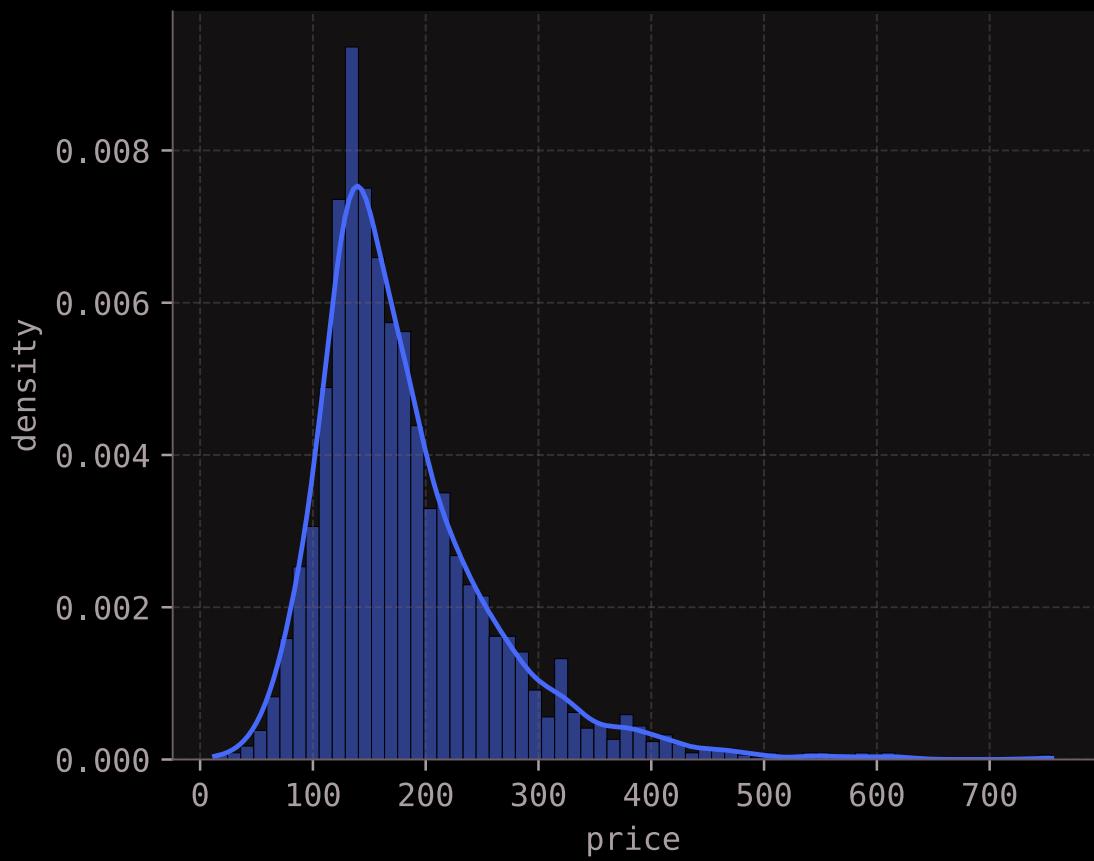
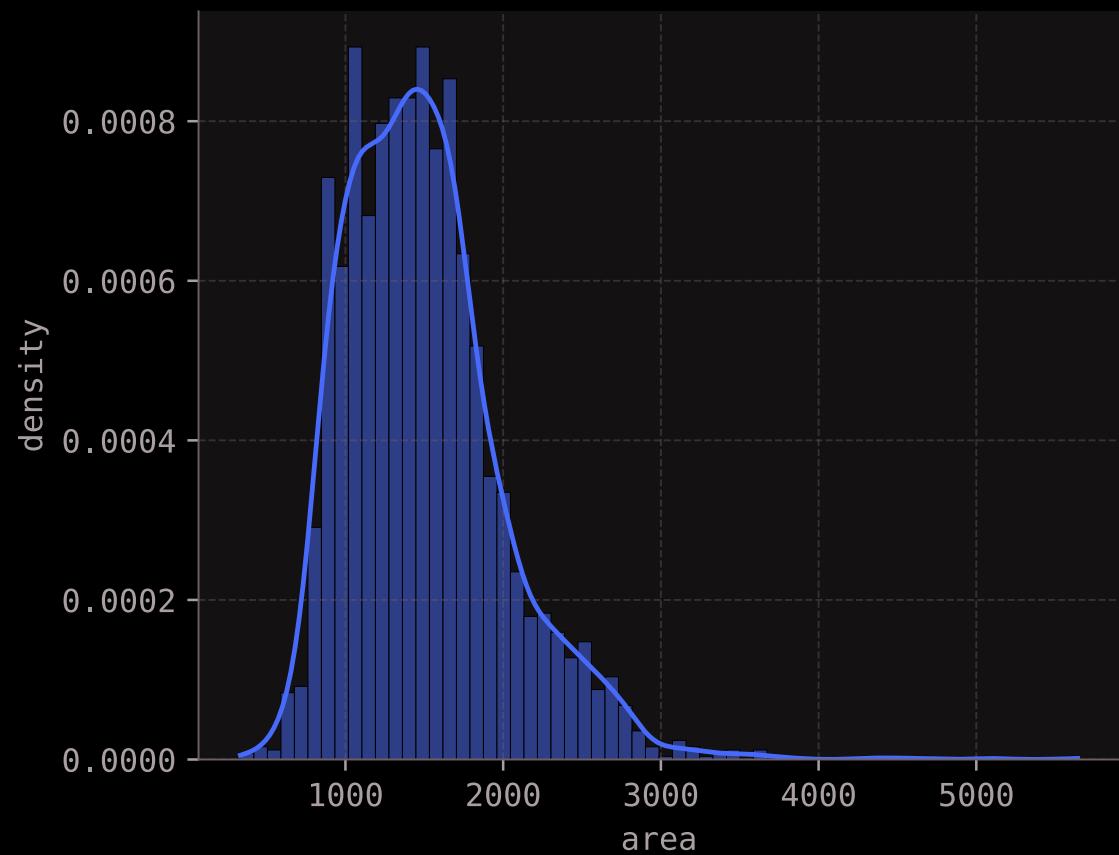
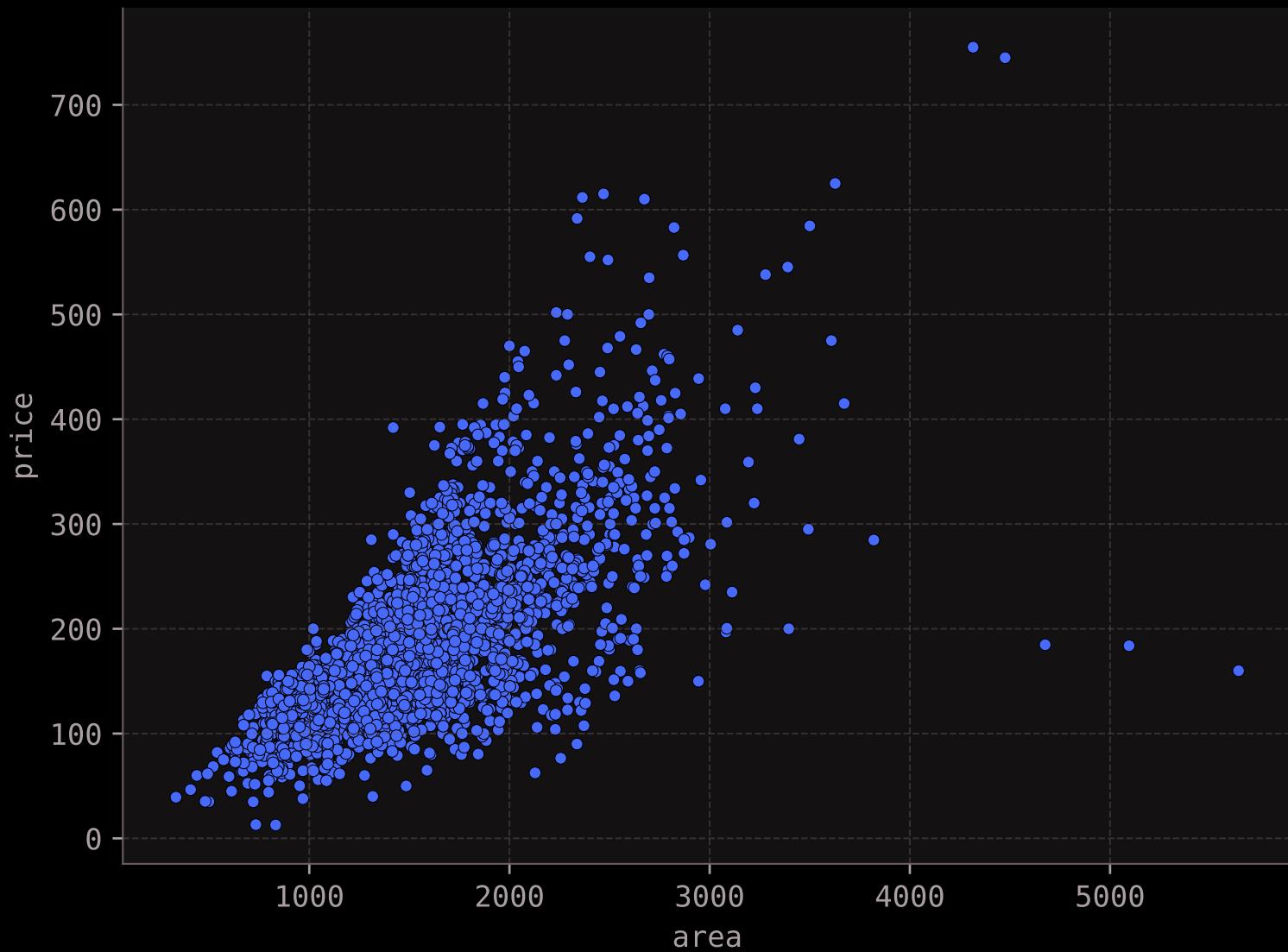


