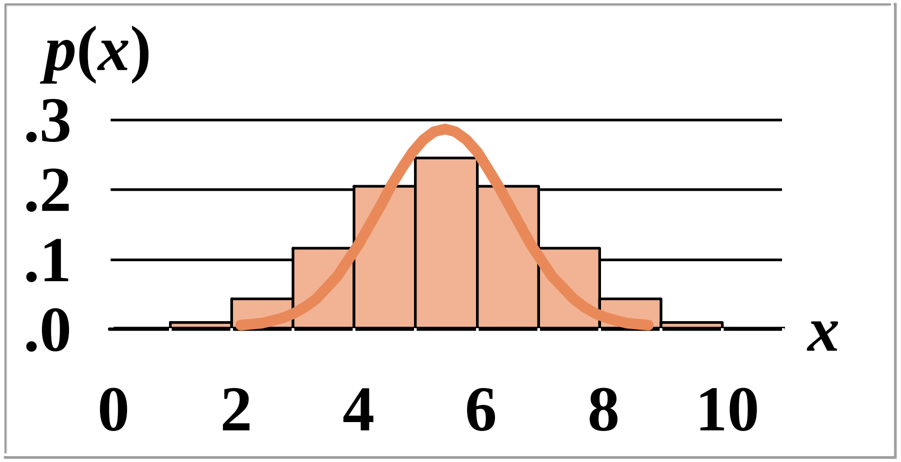
Approximating a Binomial Distribution with a Normal Distribution

Normal Approximation of Binomial Distribution

1. Useful because not all binomial tables exist
2. Requires large sample size
3. Gives approximate probability only
4. Need correction for continuity

$n = 10 \quad p = 0.50$



Normal Approximation to Binomial Distributions

- The normal distribution is used to approximate the binomial distribution when it would be impractical to use the binomial distribution to find probability.
- Theorem: Normal approximation to a binomial distribution

If $np > 5$ and $n(1-p) > 5$, then the binomial random variable X is approximately normally distributed with mean (μ) and standard deviation (σ) given as follows:

$\mu = np$, $\sigma = \sqrt{npq}$

Where n is the number of trial, p is the probability of success on each trial and q is the probability of failure on each trial= $1-p$

Normal Approximation to Binomial Distributions

Example

For each of the following cases, decide whether you can use the normal distribution to approximate X. if you can find the mean and standard deviation:

n= 65, p = 0.51, q = 0.49

np = (65)(0.51) = 33.15 > 5

nq = (65)(0.49) = 31.85 > 5

So we can use the normal approximation

Mean: $\mu = np = 33.15$, $\sigma = \sqrt{npq} = 4.03$

- N = 15, p = 0.15, q = 0.85

np = (15)(0.15) = 2.25 < 5, nq = (15)(0.85) = 12.75 > 5

We cannot use the normal approximation , because np < 5.

Continuity Correction

The binomial distribution is discrete and to calculate exact binomial probabilities, the binomial formula is used for each value of x. in order for a continuous distribution (like the normal) to be used to approximate a discrete one (like the binomial), a continuity correction should be used. There are two major reasons to employ such a corrections as follows:

