Normal Approximation to the Binomial Distribution

The normal distribution

Starter:

A drinks machine is designed to dispense an average of 200 ml of liquid. The amount of liquid is normally distributed with a standard deviation of 15 ml.

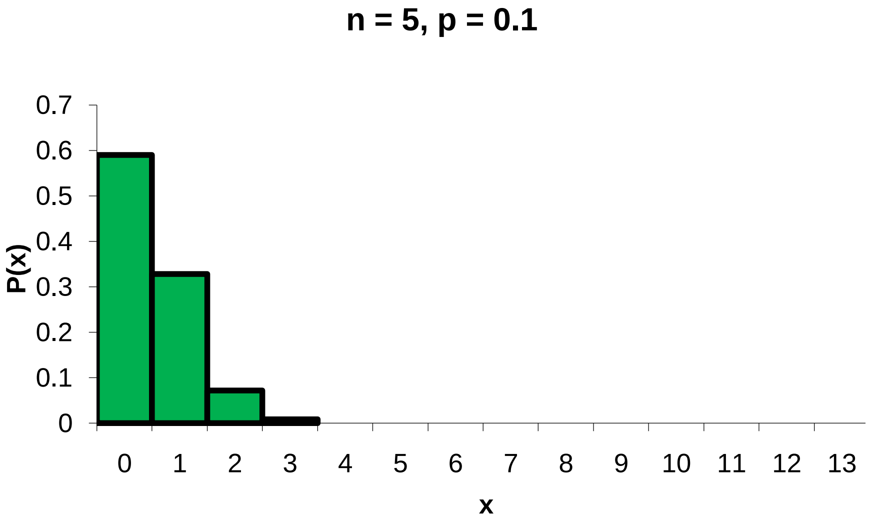
(i) If a 230 ml cup is used calculate the number of times the cup overflows, when 1000 drinks are dispensed.

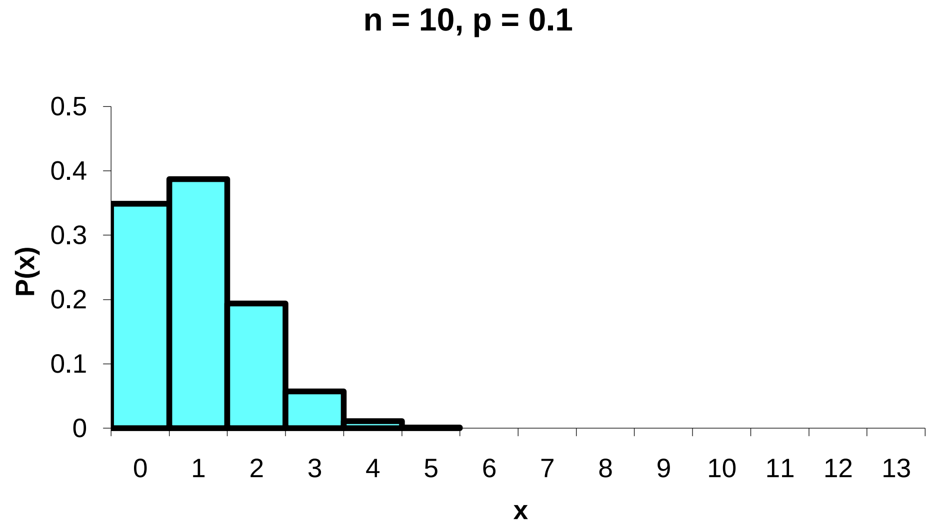
(ii) Below what value do we find the lowest 25% of the drinks?

