divided into m subintervals or **bins** of width Δx_i, $i = 1, \ldots, m$, where Δx_i is usually but not necessarily the same for each bin. The number of occurrences n_i of x in subinterval i, i.e. the number of entries in the bin, is given on the vertical axis. The area under the histogram is equal to the total number of entries n multiplied by Δx (or for unequal bin widths, $area = \sum_{i=1}^{m} n_i \cdot \Delta x_i$). Thus the histogram can be normalized to unit area by dividing each n_i by the corresponding bin width Δx_i and by the total number of entries in the histogram n. The p.d.f. $f(x)$ corresponds to a histogram of x normalized to unit area in the limit of zero bin width and an infinitely large total number of entries, as illustrated in Fig. 1.2(d).

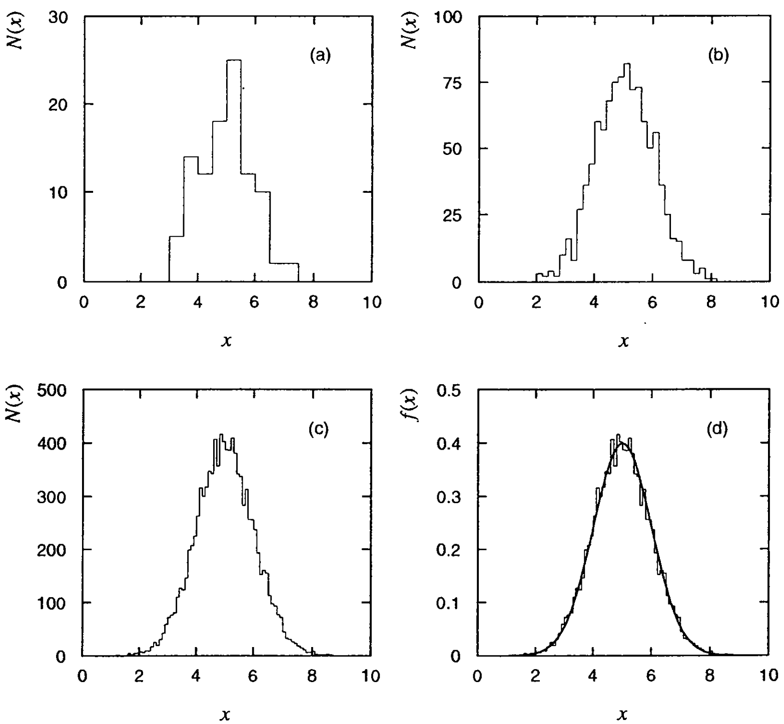

Fig. 1.2 Histograms of various numbers of observations of a random variable x based on the same p.d.f. (a) $n = 100$ observations and a bin width of $\Delta x = 0.5$. (b) $n = 1000$ observations, $\Delta x = 0.2$. (c) $n = 10000$ observations, $\Delta x = 0.1$. (d) The same histogram as in (c), but normalized to unit area. Also shown as a smooth curve is the p.d.f. according to which the observations are distributed. For (a–c), the vertical axis $N(x)$ gives the number of entries in a bin containing x. For (d), the vertical axis is $f(x) = N(x)/(n\Delta x)$.

One can consider cases where the variable x only takes on discrete values x_i, for $i = 1, \ldots, N$, where N can be infinite. The corresponding probabilities can be expressed as