7. Random vectors

7.1. Motivation





7.2. 2-dimensional random vectors

 **Definition 7.1**

Let S be a probability space. A *2-dimensional random vector* is a function

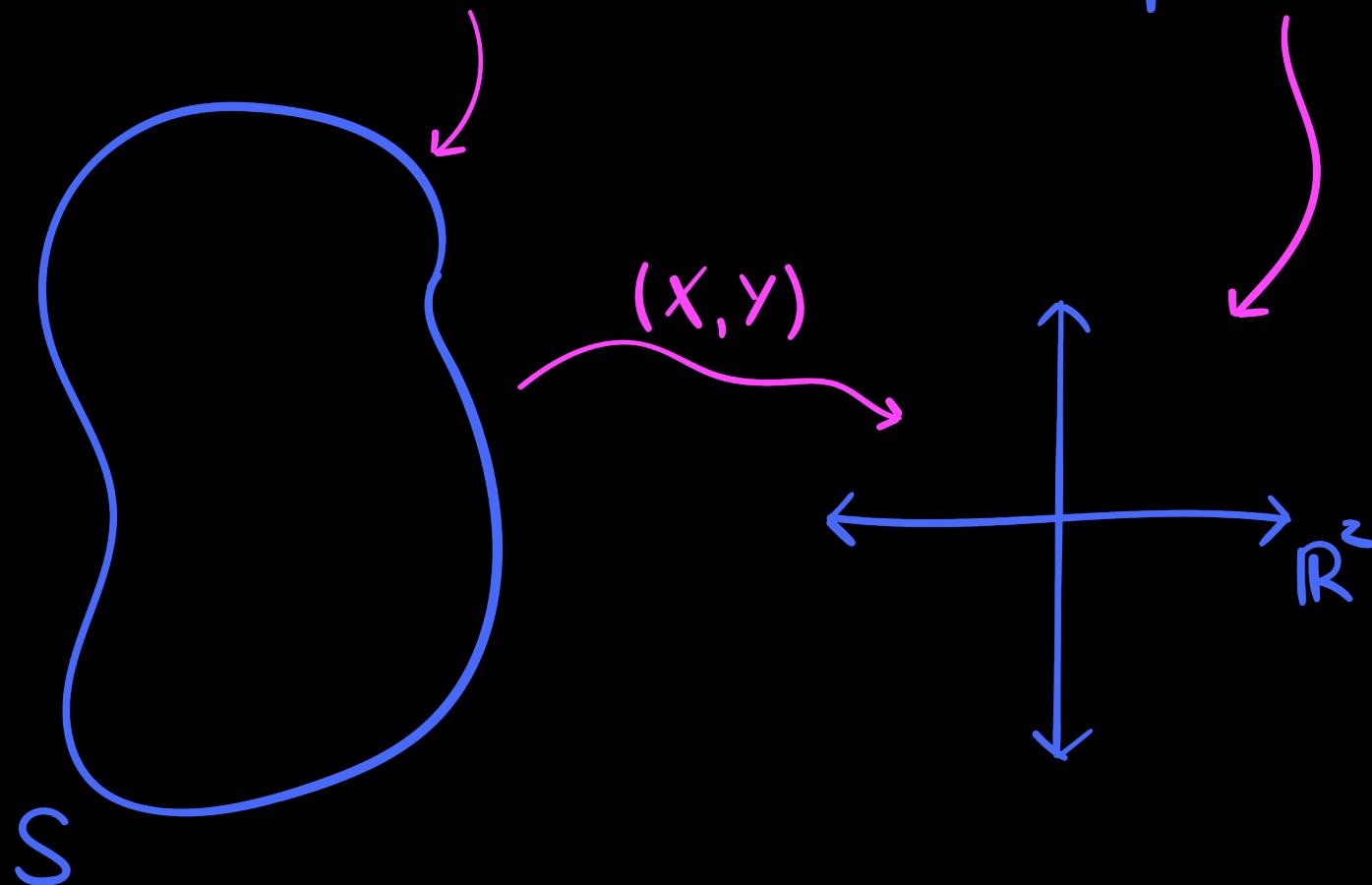
$$\mathbf{X} : S \rightarrow \mathbb{R}^2.$$

Thus, we may write $\mathbf{X}(s) = (X_1(s), X_2(s))$ for each sample point $s \in S$, where

$$X_1 : S \rightarrow \mathbb{R} \quad \text{and} \quad X_2 : S \rightarrow \mathbb{R}$$

are random variables. When we do so, the random variables X_1 and X_2 are called the *components* of the random vector \mathbf{X} .