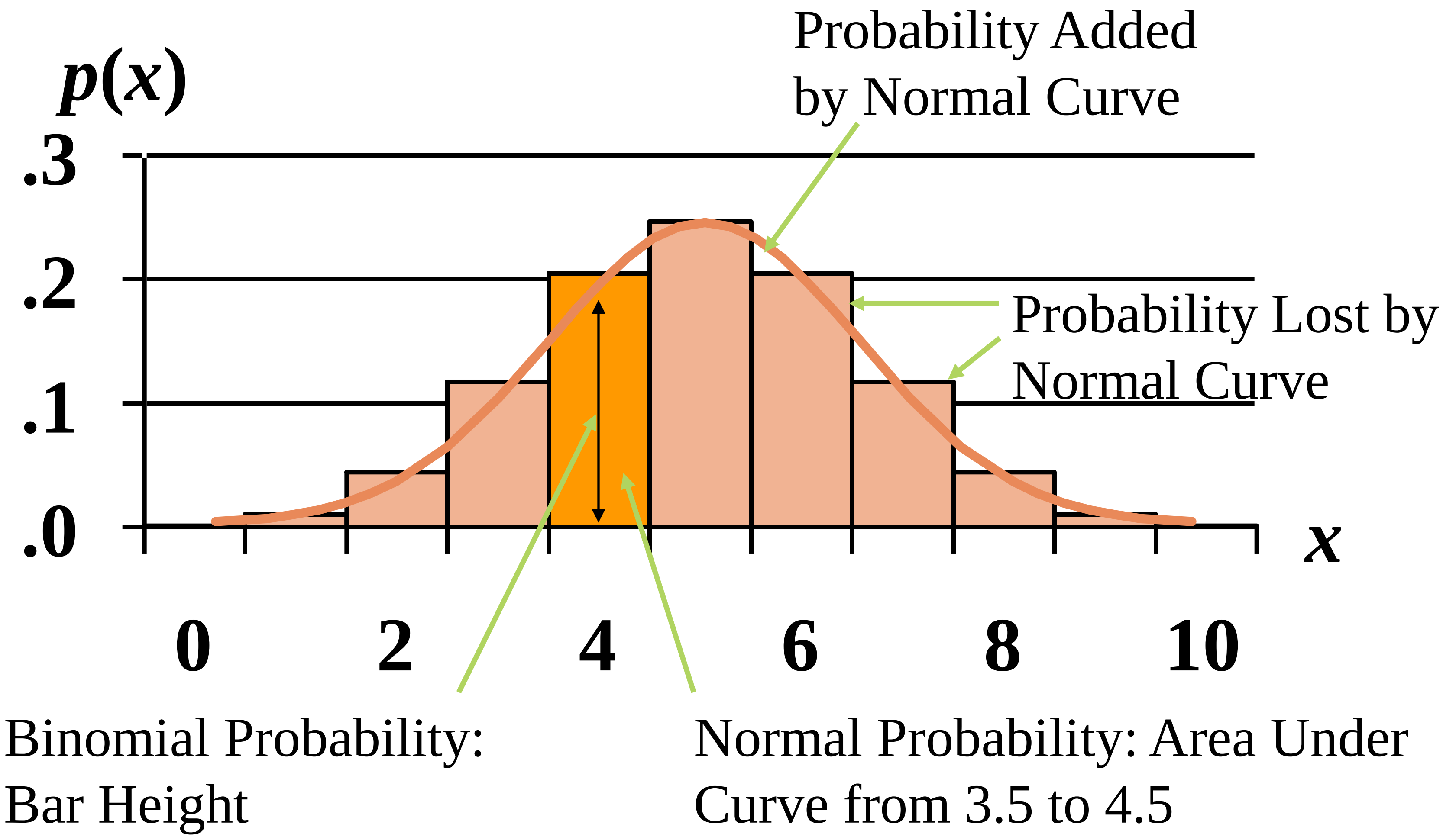
- First, recall that a discrete random variable (like the binomial) can only take on only specified values, whereas a continuous random variable used to approximate it can take on any values whatsoever.

within an interval around those specified values. Hence, when using the normal distribution to approximate the binomial distribution, more accurate approximations are likely to be obtained if a continuity correction is used.

Second, recall that with a continuous distribution (such as the normal), the probability of obtaining a particular value of a random variable is zero. On the other hand, when the normal approximation is used to approximate a discrete distribution, a continuity correction can be employed so that we can approximate the probability of a specific value of the discrete distribution (such as the binomial).

