = \frac{6xy^2}{2x} = 3y^2, \quad y \in (0, 1).$$

We see that $f_{X|Y} = f_X$ and $f_{Y|X} = f_Y$, meaning that conditioning does not actually teach us anything. This motivates a definition.

Definition 2.9. A jointly absolutely continuous pair (X, Y) are **independent** if

$$f_{XY}(x, y) = f_X(x) f_Y(y) \quad \text{for all } (x, y) \in \mathbb{R}^2. \tag{2.20}$$

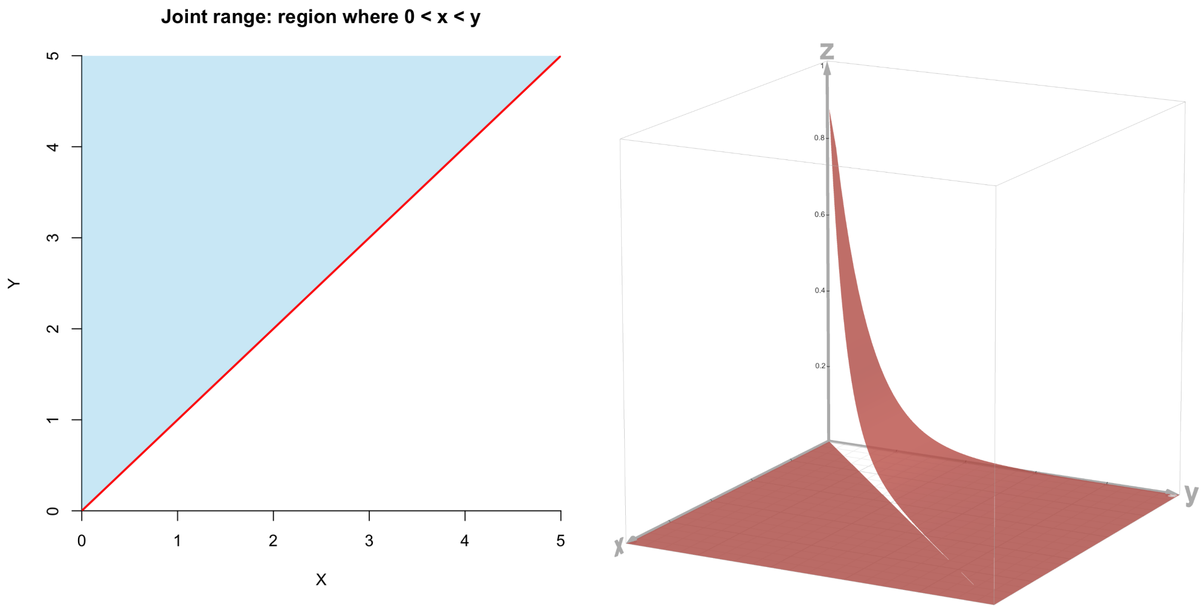

Figure 2.6: The joint density in Example 2.5.

Alternatively but equivalently, X and Y are independent if

$$f_{X|Y}(x\,|\,y) = f_X(x) \quad \text{for all } (x,\ y) \in \mathbb{R}^2 \tag{2.21}$$

$$f_{Y|X}(y\,|\,x) = f_Y(y) \quad \text{for all } (x,\ y) \in \mathbb{R}^2. \tag{2.22}$$

Example 2.5. Consider $(X,\ Y)$ with joint density

$$f_{XY}(x,\ y) = \begin{cases} e^{-y} & 0 < x < y < \infty \\ 0 & \text{else.} \end{cases}$$

The joint range and the density surface are displayed in Figure 2.6. Let's compute all of the marginals and conditionals:

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f_{XY}(x,\ y)\,\mathrm{d}y = \int_x^{\infty} e^{-y}\,\mathrm{d}y = [-e^{-y}]_x^{\infty} = e^{-x}, && 0 < x \\ f_{Y|X}(y\mid x) &= \frac{f_{XY}(x,\ y)}{f_X(x)} = \frac{e^{-y}}{e^{-x}} = e^{-(y-x)}, && x < y \\ f_Y(y) &= \int_{-\infty}^{\infty} f_{XY}(x,\ y)\,\mathrm{d}x = \int_0^{y} e^{-y}\,\mathrm{d}x = e^{-y}\int_0^{y}\mathrm{d}x = ye^{-y}, && 0 < y \\ f_{X|Y}(x\mid y) &= \frac{f_{XY}(x,\ y)}{f_Y(y)} = \frac{e^{-y}}{ye^{-y}} = \frac{1}{y}, && 0 < x < y. \end{aligned}$$