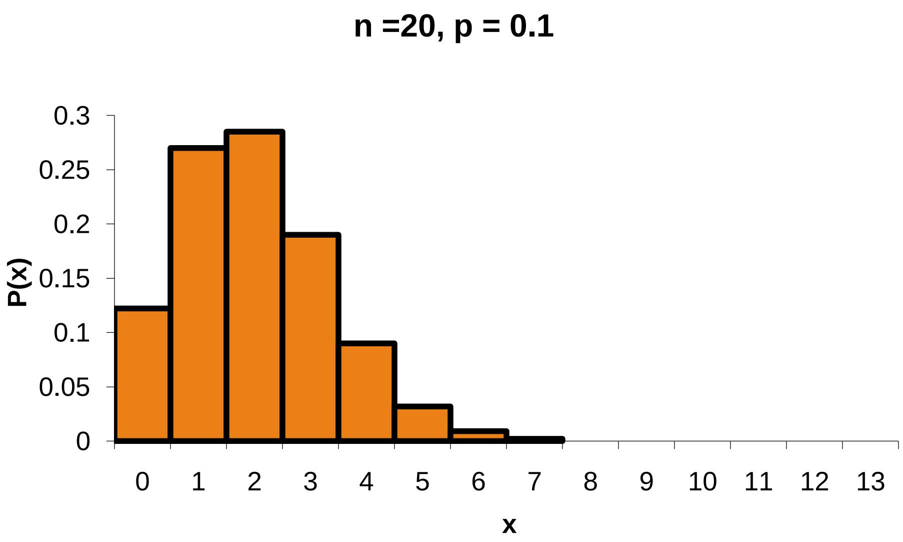
NORMAL APPROXIMATION TO BINOMIAL

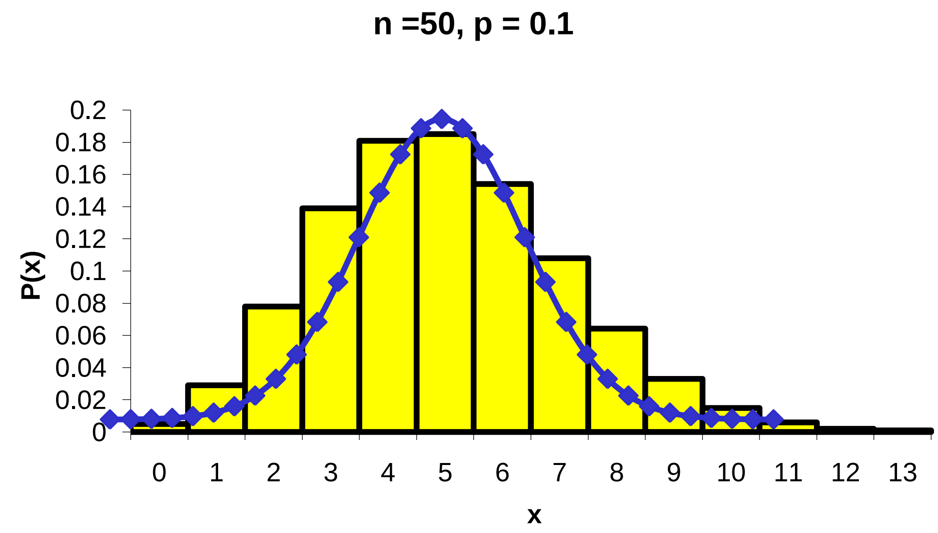
The following sketches illustrate the distributions of a Binomial distribution X with p = 0.1 for some values of n.