P measures
probability here



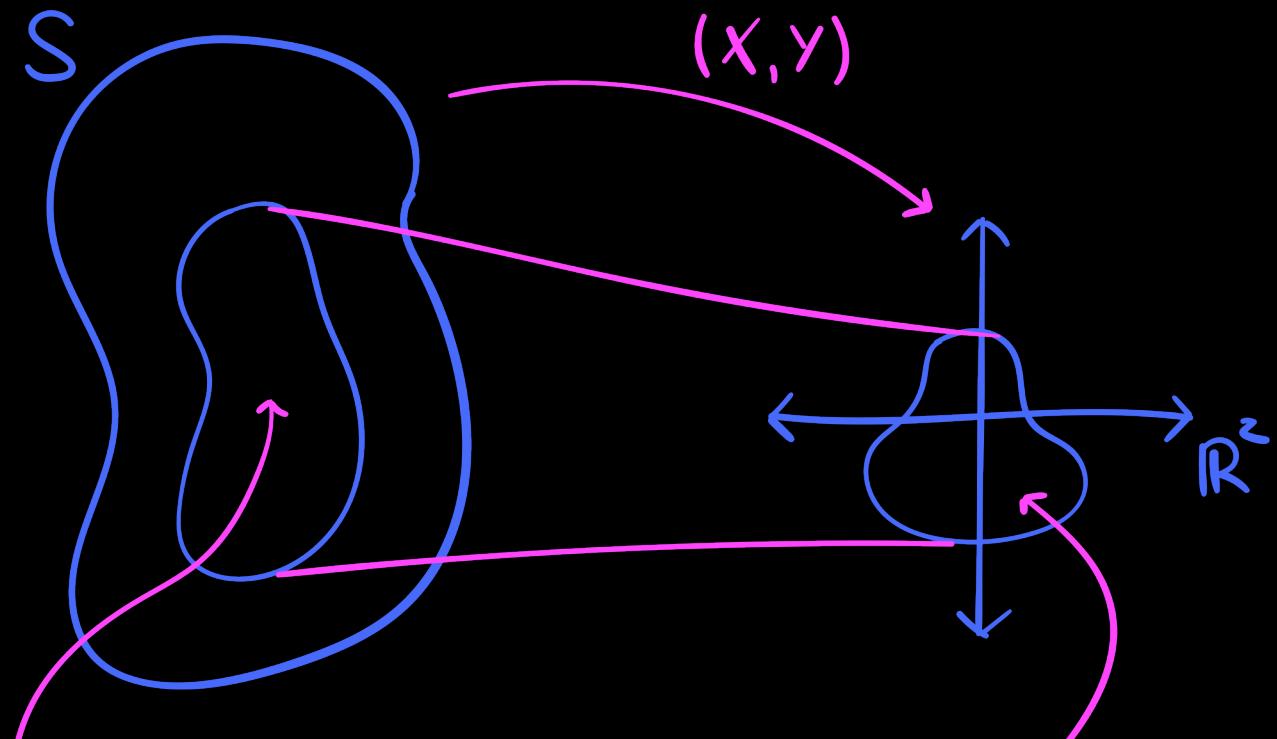


Definition 7.2

Let $(X, Y) : S \rightarrow \mathbb{R}^2$ be a 2-dimensional random vector on a probability space S with probability measure P . We define the *probability measure* of (X, Y) , denoted P_{XY} , via the formula

$$P_{XY}(C) = P(\{s \in S : (X(s), Y(s)) \in C\}), \quad (7.1)$$

for all events $C \subset \mathbb{R}^2$. The probability measure P_{XY} is also called the *joint distribution* or the *bivariate distribution* of X and Y .



$$\{s \in S : (X(s), Y(s)) \in C\}$$

\uparrow \uparrow

these have equal
probabilities by definition!



Problem Prompt

Do problem 1 on the worksheet.

Definition 7.3

Let (X, Y) be a 2-dimensional random vector.

- We shall say (X, Y) is *discrete*, or that X and Y are *jointly discrete*, if the joint probability distribution P_{XY} is discrete. In other words, we require that there exists a *joint probability mass function* $p(x, y)$ such that

$$P((X, Y) \in C) = \sum_{(x,y) \in C} p(x, y)$$

for all events $C \subset \mathbb{R}^2$.

- We shall say (X, Y) is *continuous*, or that X and Y are *jointly continuous*, if the joint probability distribution P_{XY} is continuous. In other words, we require that there exists a *joint probability density function* $f(x, y)$ such that

$$P((X, Y) \in C) = \iint_C f(x, y) \, dx \, dy$$

for all events $C \subset \mathbb{R}^2$.

Theorem 7.1

Let (X, Y) be a 2-dimensional random vector.

1. The random vector (X, Y) is discrete if and only if both X and Y are discrete.
2. If (X, Y) is continuous, then X and Y are both continuous. However, it does *not* necessarily follow that if both X and Y are continuous, then so too is (X, Y) .



Problem Prompt

Do problems 2-4 on the worksheet.

7.3. Bivariate distribution functions

Definition 7.4

Let (X, Y) be a 2-dimensional random vector. The *distribution function* of (X, Y) is the function $F : \mathbb{R}^2 \rightarrow \mathbb{R}$ defined by

$$F(x, y) = P(X \leq x, Y \leq y).$$

In particular:

1. If (X, Y) is discrete with probability mass function $p(x, y)$, then

$$F(x, y) = \sum_{x^* \leq x, y^* \leq y} p(x^*, y^*).$$

2. If (X, Y) is continuous with probability density function $f(x, y)$, then

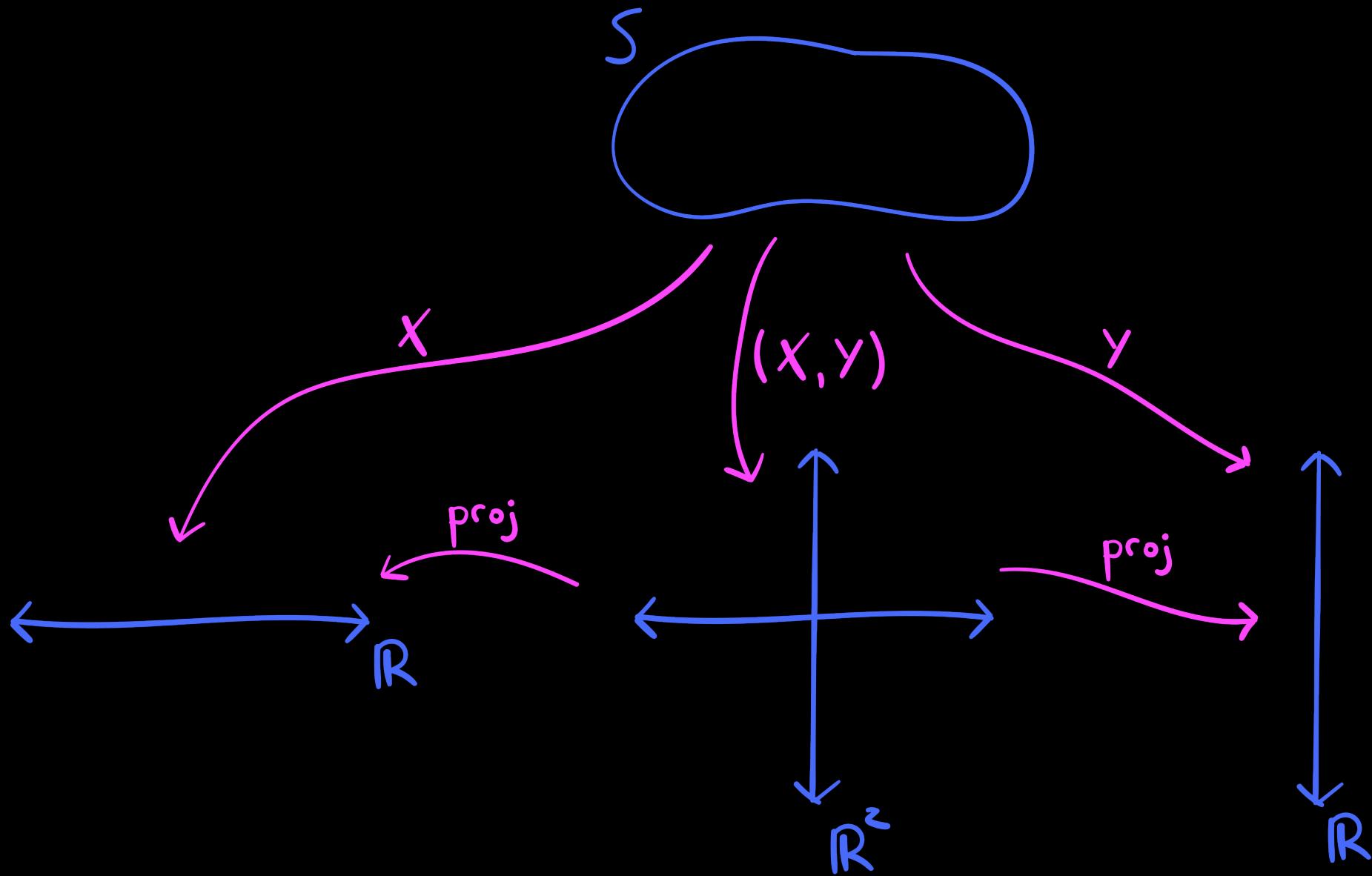
$$F(x, y) = \int_{-\infty}^y \int_{-\infty}^x f(x^*, y^*) dx^* dy^*.$$

