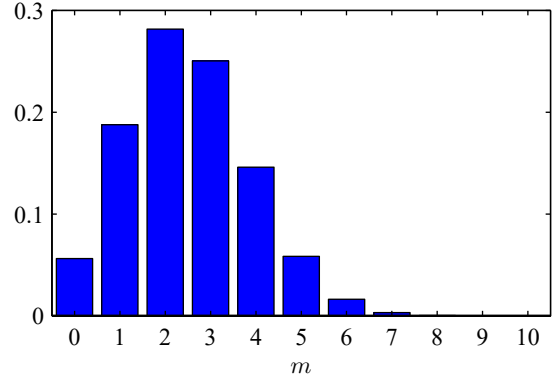


Figure 3.1 Histogram plot of the binomial distribution (3.9) as a function of m for $N = 10$ and $\mu = 0.25$.



where

$$\binom{N}{m} \equiv \frac{N!}{(N-m)!m!} \quad (3.10)$$

is the number of ways of choosing m objects out of a total of N identical objects without replacement. Figure 3.1 shows a plot of the binomial distribution for $N = 10$ and $\mu = 0.25$.

Exercise 3.3

The mean and variance of the binomial distribution can be found by using the results that, for independent events, the mean of the sum is the sum of the means and the variance of the sum is the sum of the variances. Because $m = x_1 + \dots + x_N$ and because for each observation the mean and variance are given by (3.3) and (3.4), respectively, we have

Exercise 2.10

$$\mathbb{E}[m] \equiv \sum_{m=0}^N m \text{Bin}(m|N, \mu) = N\mu \quad (3.11)$$

$$\text{var}[m] \equiv \sum_{m=0}^N (m - \mathbb{E}[m])^2 \text{Bin}(m|N, \mu) = N\mu(1 - \mu). \quad (3.12)$$

Exercise 3.4

These results can also be proved directly by using calculus.

3.1.3 Multinomial distribution

Binary variables can be used to describe quantities that can take one of two possible values. Often, however, we encounter discrete variables that can take on one of K possible mutually exclusive states. Although there are various alternative ways to express such variables, we will see shortly that a particularly convenient representation is the 1-of- K scheme, sometimes called ‘one-hot encoding’, in which the variable is represented by a K -dimensional vector \mathbf{x} in which one of the elements x_k equals 1 and all remaining elements equal 0. So, for instance, if we have a variable that can take $K = 6$ states and a particular observation of the variable happens to