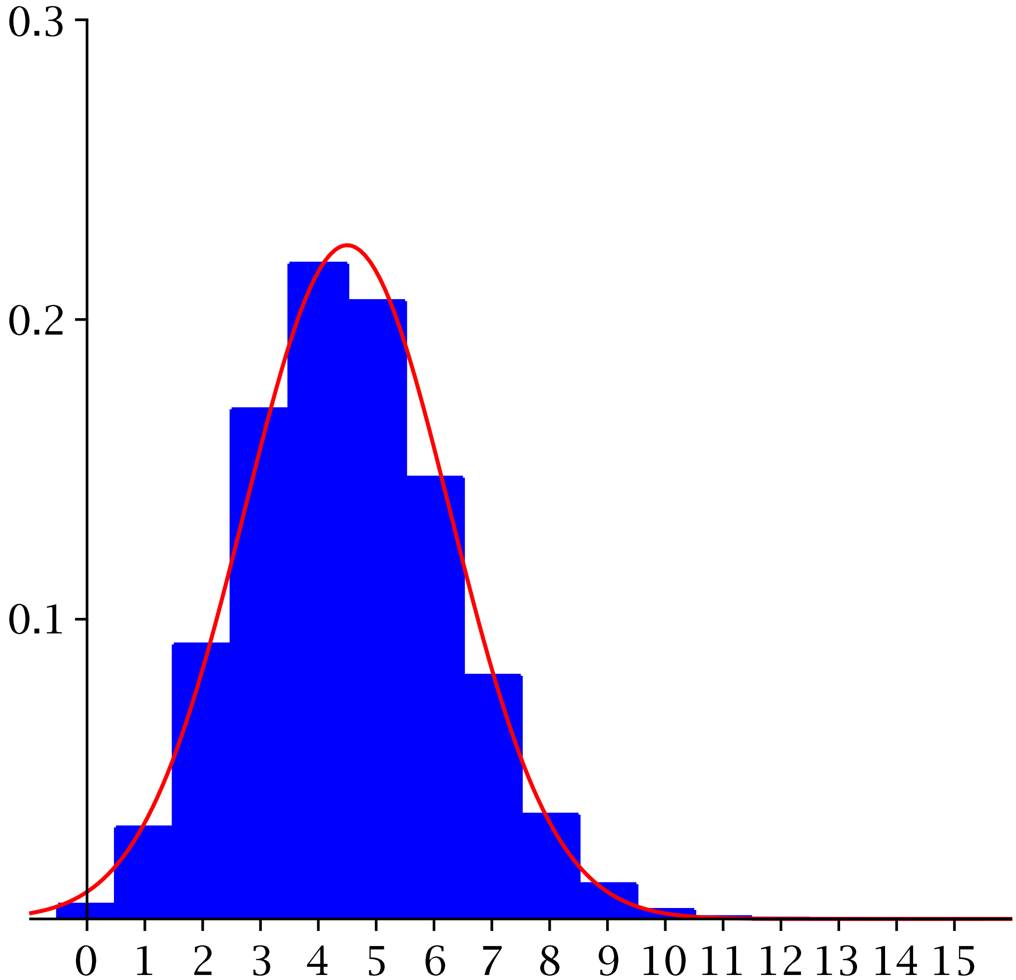
MONT 101N -- Analyzing Environmental Data
The Normal Approximation to a Binomial Probability

If $Y\_1, ..., Y\_N$ are independent Bernoulli trials with success probability p, then N times the sample mean of the Y_i is a Binomial random variable X. The Central Limit Theorem implies that for N large enough, the sample mean will be approximately normal with mean $\mu = p$ and standard deviation $\sigma = \sqrt{\frac{p\,q}{N}}$. This implies that the binomial random variable $N \cdot sample\,mean$ will be approximately normal U with $\mu = N\,p = 4.5$ and standard deviation $\sigma = \sqrt{N\,p\,q} \doteq 1.7748$. Here is a picture
illustrating what happens for $N = 15$, $p = .3$. The blue bars show the probability histogram for the binomial random variable X. The red curve is the approximating normal p.d.f. for U.



Then, for example, we can use the standardization procedure and the standard normal
table to approximate binomial probabilities. The idea is that for each of the integer values k = 0,1,2,3,..., 15, the binomial probability will be approximated by the area

under the
red curve for $k - .5 \le x \le k + .5$.

For instance, to approximate the probability that the binomial X equals 7, we would find $P(6.5 \le U \le 7.5)$. Standardizing, this is the same as the standard normal probability

$$P\left(\frac{(6.5 - 4.5)}{1.7748} \le Z \le \frac{(7.5 - 4.5)}{1.7748} \right) \doteq P(1.13 \le Z \le 1.69) = .4545 - .3708 = .0837$$

The exact binomial probability is somewhat awkward to compute by hand with *N, k* this
large, but it's certainly possible:

$$P(X = 7) = \binom{15}{7} \cdot (.3)^7 \cdot (.7)^8 = 6435 \cdot (.3)^7 (.7)^8 \doteq .0811$$

This agrees with our binomial approximation to two decimal places -- not extremely close
but good enough for many purposes(!)

▶ **Calculations**