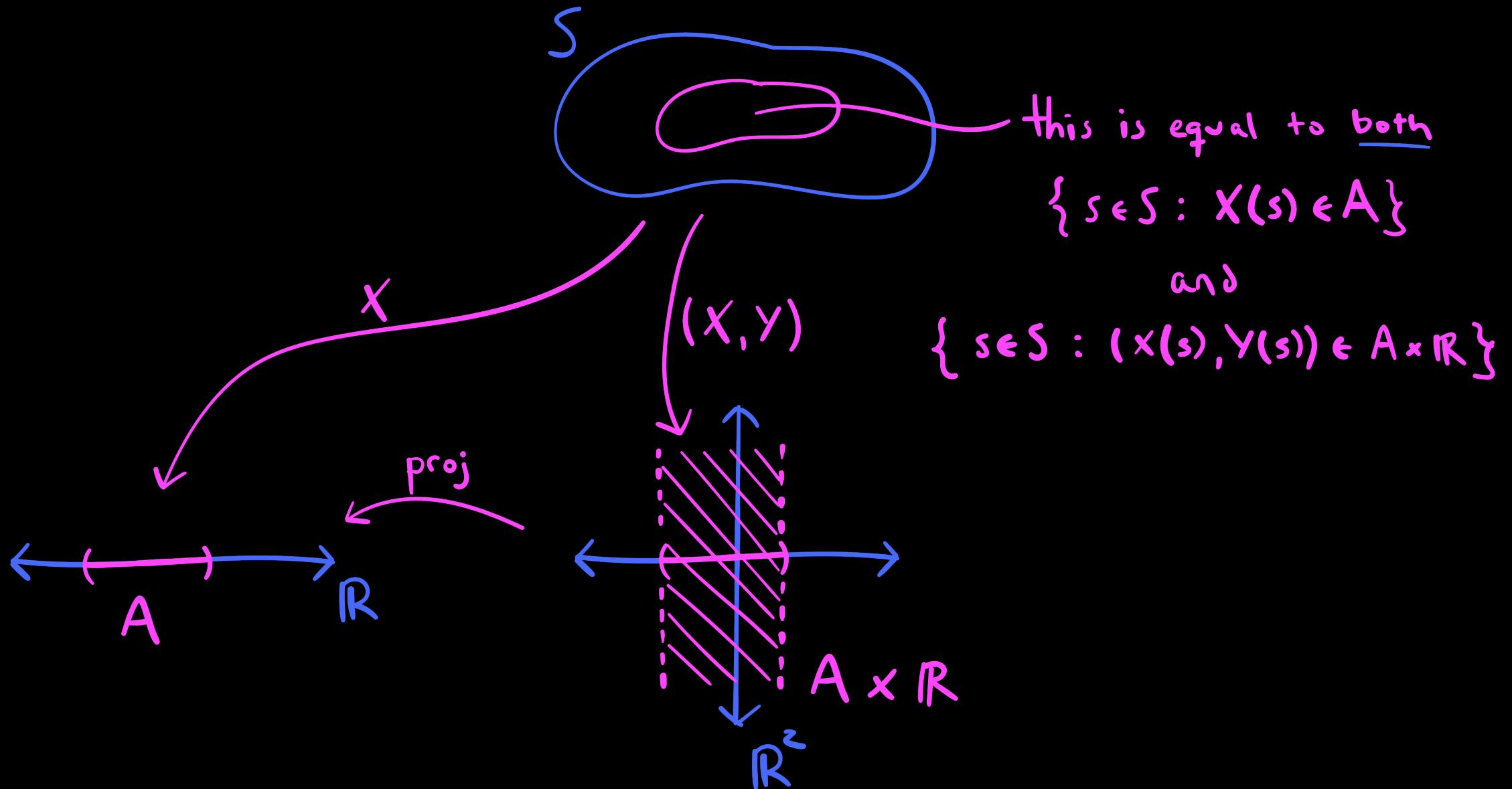


Problem Prompt

Do problem 5 on the worksheet.

7.4. Marginal distributions





Theorem 7.2

Let (X, Y) be a 2-dimensional random vector with induced probability measure P_{XY} . Then the measures P_X and P_Y may be obtained via the formulas

$$P_X(A) = P_{XY}(A \times \mathbb{R}) \quad \text{and} \quad P_Y(B) = P_{XY}(\mathbb{R} \times B)$$

for all events $A, B \subset \mathbb{R}$. In particular:

1. If (X, Y) is discrete with probability mass function $p(x, y)$, then

$$P(X \in A) = \sum_{x \in A} \sum_{y \in \mathbb{R}} p(x, y) \quad \text{and} \quad P(Y \in B) = \sum_{y \in B} \sum_{x \in \mathbb{R}} p(x, y).$$

2. If (X, Y) is continuous with probability density function $f(x, y)$, then

$$P(X \in A) = \int_A \int_{-\infty}^{\infty} f(x, y) \, dy \, dx$$

and

$$P(Y \in B) = \int_B \int_{-\infty}^{\infty} f(x, y) \, dx \, dy.$$



Definition 7.5

Let (X, Y) be a 2-dimensional random vector. Then the distributions P_X and P_Y are called the *marginal distributions* of (X, Y) .

Theorem 7.3

Let (X, Y) be a 2-dimensional random vector.

1. If (X, Y) is discrete with probability mass function $p(x, y)$, then both X and Y are discrete with probability mass functions given by

$$p_X(x) = \sum_{y \in \mathbb{R}} p(x, y) \quad \text{and} \quad p_Y(y) = \sum_{x \in \mathbb{R}} p(x, y).$$

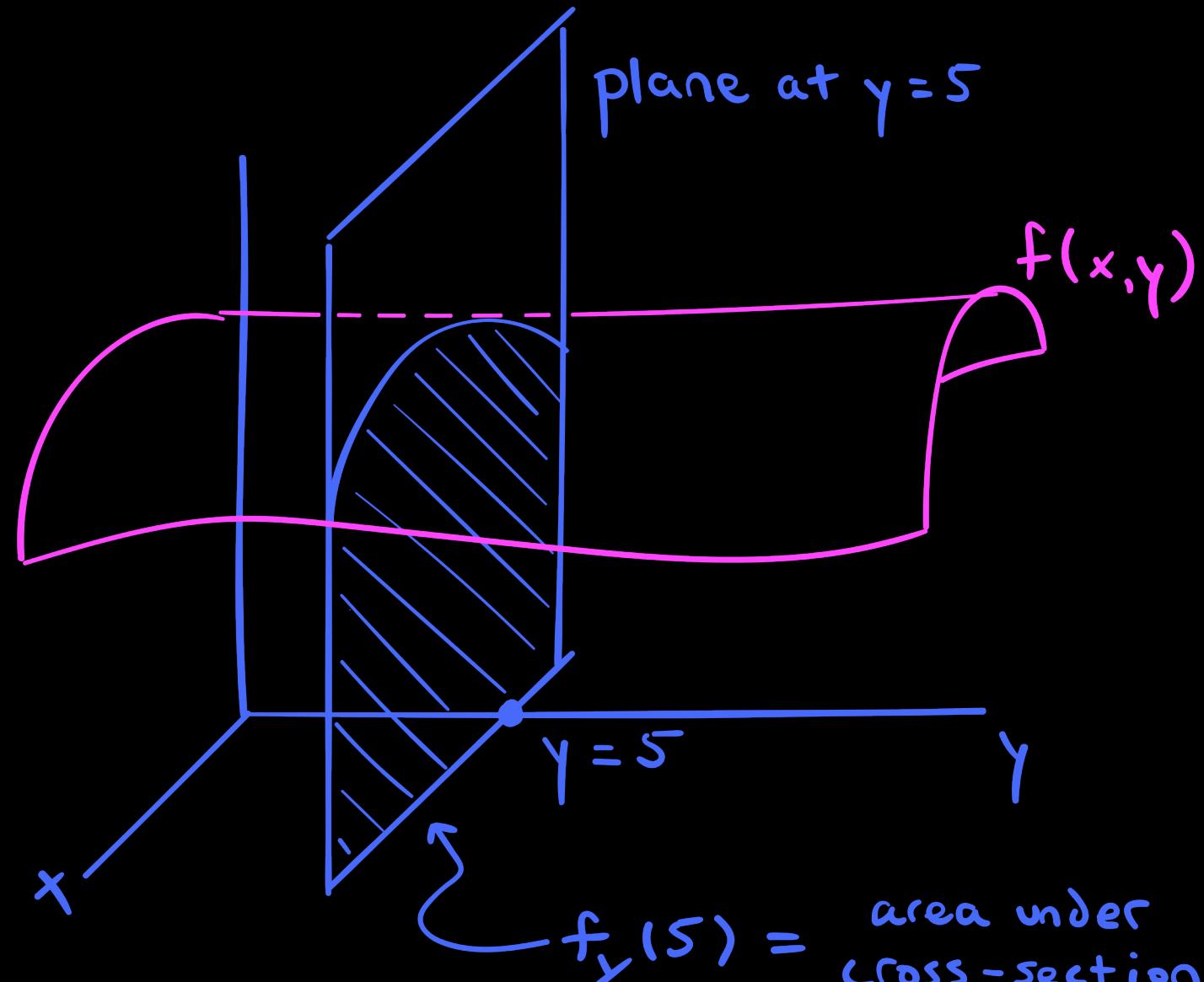
2. If (X, Y) is continuous with probability density function $f(x, y)$, then both X and Y are continuous with probability density functions given by

