2 Normal Approximation to the Binomial

When dealing with a large number of trials in a Binomial situation, making direct calculations of the probabilities can be tedious

Example 3. Suppose that $\frac{1}{3}$ of computer chips manufactured by a certain company are defective. Suppose we randomly inspect $n = 36$ chips. What is the probability that in such a sample more than 13 chips will be defective?

$$X \sim Bin(n = 36, p = \frac{1}{3})$$

Find $P(X > 13)$,

Clearly this would take a long time...
$P(X > 13) = P(X = 14) + P(X = 15) + ... + P(X = 36)$
Shorter, but still tedious...

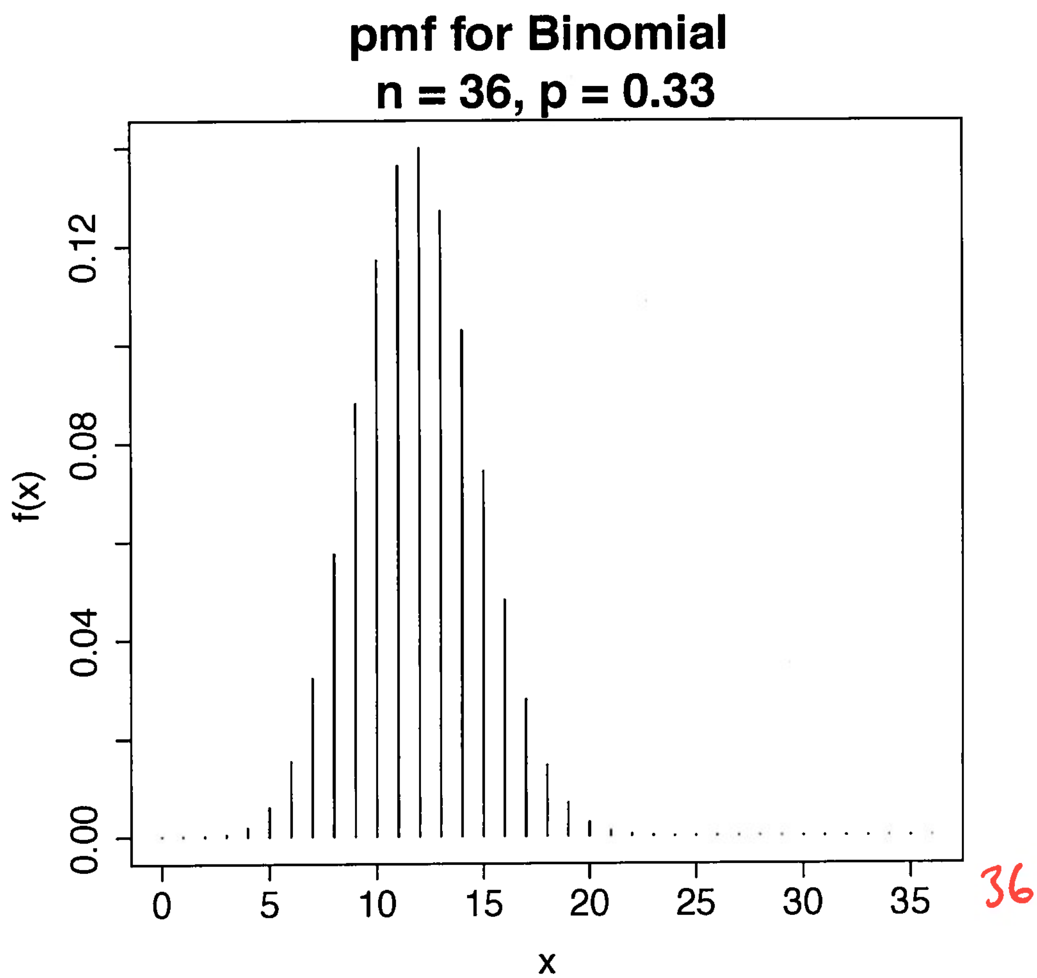
$P(X > 13) = 1 - P(X \leq 13)$
$= 1 - [P(X = 0) + P(X = 1) + ... + P(X = 13)]$