These are all familiar distributions:

$$\begin{aligned} X &\sim \text{Exponential}(1) \\ Y &\sim \text{Gamma}(2,\ 1) \\ X \mid Y = y &\sim \text{Unif}(0,\ y). \end{aligned}$$

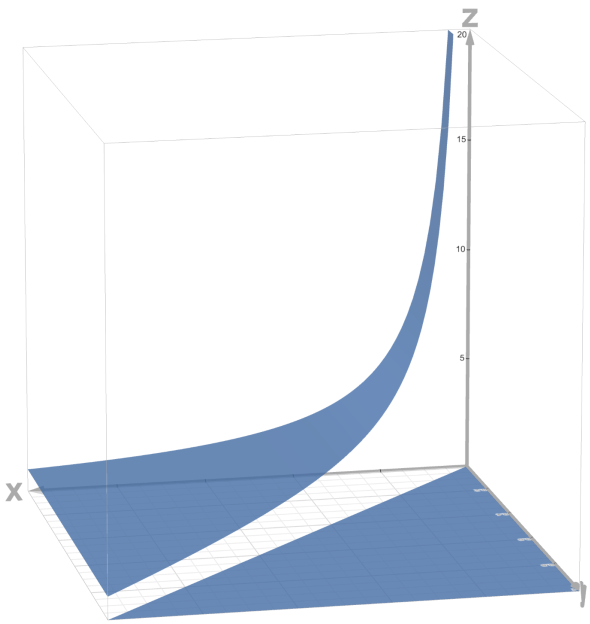

Figure 2.7: Joint density in Example 2.6.

Conditional on $X = x$, Y is Exponential(1), but shifted right by the amount x.

Example 2.6. Consider a joint distribution for (X, Y) written hierarchically:

$$X \sim \text{Unif}(0, 1)$$
$$Y \mid X = x \sim \text{Unif}(0, x).$$

So

$$\begin{aligned} f_X(x) &= 1 && 0 \leq x \leq 1 \\ f_{Y|X}(y \mid x) &= \frac{1}{x} && 0 \leq y \leq x \end{aligned}$$

$$\begin{aligned} f_{XY}(x, y) &= f_{Y|X}(y \mid x) f_X(x) \\ &= \frac{1}{x} && 0 \leq y \leq x \leq 1. \end{aligned}$$

The joint density surface is displayed in Figure 2.7. The marginal density of Y is

$$\begin{aligned} f_Y(y) &= \int_0^1 f_{XY}(x,\ y)\mathrm{d}x \\ &= \int_0^y f_{XY}(x,\ y)\mathrm{d}x + \int_y^1 f_{XY}(x,\ y)\mathrm{d}x \\ &= \int_0^y 0\,\mathrm{d}x + \int_y^1 \frac{1}{x}\mathrm{d}x \\ &= \int_y^1 \frac{1}{x}\mathrm{d}x \\ &= [\ln x]_y^1 \\ &= \ln 1 - \ln y \\ &= -\ln y, && 0 < y < 1. \end{aligned}$$

The conditional density is

$$f_{X|Y}(x \mid y) = \frac{f_{XY}(x,\ y)}{f_Y(y)} = \frac{1/x}{-\ln y}, \quad 0 \leq y \leq x \leq 1.$$

To compute the conditional expectation of X given $Y = y$, we apply the definition of expected value, but we use the *conditional* density:

$$\begin{aligned} E(X \mid Y = y) &= \int_0^1 x f_{X|Y}(x \mid y)\,\mathrm{d}x \\ &= \int_y^1 x f_{X|Y}(x \mid y)\,\mathrm{d}x \\ &= \int_y^1 x \frac{1/x}{-\ln y}\,\mathrm{d}x \\ &= \frac{1}{-\ln y}\int_y^1 \mathrm{d}x \\ &= \frac{1-y}{-\ln y}. \end{aligned}$$