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y) \, dy \quad \text{and} \quad f_Y(y) = \int_{-\infty}^{\infty} f(x, y) \, dx. \quad (7.5)$$



Tip

1. To obtain the marginal mass $p_X(x)$ from the joint mass $p(x, y)$, we "sum out" the dependence of $p(x, y)$ on y . Likewise for obtaining $p_Y(y)$ from $p(x, y)$.
2. To obtain the marginal density $f_X(x)$ from the joint density $f(x, y)$, we "integrate out" the dependence of $f(x, y)$ on y . Likewise for obtaining $f_Y(y)$ from $f(x, y)$.





Problem Prompt

Do problems 6 and 7 on the worksheet.

7.5. Bivariate empirical distributions

Definition 7.6

Let $(X_1, Y_1), (X_2, Y_2), \dots, (X_m, Y_m)$ be a sequence of 2-dimensional random vectors, all defined on the same probability space.

- The random vectors are called a *bivariate random sample* if they are *independent* and *identically distributed* (IID).

Provided that the sequence is a bivariate random sample, an *observed bivariate random sample*, or a *bivariate dataset*, is a sequence of pairs of real numbers

$$(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)$$

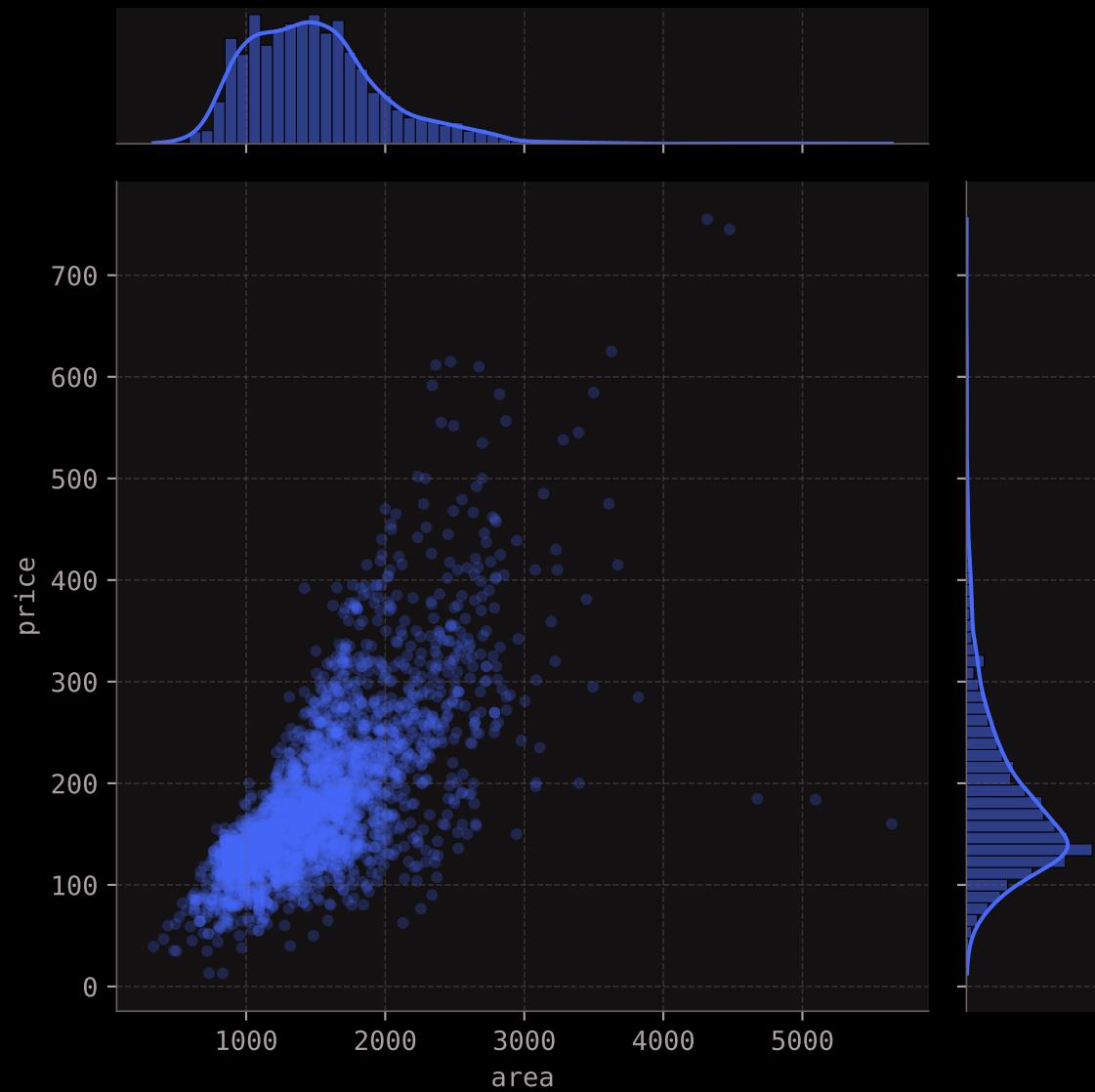
where (x_i, y_i) is an observation of (X_i, Y_i) .