(1-p)

If $X \sim Bin(n, p)$ and if n is large so that $np \geq 5$, $nq \geq 5$, we can use the normal distribution to get an approximate answer

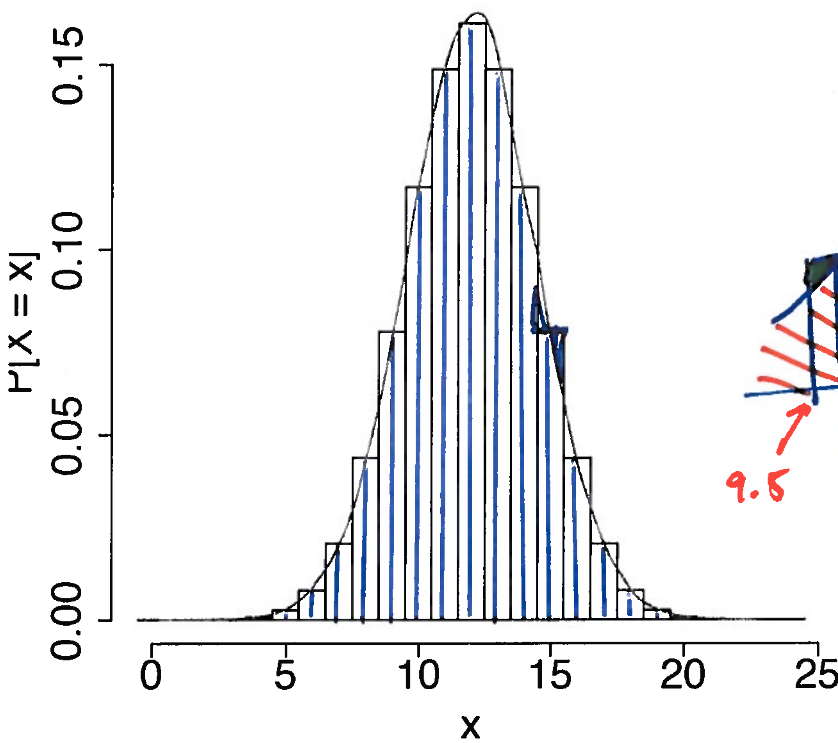
$$X \overset{approx}{\sim} N(np, np(1-p))$$

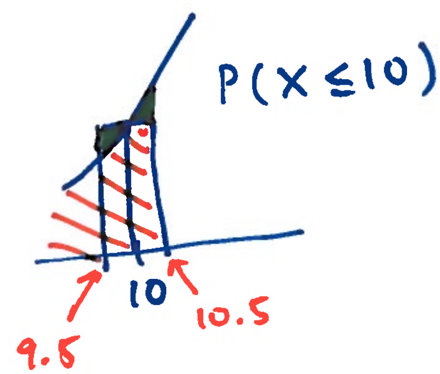
Idea: When n is large and $np \geq 5$ and $n(1-p) \geq 5$, the shape of the binomial distribution is approximately symmetrical.



Continuity Correction

Recall that a discrete random variable can take on only a specified number of values whereas a continuous random variable can take on any value within an interval. When using the normal model to approximate Binomial or Poisson distributions, we can make more accurate approximations if we use a continuity correction. With a continuous distribution (such as the normal) the probability of obtaining a particular value of a random variable is 0.

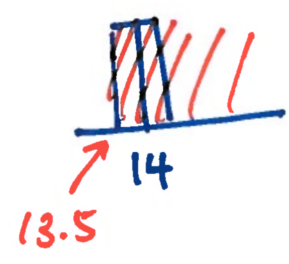


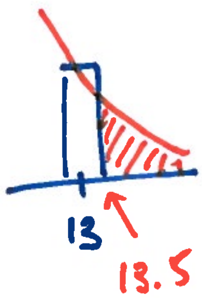


Steps when using continuity correction.