$$X \sim Bin(n, p = 0.1)$$



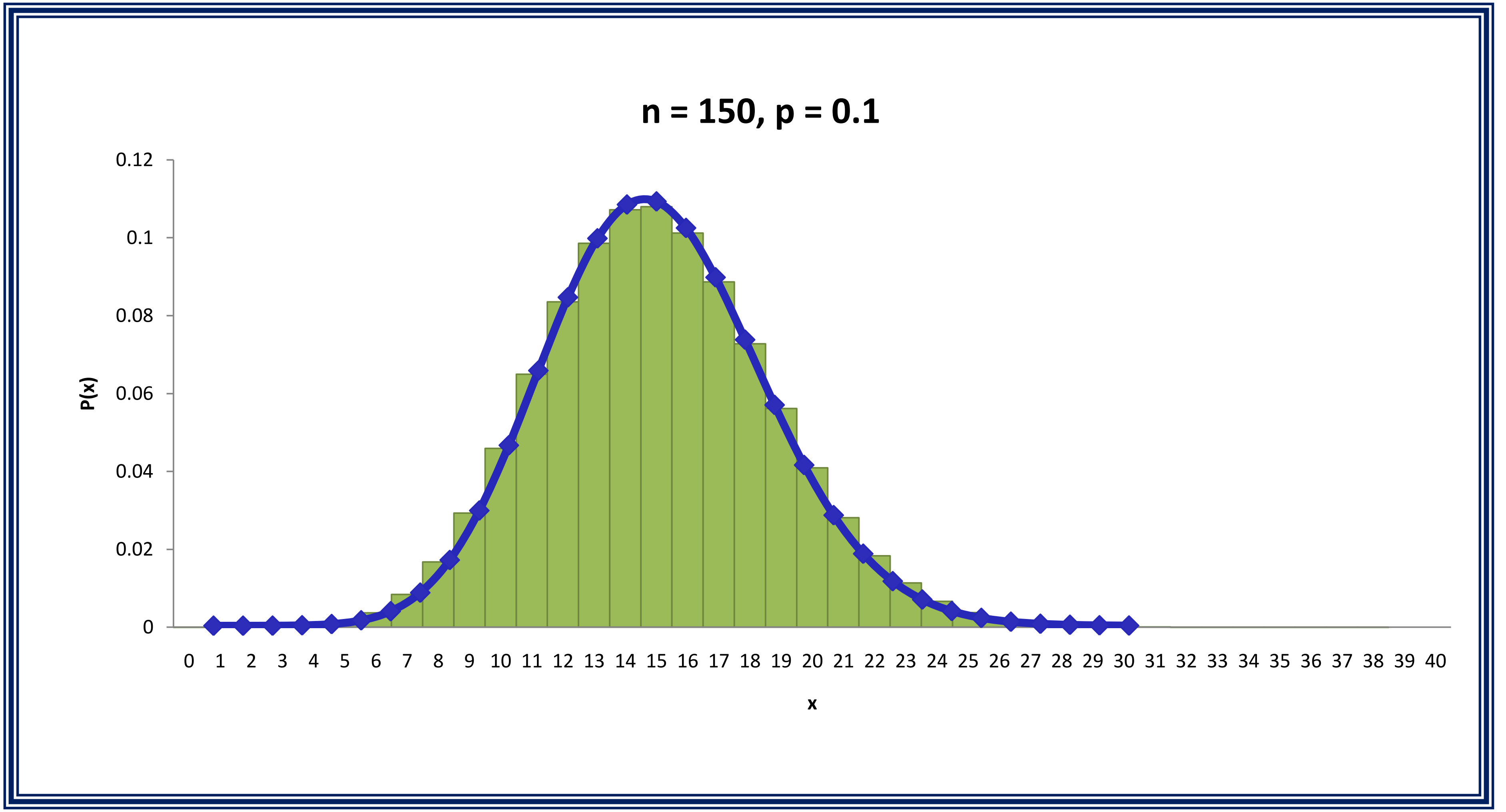






As **n** increases the binomial distribution becomes more symmetrical.

This is particularly clear when $n \rightarrow \infty$.


n = 150, p = 0.1
0.12
0.1
0.08
0.06
0.04
0.02
0
P(x)
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
x

Normal Approximations to Binomial

The **normal distribution** can be used as an approximation to the **binomial distribution** under certain circumstances.

If $X \sim B(n, p)$;

(i) When p is close to 0.5, the approximation will be good even if n is relatively small.

(ii) When p is not close to 0.5, the approximation will be good as long as $np > 5$ and $nq > 5$.

(iii) The larger the value of n the better the approximation.