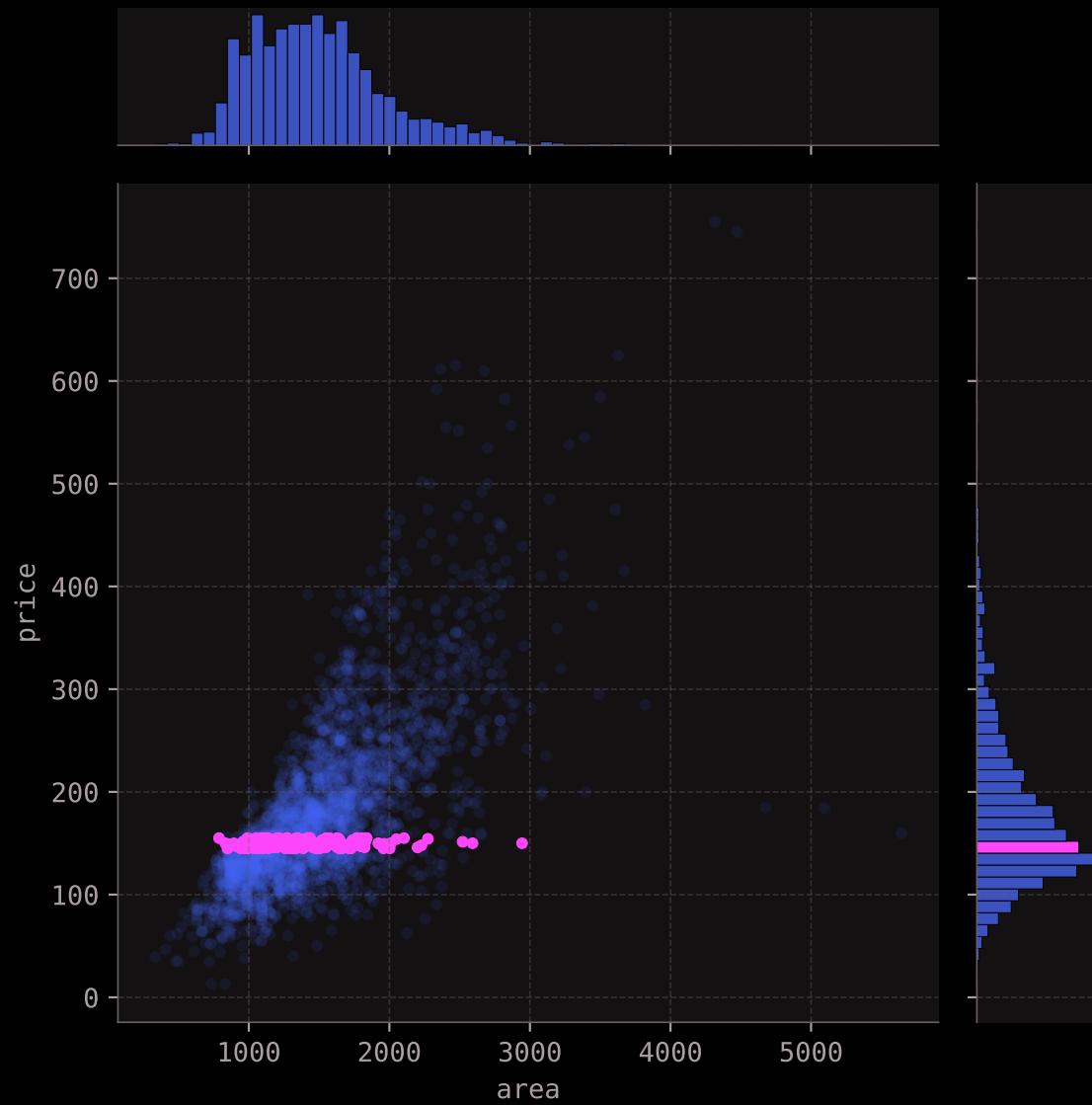


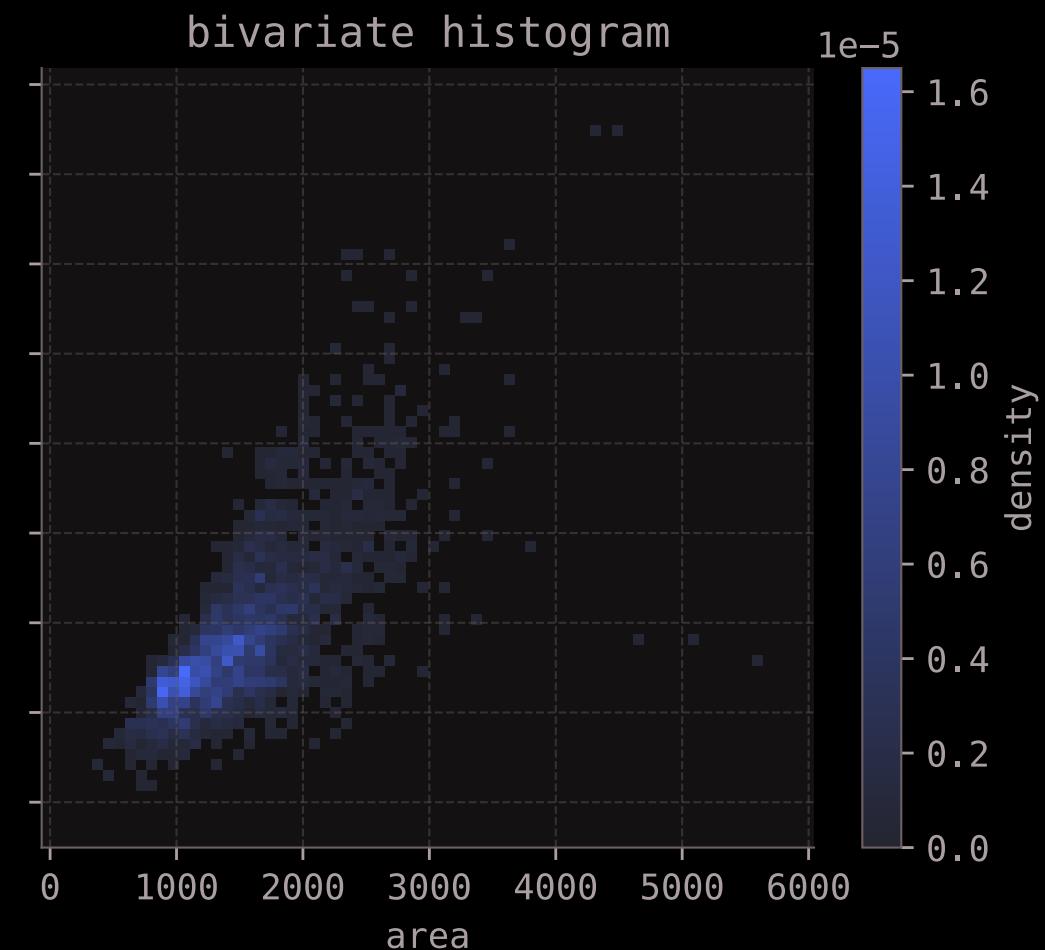
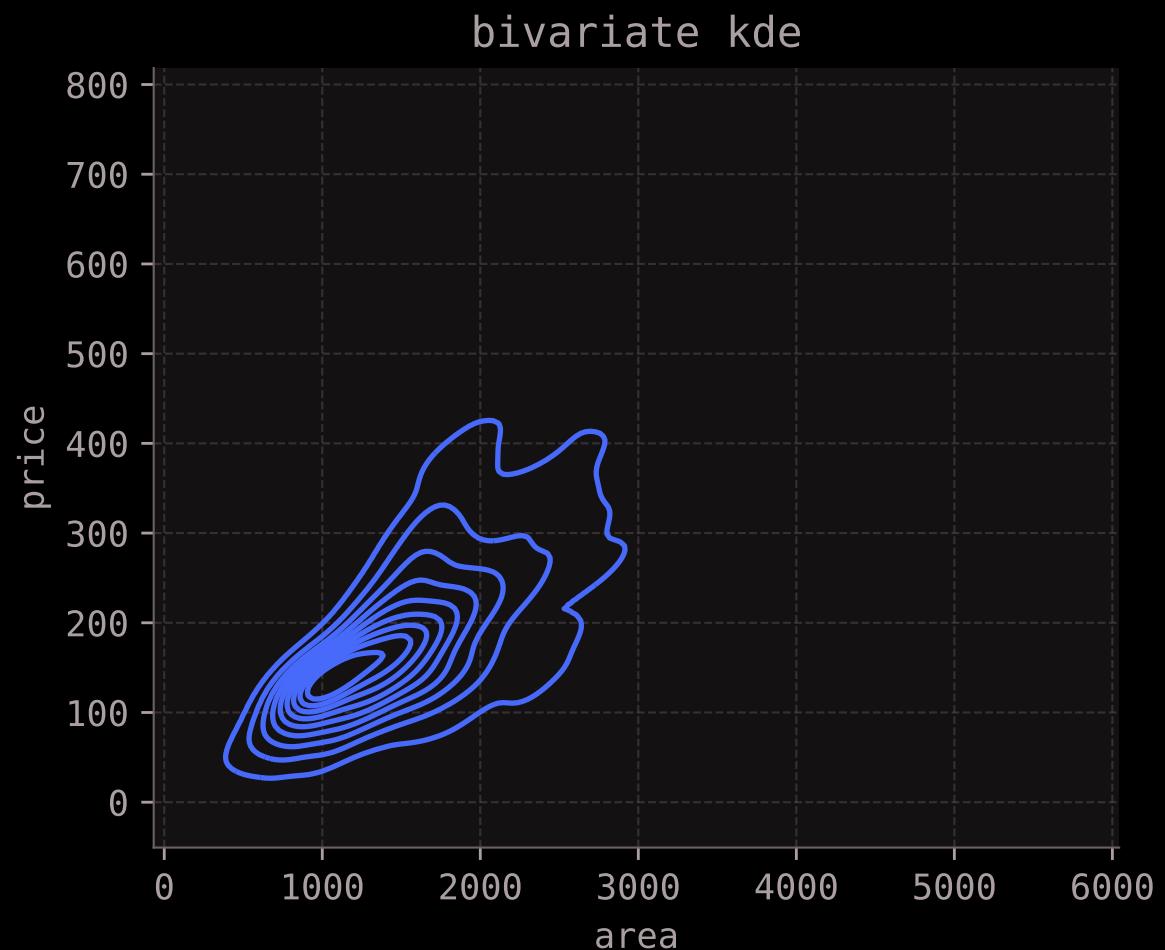
Definition 7.7

