# HYPERGEOMETRIC DISTRIBUTION

# Hypergeometric Distribution

A hypergeometric experiment possesses the following properties:

- A sample of a particular number of items is taken from a finite population without replacement.
- The population is divided into success or failure with a fixed number of successes.
- The result is the number of successes in the sample.
- The sample size is written as $n$, the population size as $N$, and the number of successes in the population as $k$.

People with O– blood are known as universal donors because their blood can be used in transfusions for people with any blood type without side effects. A researcher goes to a small school with 250 students. It is known from school records that 15 of the students have blood type O–. The researcher randomly selects a sample of 30 students and finds that 5 have O– blood.

a Is this an example of a hypergeometric experiment?

b List the values of the random variable $X$, the number of students who have O– blood type.

There is a finite population of 250.

A sample of 30 is taken without replacement.

The population is divided into two groups with 15 successes.

It is a hypergeometric experiment with $n = 30$, $N = 250$ and $k = 15$.

$X = 0, 1, 2, 3, \ldots, 15$

Hence $X$ represents the discrete random variable associated with this hypergeometric experiment.

A **hypergeometric distribution** is the probability distribution of a hypergeometric experiment. The random variable is the number of successes.

- There are $\binom{k}{x}$ ways of choosing $x$ successes from the $k$ successes in the population.
- There are $\binom{N-k}{n-x}$ ways of choosing $n - x$ failures from the $N - k$ failures in the population.
- Thus there are $\binom{k}{x}\binom{N-k}{n-x}$ ways of choosing a sample of $n$ with $k$ successes.
- There are $\binom{N}{n}$ ways of choosing a sample of $n$ from the population of $N$.

For the hypergeometric random variable $X$, $P(X = x) = \dfrac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}$ where $x \le n$ and$x \le k$.

$N =$ population size $k =$ number of successes in population $n =$ number of trials
$x =$ successes we are looking for.

Note: $E[X] = n\dfrac{k}{N}$

A box contains 50 electrical safety switches, of which 20% are known to be faulty. If a sample of 8 switches is selected at random and tested, calculate the probability that:

**a** two will be defective

**b** at least two will be defective.

Let $X$ = number of faulty switches

$$P(X=2)=\frac{\binom{10}{2}\binom{50-10}{8-2}}{\binom{50}{8}}$$

$$=\frac{\binom{10}{2}\binom{40}{6}}{\binom{50}{8}}$$

$$=0.321\ 724\ldots$$

The probability of two faulty switches is about 0.3217.

$$P(X<2)=P(X=0)+P(X=1)$$

$$=\frac{\binom{10}{0}\binom{40}{8}}{\binom{50}{8}}+\frac{\binom{10}{1}\binom{40}{7}}{\binom{50}{8}}$$

$$=0.490\ 502\ldots$$

$$P(X\geq 2)=1-0.490\ 502\ldots$$
$$=0.509\ 497\ldots$$

The probability of at least two faulty switches is about 0.5095

A computer manufacturer orders 500 printer circuit boards from a supplier. To determine whether or not to accept the order, the manufacturer randomly selects 15 circuit boards and tests them. The manufacturer will only accept the full order if there are no defective circuit boards in those that have been randomly selected for testing. Find each of the following probabilities.

a The probability of no faulty boards if 5% of the circuit boards are defective.

b The probability of no faulty boards if 10% of the circuit boards are defective.

Success = a defective circuit board

$N = 500,\ n = 15,\ k = 0.05 \times 500 = 25$

$$P(X = 0) = \frac{\binom{25}{0}\binom{500-25}{15-0}}{\binom{500}{15}}.$$

$$= 0.458\,096\ldots$$

When 5% of all boards are defective, the probability that none of the sample boards is defective is about 46%.

$N = 500,\ n = 15,\ k = 0.1 \times 500 = 50$

$$P(X = 0) = \frac{\binom{50}{0}\binom{500-50}{15-0}}{\binom{500}{15}}$$

$$= 0.201\,045\ldots$$

When 10% of all boards are defective, the probability that none of the sample boards is defective is about 20%.

Which of the following in not a characteristic of a hypergeometric probability experiment?

- A The population to be sampled must be finite.
- B For each trial of the experiment, there are only two possible outcomes.
- C A finite sample is randomly selected from the population.
- D The probability of success in each trial is constant.
- E The exact number of successes in the population is known.

D The probability of success in each trial is constant is not true because if one is already chosen and not replaced, it reduces the probability of it being selected again

Four cards are drawn from a well shuffled standard deck of playing cards. What is the probability two of the four cards are hearts?

An electrical appliance has 6 transistors. It is know that two of the transistors are faulty but it is not known which two. If three transistors are randomly selected what is the probability that one of the three is faulty?