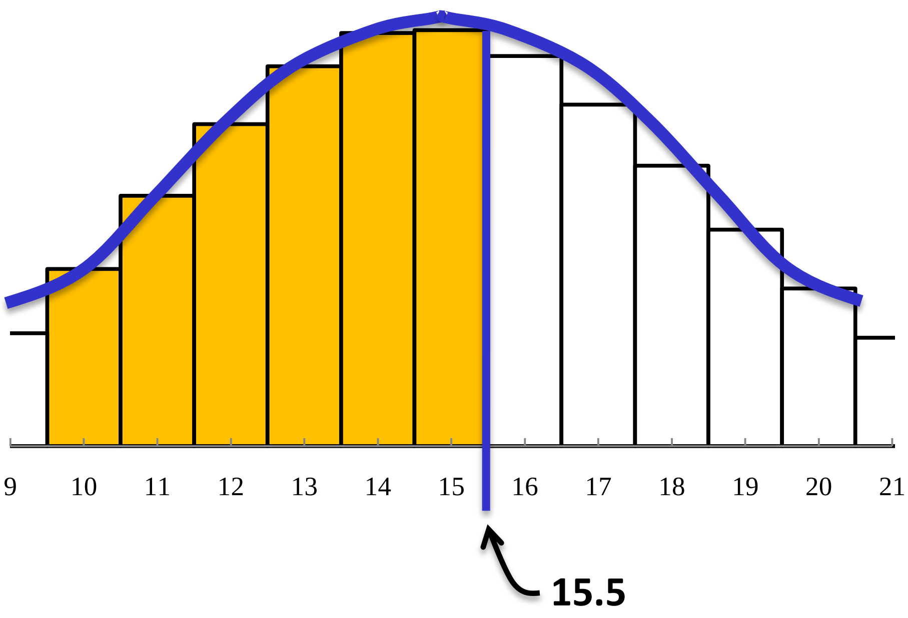
Continuity Correction

The binomial distribution is a discrete random variable, whereas the normal distribution is continuous.

We need to take this into account when we are using the normal distribution to approximate a binomial using a **continuity correction**.

In the discrete distribution, each probability is represented by a rectangle.

When converting probabilities, we want to include whole rectangles from the binomial into the normal, the continuity correction.



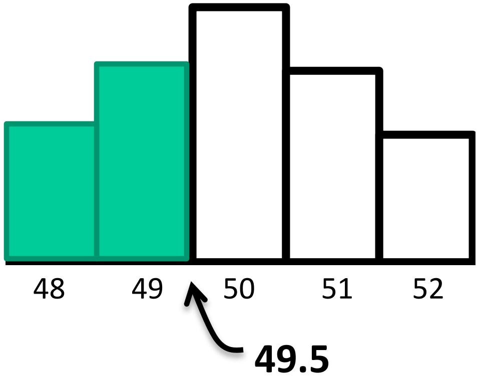
$P(x < 16)_{Bin} = P(x < 15.5)_{Norm}$

[The continuity correction]

Example:
The probability that a patient recovers from a blood disease is 0.6. If 100 people contracted the disease, what is the probability that less than a half recover.

X = Number of recoveries
$X \sim Bin(n = 100, p = 0.6)$ $\Rightarrow X \sim Norm\ (60, 24)_{approx}$
$\mu = np = 60,\ \sigma^2 = npq = 24$

$$\begin{aligned} P(X < 50)_{Bin} &= P(X < 49.5)_{Norm} \\ &= P(z < -2.14) \\ &= 1 - P(z < 2.14) \\ &= 1 - 0.98382 \\ &= 0.01618 \end{aligned}$$



$$Z = \frac{X - \mu}{\sigma} = \frac{49.5 - 60}{\sqrt{24}} = -2.14$$

Example:

A multiple choice paper has 80 questions, each with 4 possible answers of which only one is correct. What is the probability that by sheer guesswork a student will obtain between 25 and 30 correct answers inclusive.

X = Number of correct answers

$X \sim \text{Bin}(n = 80, p = 0.25)$

$\mu = np = 20$, $\sigma^2 = npq = 15$

$\Rightarrow X \sim \text{Norm}(20, 15)_{\text{approx}}$