Let $(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)$ be an observed bivariate random sample, i.e., a bivariate dataset. The *empirical distribution* of the dataset is the discrete probability measure on \mathbb{R}^2 with joint probability mass function

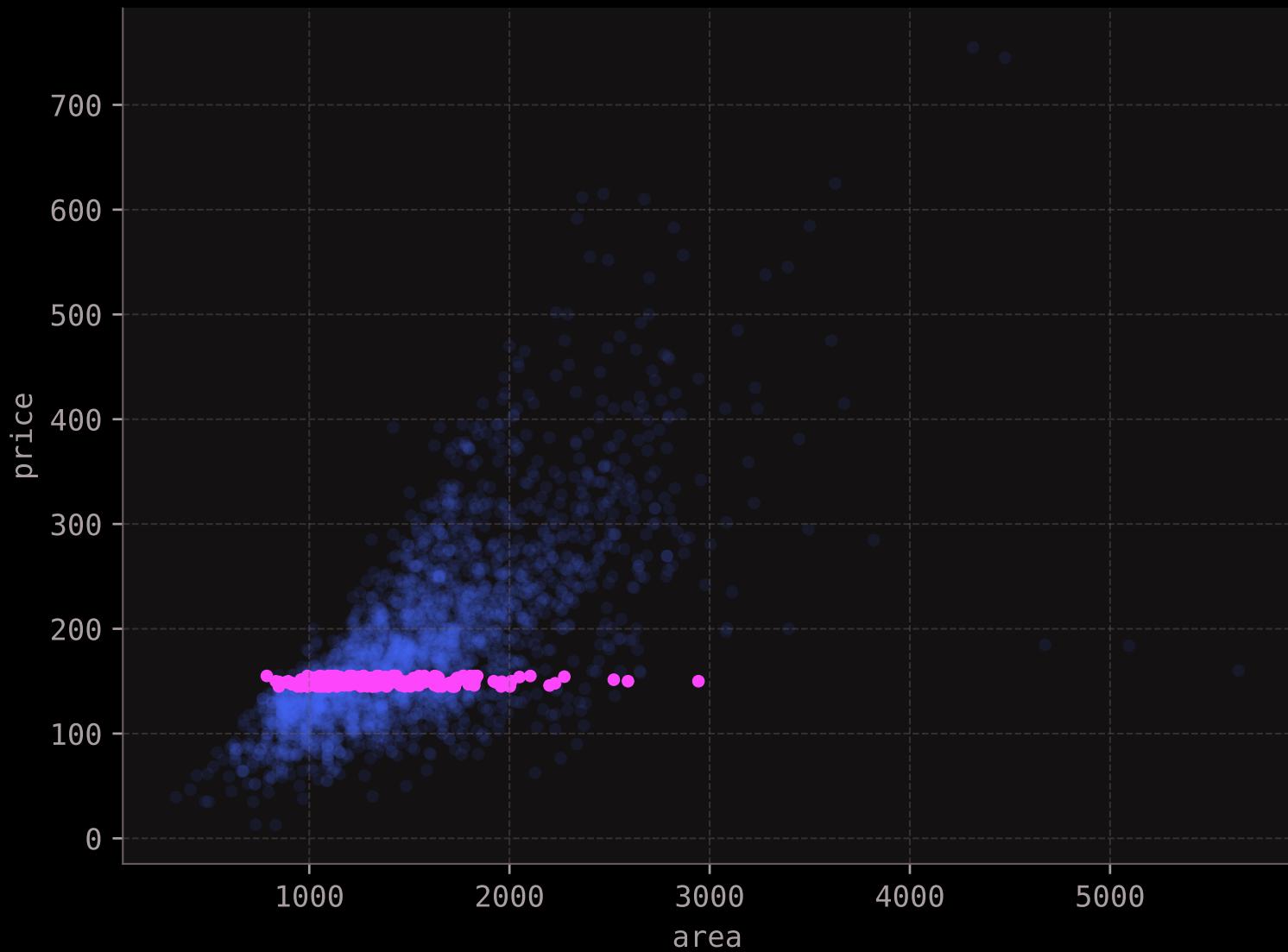
$$p(x, y) = \frac{\text{number of data points } (x_i, y_i) \text{ that match } (x, y)}{m}.$$