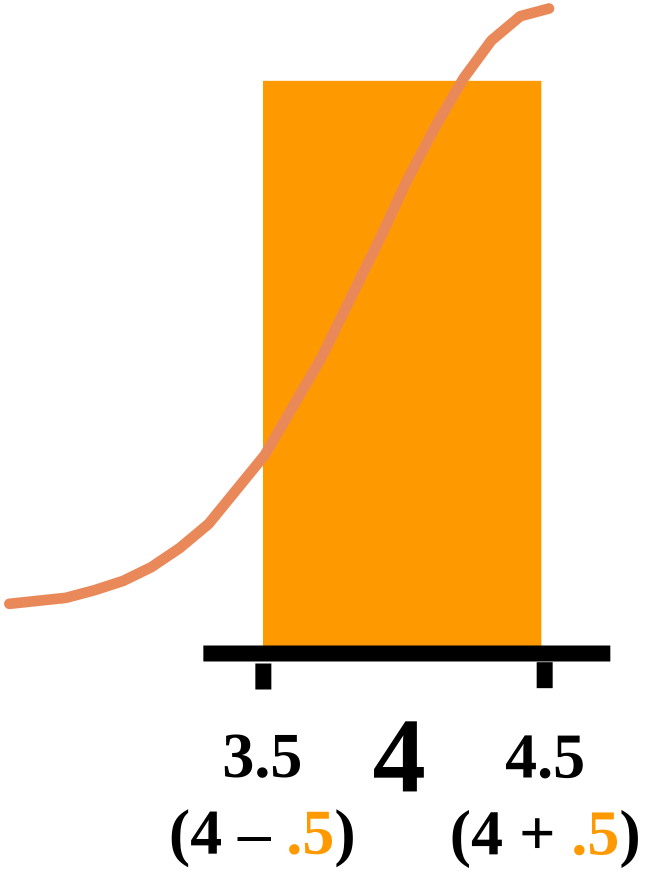
Therefore, when we use a normal distribution which is continuous to approximate a binomial probability which is discrete, we need to move 0.5 unit to the left or to the right of the x-value to include all possible x-values in the interval by using what is called continuity correction. To use a continuity correction to convert the binomial distribution x-values to a normal distribution x-values, we use the following rules:.

Why Probability Is Approximate



Correction for Continuity

1. A 1/2 unit adjustment to discrete variable
2. Used when approximating a discrete distribution with a continuous distribution
3. Improves accuracy



Using a Normal Distribution to Approximate Binomial Probabilities

1. Determine *n* and *p* for the binomial distribution, then calculate the interval:

$$\mu \pm 3\sigma = np \pm 3\sqrt{np(1-p)}$$

If interval lies in the range 0 to *n*, the normal distribution will provide a reasonable approximation to the probabilities of most binomial events.

Using a Normal Distribution to Approximate Binomial Probabilities

2. Express the binomial probability to be approximated by the form

$$P(x \le a) \text{ or } P(x \le b) - P(x \le a)$$

For example,

$$P(x < 3) = P(x \le 2)$$

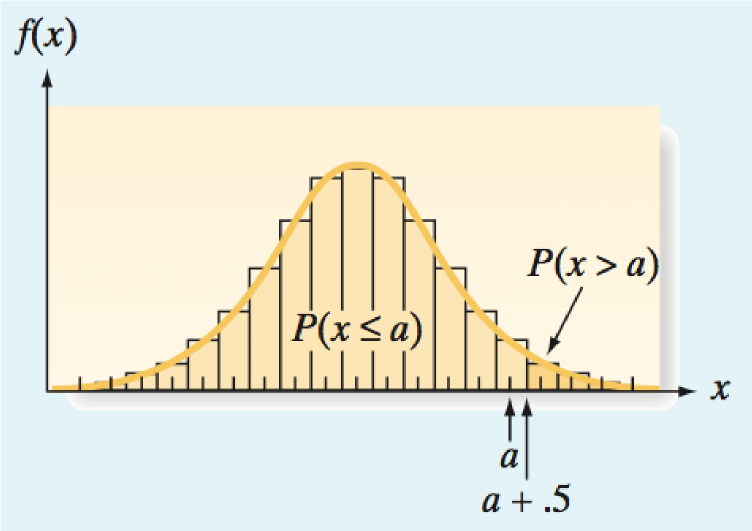
$$P(x \ge 5) = 1 - P(x \le 4)$$

$$P(7 \le x \le 10) = P(x \le 10) - P(x \le 6)$$

Using a Normal Distribution to Approximate Binomial Probabilities

3. For each value of interest *a*, the correction for continuity is (a + .5), and the corresponding standard normal *z*-value is

