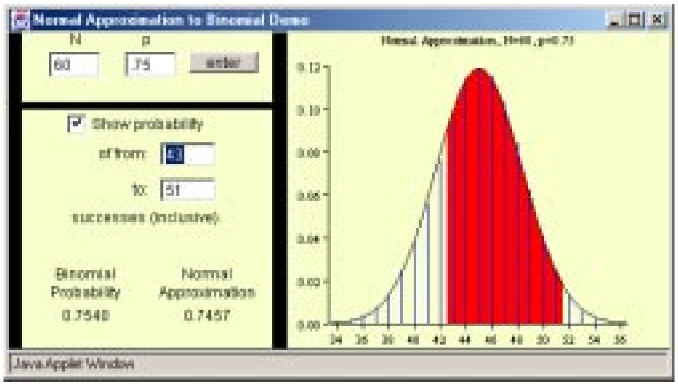
Math 180A, Fall 2005, Prof. Tesler – October 7, 2005
Approximating the binomial distribution by the normal curve

You can try this at home: `http://www.ruf.rice.edu/~lane/stat_sim/normal_approx/index.html`



k	$P(X=k)=\binom{60}{k}(.75)^k(.25)^{60-k}$
43	0.09562
44	0.11083
45	0.11822
46	0.11565
47	0.10335
48	0.08397
49	0.06169
50	0.04071
51	0.02395
total	**0.75404**

(a) **The binomial distribution** for $n = 60$, $p = 3/4$, is shown as vertical lines.

The mean is $\mu = np = 60(3/4) = 45$. In this case it's an integer, but it doesn't have to be.

The standard deviation is $\sigma = \sqrt{np(1-p)} = \sqrt{60(3/4)(1/4)} = \sqrt{11.25} \approx 3.354101966$.

The mode is $\lfloor np + p \rfloor = \lfloor 60(3/4) + (3/4) \rfloor = \lfloor 45\frac{3}{4} \rfloor = 45$. It only agrees with the mean because the mean is an integer. Notice that the mode has the largest probability density: $P(X = 45) = \binom{60}{45}(.75)^{45}(.25)^{15} \approx 0.1182280046$. Although this is the largest density, it is not even close to 1. Also, note that this applet only shows the x-axis betwen 34 and 56, even though X takes on integer values between 0 and 60. This is because the probabilities outside the range shown are very small; they are still positive, but on this scale you would not be able to tell that they are not zero.

We can compute $P(43 \le X \le 51)$ for the binomial distribution by adding up $P(X = 43) + P(X = 44) + \cdots + P(X = 51) \approx 0.75404$; see the table at top right.

(b) **The normal distribution** for those same values of μ and σ ($\mu = 45$, $\sigma = \sqrt{11.25}$) is shown as a smooth curve with a bell shape. Notice that at integer values of X, the binomial distribution and the normal distribution are nearly the same.

Let's approximate the probability $P(43 \le X \le 51)$ we had above by the corresponding one for the normal distribution.

(i) Apply the "continuity correction": the binomial distribution is discrete (inputs are integers) but the normal distribution is continuous. Use $P(43 \le X \le 51) = P(42.5 \le X \le 51.5)$ for the binomial distribution. There were 9 terms added together in (a), and here, we get $51.5 - 42.5 = 9$. The integral under the normal curve from 42.5 to 51.5 could be approximated as summing 9 rectangles of width 1, centered at $43, \ldots, 51$.

(ii) Convert to z-scores: $z = \dfrac{x-\mu}{\sigma} = \dfrac{x-np}{\sqrt{np(1-p)}} = \dfrac{x-45}{\sqrt{11.25}}$

$$P(42.5 \le X \le 51.5) = P\left(\frac{42.5-np}{\sqrt{np(1-p)}} \le \frac{X-np}{\sqrt{np(1-p)}} \le \frac{51.5-np}{\sqrt{np(1-p)}}\right)$$

$$= P\left(\frac{42.5-45}{\sqrt{11.25}} \le Z \le \frac{51.5-45}{\sqrt{11.25}}\right) = P(-.7453559926 \le Z \le 1.937925581)$$

(iii) Approximate this by the standard normal distribution:

$$P(-.7453559926 \le Z \le 1.937925581) \approx \Phi(1.937925581) - \Phi(-.7453559926)$$

(iv) Look in the table in the book on page 531; it only has z's to two decimal places, so round them: $\Phi(1.94) = 0.9738$ and $\Phi(-0.75) = 1 - \Phi(0.75) = 1 - 0.7734 = 0.2266$. Then $\Phi(1.937925581) - \Phi(-.7453559926) \approx$ $\boxed{0.7472}$.

If you use a calculator or software that can compute $\Phi(z)$ more accurately than the table in the book, you will get

$$\Phi(1.937925581) - \Phi(-.7453559926) = 0.9736838487 - 0.2280282702 = \boxed{\mathbf{0.7456555785}}$$