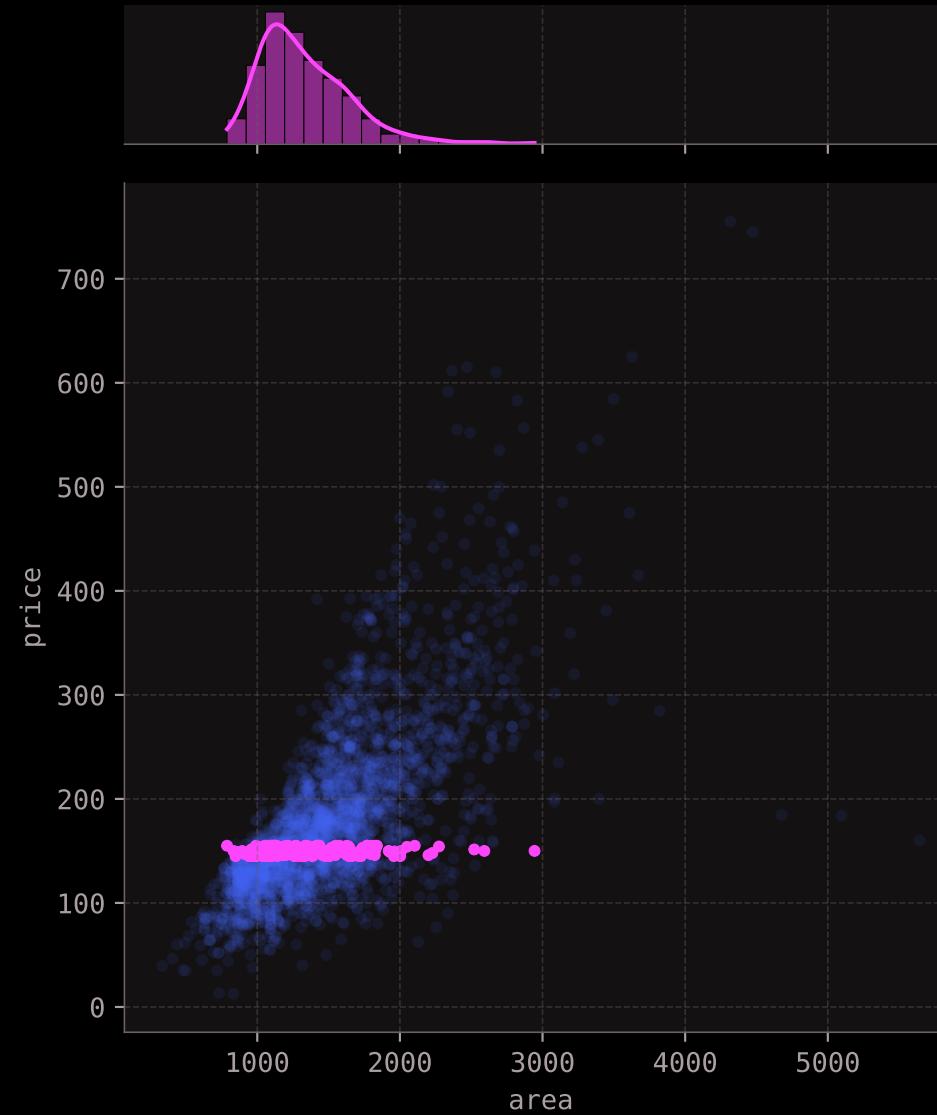






7.6. Conditional distributions





Definition 7.8

Let X and Y be random variables.

- Suppose (X, Y) is discrete, so that both X and Y are discrete as well. The *conditional probability mass function of X given Y* is the function

$$p_{X|Y}(x|y) = \frac{p_{XY}(x,y)}{p_Y(y)},$$

defined for all those y such that $p_Y(y) \neq 0$.

- Suppose (X, Y) is continuous, so that both X and Y are continuous as well. The *conditional probability density function of X given Y* is the function

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

defined for all those y such that $f_Y(y) \neq 0$.



Problem Prompt

Do problems 8 and 9 on the worksheet.

Theorem 7.4

Let X and Y be random variables.

- In the case that (X, Y) is discrete, for fixed y with $p_Y(y) \neq 0$, the function $p_{X|Y}(x|y)$ is a probability mass function in the variable x . In particular, we have

$$P(X \in A|Y = y) = \sum_{x \in A} p_{X|Y}(x|y), \quad (7.6)$$

for all events $A \subset \mathbb{R}$.

- In the case that (X, Y) is continuous, for fixed y with $f_Y(y) \neq 0$, the function $f_{X|Y}(x|y)$ is a probability density function in the variable x .



Problem Prompt

Do problems 10 and 11 on the worksheet.

7.7. The Law of Total Probability and Bayes' Theorem for random variables

🔔 Theorem 7.5 (The Law of Total Probability (for random variables))

Let X and Y be random variables.

- If X and Y are jointly discrete, then

$$p_X(x) = \sum_{y \in \mathbb{R}} p_{X|Y}(x|y)p_Y(y) \quad (7.7)$$

for each $x \in \mathbb{R}$.

- If X and Y are jointly continuous, then

$$f_X(x) = \int_{\mathbb{R}} f_{X|Y}(x|y)f_Y(y) \, dy. \quad (7.8)$$

for each $x \in \mathbb{R}$.

🔔 Theorem 7.6 (Bayes' Theorem (for random variables))

Let X and Y be random variables.

- If X and Y are jointly discrete, then

$$p_{X|Y}(x|y) = \frac{p_{Y|X}(y|x)p_X(x)}{p_Y(y)}. \quad (7.9)$$

- If X and Y are jointly continuous, then

$$f_{X|Y}(x|y) = \frac{f_{Y|X}(y|x)f_X(x)}{f_Y(y)}. \quad (7.10)$$



Problem Prompt

Do problem 12 on the worksheet.

7.8. Random vectors in arbitrary dimensions

Definition 7.9

Let S be a probability space and $n \geq 1$ an integer. An *n -dimensional random vector* is a function

$$\mathbf{X} : S \rightarrow \mathbb{R}^n.$$

Thus, we may write

$$\mathbf{X}(s) = (X_1(s), X_2(s), \dots, X_n(s))$$

for each sample point $s \in S$. When we do so, the functions X_1, X_2, \dots, X_n are ordinary random variables that are called the *components* of the random vector \mathbf{X} .

Definition 7.10

Let $(X_1, X_2, \dots, X_n) : S \rightarrow \mathbb{R}^n$ be an n -dimensional random vector on a probability space S with probability measure P . We define the *probability measure* of the random vector, denoted $P_{X_1 X_2 \dots X_n}$, via the formula

$$P_{X_1 X_2 \dots X_n}(C) = P(\{s \in S : (X_1(s), X_2(s), \dots, X_n(s)) \in C\}), \quad (7.11)$$

for all events $C \subset \mathbb{R}^n$. The probability measure $P_{X_1 X_2 \dots X_n}$ is also called the *joint distribution* of the component random variables X_1, X_2, \dots, X_n .



Problem Prompt

Do problems 13 and 14 on the worksheet.

7.9. Independence

Definition 7.11

Let $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_m$ be random vectors, all defined on the same probability space. Then these random vectors are said to be *independent* if

$$P(\mathbf{X}_1 \in C_1, \mathbf{X}_2 \in C_2, \dots, \mathbf{X}_m \in C_m) = P(\mathbf{X}_1 \in C_1)P(\mathbf{X}_2 \in C_2) \cdots P(\mathbf{X}_m \in C_m)$$

for all events C_1, C_2, \dots, C_m . If the vectors are not independent, they are called *dependent*.

🔔 Theorem 7.7 (Mass/Density Criteria for Independence)

Let X_1, X_2, \dots, X_m be random variables.

- Suppose that the random variables are jointly discrete. Then they are independent if and only if

$$p_{X_1 X_2 \cdots X_m}(x_1, x_2, \dots, x_m) = p_{X_1}(x_1)p_{X_2}(x_2) \cdots p_{X_m}(x_m)$$

for all $x_1, x_2, \dots, x_m \in \mathbb{R}$.

- Suppose that the random variables are jointly continuous. Then they are independent if and only if

$$f_{X_1 X_2 \cdots X_m}(x_1, x_2, \dots, x_m) = f_{X_1}(x_1)f_{X_2}(x_2) \cdots f_{X_m}(x_m) \quad (7.15)$$

for all $x_1, x_2, \dots, x_m \in \mathbb{R}$.



Theorem 7.8 (Invariance of Independence)

Let $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_m$ be independent n -dimensional random vectors, and let

$g : \mathbb{R}^n \rightarrow \mathbb{R}^k$ be a function. Then the k -dimensional random vectors

$g(\mathbf{X}_1), g(\mathbf{X}_2), \dots, g(\mathbf{X}_m)$ are independent.



Corollary 7.1 (Independence of Components)

Suppose $(X_1, Y_1), (X_2, Y_2), \dots, (X_m, Y_m)$ are independent 2-dimensional random vectors. Then the sequences X_1, X_2, \dots, X_m and Y_1, Y_2, \dots, Y_m of random variables are independent.

Theorem 7.9 (Conditional Criteria for Independence)

Let X and Y be two random variables.

- Suppose X and Y are jointly discrete. Then they are independent if and only if

$$p_{X|Y}(x|y) = p_X(x)$$

for all $x \in \mathbb{R}$ and all $y \in \mathbb{R}$ such that $p_Y(y) > 0$.

- Suppose X and Y are jointly continuous. Then they are independent if and only if

$$f_{X|Y}(x|y) = f_X(x)$$

for all $x \in \mathbb{R}$ and all $y \in \mathbb{R}$ such that $f_Y(y) > 0$.



Problem Prompt

Do problems 15-18 on the worksheet.