Example:

- To find: $P(X \leq 10)$, use $P(X \leq 10.5)$
- To find: $P(X \geq 14)$, use $P(X \geq 13.5)$





What about $P(X > 13)$?
First rewrite it using the equality sign, $P(X > 13) = P(X \geq 14) = P(\geq 13.5)$

$P(10 \leq X \leq 14)$? Use $P(9.5 \leq X \leq 14.5)$

Let's continue example 3, using the normal approximation
<u>Solution:</u>

$P(X = n)$ use $P(n - 0.5 \leq X \leq n + 0.5$

$P(X > n)$ use $P(X \geq n + 0.5)$

$P(X \leq n)$ use $P(X \leq n + 0.5)$

$P(X < n)$ use $P(X \leq n - 0.5)$

$P(X \geq n)$ use $P(X \geq n - 0.5)$

$X \sim Bin(36, p = \frac{1}{3})$

$P(X > 13) = ?$

$n \times p = 36 \times \frac{1}{3} = 12 \geq 5$

$n(1-p) = 36 \times \frac{2}{3} = 24 \geq 5$

$\therefore$ we can use the normal approx to Binomial dist.

$E(X) = np = 12$

$Var(X) = np(1-p)$
$= 36 \times \frac{1}{3} \times \frac{2}{3}$
$= 8$

approx.
$X \sim N(12, 8)$

$P(X > 13) = P(X \geq 13.5)$ (continuity correction)

$= 1 - P(X < 13.5)$

$= 1 - P\left(Z < \frac{13.5 - 12}{\sqrt{8}}\right)$

$= 1 - 0.7019 = 0.298$

3 Normal Approximation to the Poisson

If $X \sim pois(\lambda t)$ and λ is large ($\lambda t \geq 20$), we can use the normal distribution to approximate the Poisson distribution.

$$X \sim N(\lambda t, \lambda t)$$

Like the Binomial case, we are approximating a discrete distribution with a continuous distribution, thus we use the continuity correction.

Example 4. $X \sim Pois(25)$

(a) Find $P(X = 27)$

(b) Find $P(24 \leq X \leq 27)$

a) Actual calculation using poisson:

$$P(X = 27) = \frac{e^{-25}(25)^{27}}{27!} = 0.0708$$

Compare using normal approx.

$\lambda t = 25 \geqslant 20$

$X \sim N(25, 25)$

E(X) for poisson is 25 when $\lambda t = 25$

Var(X) = 25 for pois($\lambda t = 25$)

$P(X=27) = P(26.5 \le X \le 27.5)$ (apply continuity correction)

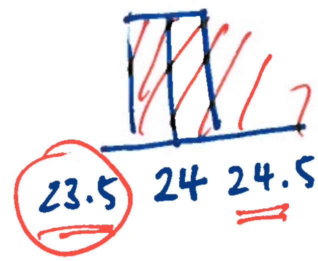
$$= P\left(\frac{26.5-25}{\sqrt{25}} \le Z \le \frac{27.5-25}{\sqrt{25}}\right)$$

$$= P(0.3 \le Z \le 0.5)$$

$$= 0.6915 - 0.6179 = 0.0736$$

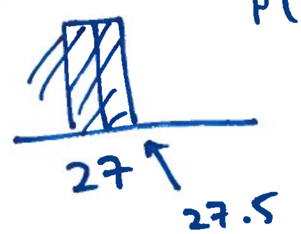
pretty close to exact.

b) Find $P(24 \le X \le 27)$



$$= P(23.5 \le X \le 27.5)$$

$$= P\left(\frac{23.5-25}{\sqrt{25}} \le Z \le \frac{27.5-25}{\sqrt{25}}\right)$$



$$= P(-0.3 \le Z \le 0.5)$$

$$= P(Z \le 0.5) - P(Z \le -0.3)$$

$$= P(Z \le 0.5) - [1 - P(Z \le 0.3)]$$

$$= 0.6915 - 0.3821 = 0.3094$$