It is known that 18% of a batch of 50 computer chips are defective. If 10 chips are selected for testing what is the probability that 4 of the chips are defective?

$$P(x=2) = \frac{{}^{39}C_2 \times {}^{13}C_2}{{}^{52}C_4} = 0.2135$$

$$P(x=4) = \frac{{}^{41}C_6 \times {}^{9}C_4}{{}^{50}C_{10}} = 0.0552$$

$$P(x=1) = \frac{{}^{4}C_2 \times {}^{2}C_1}{{}^{6}C_3} = 0.6$$

9 Six cards are drawn from a standard deck of playing cards without replacement. What is the probability of drawing two spades?

10 A batch of 100 circuit breakers contains 15% that are defective. If a sample of 19 circuit breakers is selected at random and tested, calculate the probability that:

a three will be defective

b at least three will be defective.

11 In a group of 200 people, 12 are known to suffer from colour blindness. A researcher randomly selects a sample of 20 people from the group and tests them.

a What is the probability that exactly three are colour blind?

b What is the probability that at least one is colour blind?

**9** $P(x = 2 \text{ spades}) = \dfrac{{}^{39}C_4 \times {}^{13}C_2}{{}^{52}C_6} = 0.3151$

**10 a** $P(\text{three will be defective}) = \dfrac{{}^{85}C_{16} \times {}^{15}C_3}{{}^{100}C_{19}} = 0.2705$

**b** $P(\text{at least three will be defective}) = 1 - P(0) - P(1) - P(2)$

$= 1 - 0.0321\ldots - 0.1367\ldots - 0.253$

$= 0.5777$

11 a $200 = 188 + 12,\ N = 200,\ k = 12,\ n = 20$

$P(\text{exactly three are colourblind}) = \dfrac{{}^{188}C_{17} \times {}^{12}C_3}{{}^{200}C_{20}} = 0.0833$

b $P(\text{at least one is colourblind}) = 1 - P(0)$

$= 1 - \dfrac{{}^{188}C_{20} \times {}^{12}C_0}{{}^{200}C_{20}} = 1 - 0.2718 = 0.7282$

In the game of Lotto, a player picks 6 numbers from the numbers 1, 2, 3, …, 45 and marks them on a card. This is called an entry. Six balls numbered 1 to 45 are then randomly selected by the Lotto organisers. Prizes are won by players who select 3, 4, 5 or 6 numbers that match the ones that have been randomly selected. The jackpot (major prize) is won if the player has 6 matching numbers.

a What is the probability that a single entry will win the jackpot?

b What is the probability that a single entry will win a prize?

**a** $P(\text{a single entry will win the jackpot}) = \dfrac{{}^{39}C_0 \times {}^{6}C_6}{{}^{45}C_6} = 0.000\,000\,1228$

**b** $P(\text{a single entry will win a prize}) = P(3) + P(4) + P(5) + P(6)$

$$= \frac{{}^{39}C_3 \times {}^{6}C_3}{{}^{45}C_6} + \frac{{}^{39}C_2 \times {}^{6}C_4}{{}^{45}C_6} + \frac{{}^{39}C_1 \times {}^{6}C_5}{{}^{45}C_6} + \frac{{}^{39}C_0 \times {}^{6}C_6}{{}^{45}C_6}$$

$= 0.022\,441 + 0.001\,364 + 0.000\,029 + 0.000\,000\,1$

$= 0.0238$

A manufacturer received an order of 250 machined components. The components are only acceptable if they are within ±1 mm of the specified dimensions. It is known that 12 of the components are not within ±1 mm of the required dimensions and are therefore unusable. The order will be rejected if there are four or more unusable components. The manufacturer decides to randomly select 20 components for testing to see if they are within the specified dimensions.

a What is the probability that no unusable components are found?

b What is the probability that three unusable components are found?

c What is the probability that the order will be rejected?

**a** $$P(\text{no unusable components are found}) = \frac{^{238}C_{20}}{^{250}C_{20}} = 0.3590$$

**b** $$P(\text{three unusable components are found}) = \frac{^{238}C_{17} \times {^{12}C_3}}{^{250}C_{20}} = 0.0507$$

**c** $$P(\text{the order will be rejected}) = P(x \geq 4)$$

$$= 1 - P(0) - P(1) - P(2) - P(3)$$

$$= 1 - 0.3598 - 0.3935 - 0.1869 - 0.0507$$

$$= 0.0098$$

In certain criminal trials, a jury is required to reach a unanimous verdict in order to convict an accused person. If a unanimous verdict is not reached, the jury is said to be 'hung' and the trial is abandoned. A particular jury consists of 12 people randomly selected from a pool of 50 potential jurors, of which 3 would never be willing to convict, regardless of the evidence presented at the trial. What is the probability that the trial will result in a hung jury, regardless of the evidence presented?

$$P(\text{the trial will result in a hung jury}) = P(x \geq 1)$$

$$= 1 - P(0)$$

$$= 1 - \frac{{}^{47}C_{12} \times {}^{3}C_{0}}{{}^{50}C_{12}}$$

$$= 1 - 0.4304$$

$$= 0.5696$$

# EXERCISE 2.05 (PAGE 70)

- Question 2, 3, 5, 6, 10