is 1. Then the function can be used to express probabilities regarding X. Specifically, for $a < b$, $P(a < X < b)$ can represented by the area under the curve between points a, b as shown in Figure 3.2. In fact, such an approach can be made precise by relating the function in the histogram to the "probability density function" of the random variable and calculating ares using calculus.

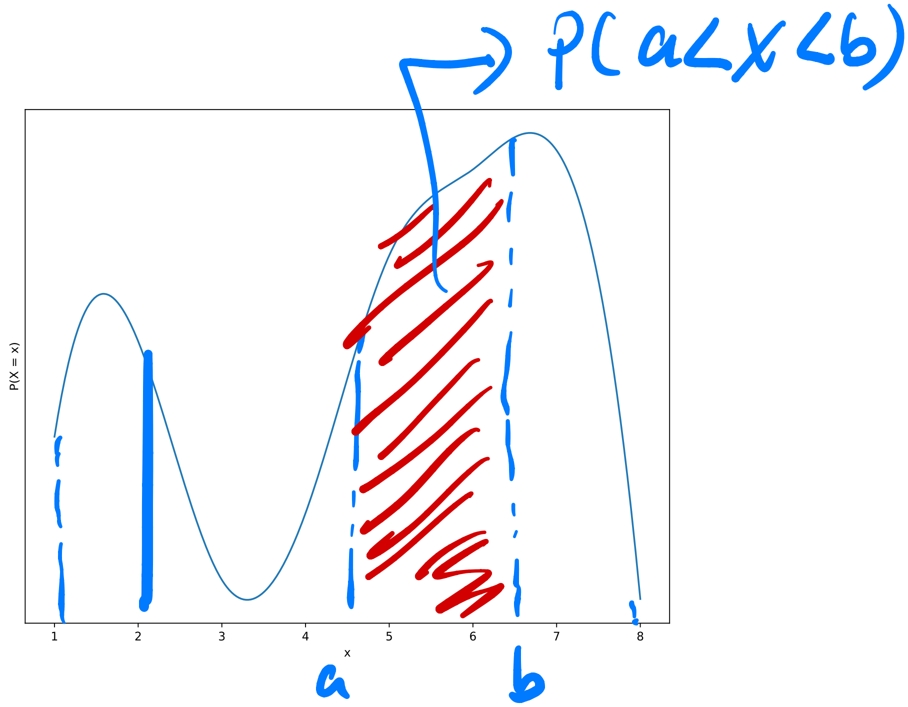

Figure 3.2: $P(a < X < b)$

One consequence of this approach to continuous random variables is that, for any choice of a, $P(X = a)$ is the area under the curve at a. Because, in mathematics, the area of a line is always 0, so $P(X = a) = 0$ for any a.

3.4 Summarizing The Distribution of a Random Variable

The distribution of a random variable can be complex and contains much information. Hence, it is often useful to summarize probability distributions by a few numbers, in the same way that we summarized a dataset in Chapter 2. In fact, any type of summary that can be applied to a dataset can be applied to a random variable by considering random variables $X_1, X_2, \ldots$ obtained from a long sequence of experiments. For instance, the mean of a random variable X can be viewed as the limiting value of the sample mean based on a long sequence of repetitions of the experiment. Therefore, the relationship between the mean of a random variable and a sample mean is the same as the relationship between a probability and a sample frequency. The mean of a random variable is sometimes called its *expected value.*

Consider the example in which X denotes the number of touch-downs passes thrown by Aaron Rodgers; using the distribution in Table 3.1, the mean or expected value of X is 2.15. Therefore, according to this result, in a large number of Bears-Packers games, we expect Rodgers to throw 2.15 touch-downs passes per game.

In the above example, X denotes the number of touch-downs passes thrown by Rodgers, $E(X) = 2.15$. This notation is particularly convenient when describing properties of means. For instance, if X and Y are random variables, then

$$E(X + Y) = E(X) + E(Y) \tag{3.1}$$

a is a constant

That is, the average value of a sum of two random variables is simply the sum of the average values. For example, let Y denote the number of rushing touch-downs scored by Rodgers in a given game and assume that $E(Y) = 0.23$; this corresponds to Rodgers' career average through the 2012 season. Then, Rodgers' average total number of touchdown is $2.15 + 0.23 = 2.38$. More generally, if a, b, and c are constants, then

$E(ax) = aE(x)$

$$E(aX + bY + c) = aE(X) + bE(Y) + c \tag{3.2}$$

For instance, in the Rodgers example, let $P = 6X + 6Y$ denote the total number of points results from his touch-down passes and rushing touch-downs. Then,

$P = 6x + 6y \Rightarrow E(P) = E(6x + 6y)$

$= E(6x) + E(6y)$

$$= 6\underbrace{E(X)}_{2.15} + 6\underbrace{E(Y)}_{0.23}$$

$$= 6(2.15) + 6(0.23)$$

$$= 14.28$$

The same approach used to define the mean of a random variable can be used to define the median, standard deviation, and variance of a random variable. These quantities can all be calculated from a random variable's probability distribution. For example, in the Rodgers' example, it can be shown that X has median 2, standard deviation 1.31, and variance 1.73.