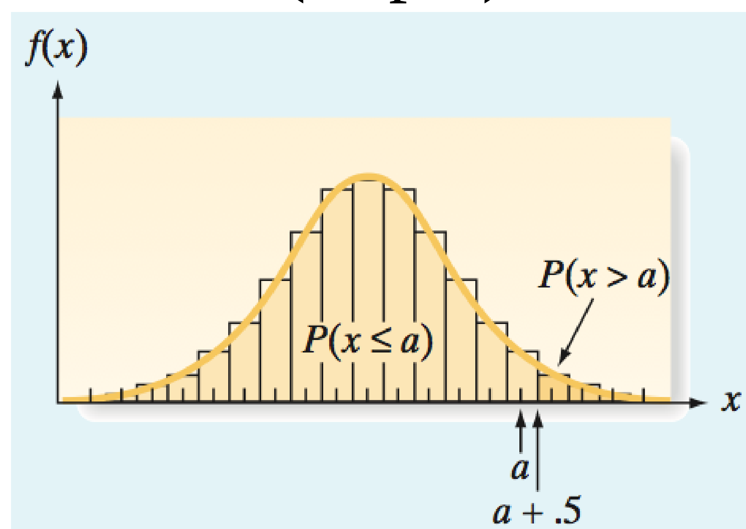
$$z = \frac{(a + .5) - \mu}{\sigma}$$



Using a Normal Distribution to Approximate Binomial Probabilities

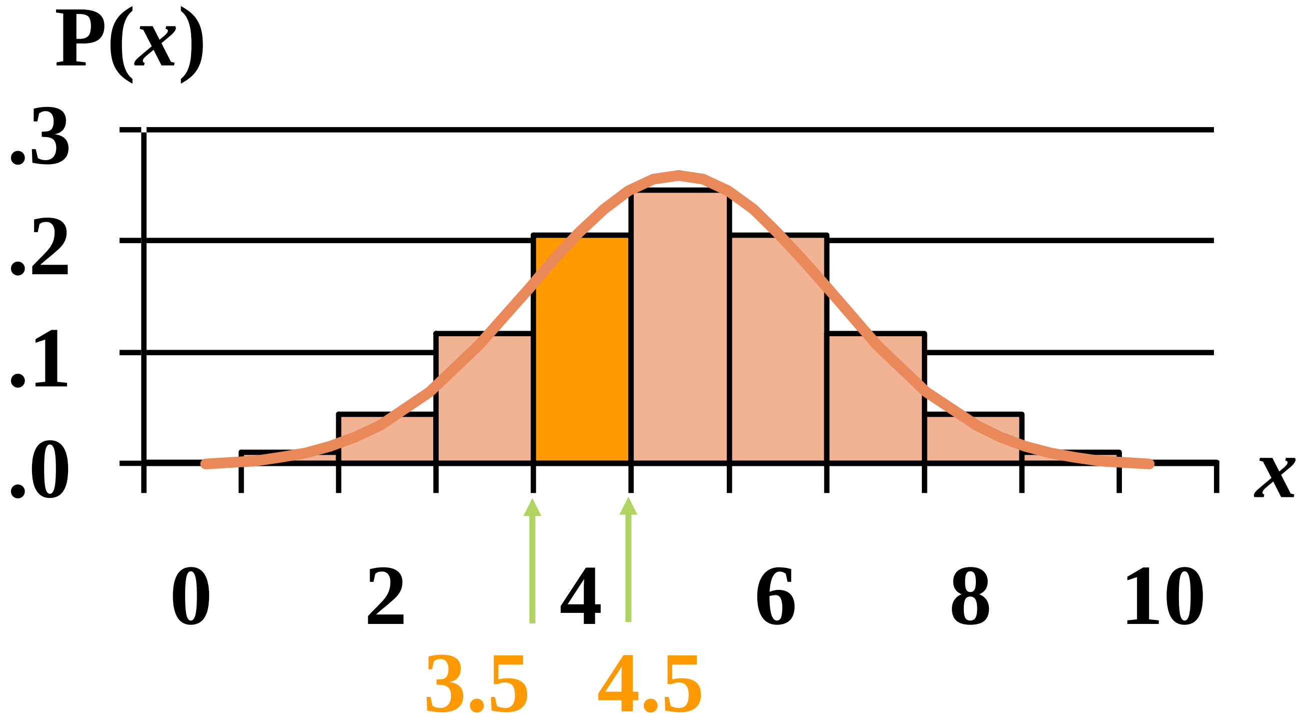
4. Sketch the approximating normal distribution and shade the area corresponding to the event of interest. Using z score table and the *z*-value (step 3),

To find the shaded area. This is the approximate probability of the binomial event.



