Example 4.1.4 (Uniform Distribution). Let $X_1, \ldots, X_n$ be iid with the uniform $(0, \theta)$ density; i.e., $f(x) = 1/\theta$ for $0 < x < \theta$, 0 elsewhere. Because θ is in the support, differentiation is not helpful here. The likelihood function can be written as

$$L(\theta) = \theta^{-n} I(\max\{x_i\}, \theta), \quad \text{for all } \theta > 0,$$

where $I(a, b)$ is 1 or 0 if $a \leq b$ or $a > b$, respectively. The function $L(\theta)$ is a decreasing function of θ for all $\theta \geq \max\{x_i\}$ and is 0 otherwise [sketch the graph of $L(\theta)$]. So the maximum occurs at the smallest value that θ can assume; i.e., the mle is $\widehat{\theta} = \max\{X_i\}$. ■

4.1.1 Histogram Estimates of pmfs and pdfs

Let $X_1, \ldots, X_n$ be a random sample on a random variable X with cdf $F(x)$. In this section, we briefly discuss a histogram of the sample, which is an estimate of the pmf, $p(x)$, or the pdf, $f(x)$, of X depending on whether X is discrete or continuous. Other than X being a discrete or continuous random variable, we make no assumptions on the form of the distribution of X. In particular, we do not assume a parametric form of the distribution as we did for the above discussion on maximum likelihood estimates; hence, the histogram that we present is often called a **nonparametric** estimator. See Chapter 10 for a general discussion of nonparametric inference. We discuss the discrete situation first.

The Distribution of X Is Discrete

Assume that X is a discrete random variable with pmf $p(x)$. Suppose first that the space of X is finite, say, $\mathcal{D} = \{a_1, \ldots, a_m\}$. An intuitive estimator of $p(a_j)$ is the relative frequency of sample observations, which are equal to a_j. For $j = 1, 2, \ldots, m$, define the statistics

$$I_j(X_i) = \begin{cases} 1 & X_i = a_j \\ 0 & X_i \neq a_j. \end{cases}$$

Then the intuitive estimate of $p(a_j)$ can be expressed by the average

$$\widehat{p}(a_j) = \frac{1}{n} \sum_{i=1}^{n} I_j(X_i). \tag{4.1.9}$$

Thus the estimates $\{\widehat{p}(a_1), \ldots, \widehat{p}(a_m)\}$ constitute the nonparametric estimate of the pmf $p(x)$. Note that $I_j(X_i)$ has a Bernoulli distribution with probability of success $p(a_j)$. As Exercise 4.1.6 shows, our estimator of the pmf is unbiased.

Suppose next that the space of X is infinite, say, $\mathcal{D} = \{a_1, a_2, \ldots\}$. In practice, we select a value, say, a_m, and make the groupings

$$\{a_1\}, \{a_2\}, \ldots, \{a_m\}, \tilde{a}_{m+1} = \{a_{m+1}, a_{m+2}, \ldots\}. \tag{4.1.10}$$

Let $\widehat{p}(\tilde{a}_{m+1})$ be the proportion of sample items that are greater than or equal to a_{m+1}. Then the estimates $\{\widehat{p}(a_1), \ldots, \widehat{p}(a_m), \widehat{p}(\tilde{a}_{m+1})\}$ form our estimate of