3.5 Conditional Probability

Probabilities are interpreted as long-run relative frequencies in a large sequence of experiments. For example, in the Bears-Packers experiment in which the Bears play the Packers at home, if the event that the Bears win has probability 0.25, this means that in a hypothetical long sequence of games, the Bears will win about 25% of the time. An important part of this interpretation is that the 25% applies to all of the experiments in the sequence. However, in some cases, we might only be interested in those experiments satisfying some further conditions.

Continuing the example, let B be the event that the Bears win so that $P(B) = 0.25$. Now, consider another event, the one in which jay Cutler throws 4 interceptions; denote this event by C. Suppose we are interested in the Bears' probability of winning in those games in which Cutler throws 4 interceptions. We can describe this probability as "the probability that the Bears win *given that* Cutler throws 4 interceptions;" symbolically, we write

$$\boxed{P(B|C)}$$

where the vertical line is read "given that." Probabilities such as $P(B|C)$ are called *conditional probability* because they include additional conditions. Note that, in $P(B|C)$, we are interested in the probability of B; C simply describe the conditions under which the probability is to be calculated. Thus, if, for example, $P(B|C) = 0.05$, then in a long sequence of games *in which Cutler throws 4 interceptions*, the Bears win about 5% of the time. We might write this conditional probability more informally as

$$\boxed{P(\text{Bears beat the Packers} \mid \text{Cutler throws 4 interceptions}) = 0.05}$$

Conditional probabilities are useful because they allow us to incorporate additional assumptions, or additional information, into the probability calculation. Conditional probabilities can be determined from standard, unconditional probabilities. To see how this can be done, consider how we would calculate the conditional in the example if we had access to the results from the long sequence of games.

We are interested in those games in which Cutler throws 4 interceptions, so if we are reading a sequence of game summaries, one per page; we place those in which Cutler threw 4 interceptions in a separate pile, a C pile. We now go through that pile and put those that the Bears won in a second pile. The probability $P(B|C)$ is the ratio of the number of games in the second pile to the number of games in the C pile. Note that the second pile consists of those games that Cutler throws 4 interceptions *and* the Bears win; denote this event by "B and C." Thus, in probability terms,

$$P(B|C) = \frac{P(B \text{ and } C)}{P(C)}$$

Probability of the interception

In general, consider an experiment and let A, B be events. Let A *and* B be the event in which both A and B occur. Then,

$$P(A|B) = \frac{P(A \textit{ and } B)}{P(B)} \tag{3.3}$$

Note that this relationship can be written as

$$P(A \textit{ and } B) = P(A|B)P(B) \tag{3.4}$$

so that it gives a formula for finding the probability that two events both occur. That is, the probability that both A, B occur is the probability that B occurs times the probability that A occurs given that B occurs. Note that, because A *and* B, and B *and* A mean the same thing, we also have

$$P(A \textit{ and } B) = P(B|A)P(A) \tag{3.5}$$

giving two options for determining $P(A \textit{ and } B)$. The conditional probability $P(A|B)$ can be viewed as an "updated" version of the probability, updated to take into account that B occurs. In many cases, this additional information is important in assessing the probability of interest. For instance, in the Bears-Packers example, knowing that Cutler throws 4 interceptions will change the probability that the Bears win. However, in other cases, knowing that B occurs will not change the probability that A occurs. For instance, if K is the event that the Bears kick off to start the game, then it may be reasonable to assume that

$$P(B|K) = P(B)$$

that is, the fact that the Bears kick off to start the game does not change their probability of winning the game. Events A, B for which $P(A|B) = P(A)$ are said to *independent*. For independent events, knowledge about one of them does not change our probability of the other. By rearranging the formula for $P(A|B)$, it can be shown that A, B are independent if and only if

$$P(A \textit{ and } B) = P(A)P(B) \tag{3.6}$$

That is, for independent events A and B, the probability that both occur is simply the product of the individual probabilities. Consideration of conditional probabilities shows that it is important to be aware of the conditions under which a probability is calculated and to pay close attention

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \quad \cdots \text{①}$$

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} \quad \cdots \text{②}$$

$$P(A \cap B) = P(A \mid B)\,P(B) = P(B \mid A)\,P(A)$$

1

x	$P(X=x)$
0	0.10
1	0.25
2	0.25
3	0.25
4	0.10
5	0.05

50%

50%

median = 2

Random Variable

$$E(X) = \sum_x x \cdot P(X=x)$$

outcome

$$= 0 \cdot \underbrace{P(X=0)}_{0.10} + 1\,\underbrace{P(X=1)}_{0.25} + 2\,\underbrace{P(X=2)}_{0.25}$$

$$+ 3\,\underbrace{P(X=3)}_{0.25} + 4\,\underbrace{P(X=4)}_{0.10} + 5\,\underbrace{P(X=5)}_{0.05}$$

$$= 0 + 0.25 + 0.5 + 0.75 + 0.4 + 0.25$$

$$= 2.15$$

Since $E(X) > \text{Median}(X)$, the distribution of X is a little right-skewed.

(2) B: Barcelona wins a home game

M: Messi scores 2 or more goals in a home game

$P(B) = 0.65$ $P(B \cap M) = 0.35$

$P(M) = 0.45$

$$P(B|M) = \frac{P(B \cap M)}{P(M)} = \frac{0.35}{0.45} = \frac{7}{9}$$

$$P(M|B) = \frac{P(B \cap M)}{P(B)} = \frac{0.35}{0.65} = \frac{7}{13}$$

$$P(A \cap B) = P(B \cap A)$$