Normal Approximation Example

What is the normal approximation of $p(x = 4)$ given $n = 10$, and $p = 0.5$?



Normal Approximation Solution

1. Calculate the interval:

$$np \pm 3\sqrt{np(1-p)} = 10(0.5) \pm 3\sqrt{10(0.5)(1-0.5)}$$
$$= 5 \pm 4.74 = (0.26,\ 9.74)$$

Interval lies in range 0 to 10, so normal approximation can be used

2. Express binomial probability in form:

$$P(x=4) = P(x \le 4) - P(x \le 3)$$

Normal Approximation Solution

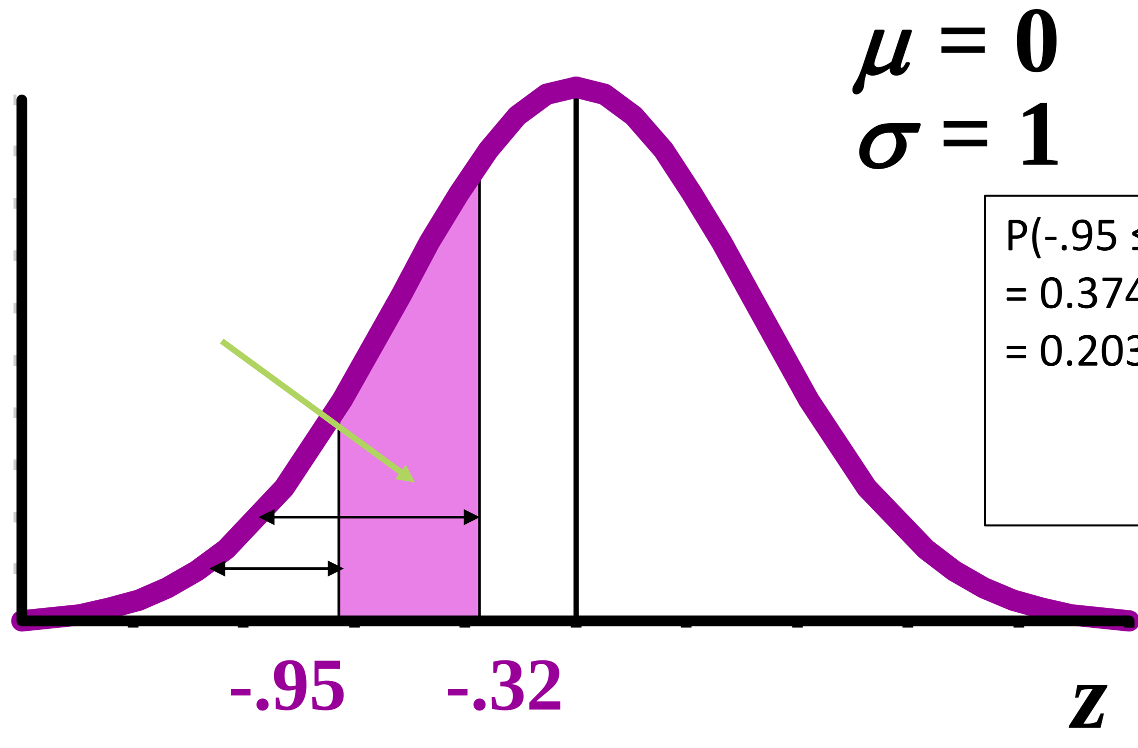
3. Compute standard normal z values:

$$z \approx \frac{(a+.5)-n\cdot p}{\sqrt{n\cdot p(1-p)}} = \frac{3.5-10(.5)}{\sqrt{10(.5)(1-.5)}} = -0.95$$

$$z \approx \frac{(a+.5)-n\cdot p}{\sqrt{n\cdot p(1-p)}} = \frac{4.5-10(.5)}{\sqrt{10(.5)(1-.5)}} = -0.32$$

Normal Approximation Solution

4. Sketch the approximate normal distribution:



P(-.95 ≤ z ≤ -.32)
= 0.3744-0.1710
= 0.2034

Normal Approximation Solution

5. The exact probability from the binomial formula is 0.2051 (versus .2034)