Hypergeometric Distribution : A Distribution of Dependent Events

Bindeshwar Singh Kushwaha

PosstNetwork Academy

Introduction

- In the previous sections, we studied distributions such as the binomial distribution.

Introduction

- In the previous sections, we studied distributions such as the binomial distribution.
- The binomial distribution assumes that each trial is independent and the probability of success remains constant.

Introduction

- In the previous sections, we studied distributions such as the binomial distribution.
- The binomial distribution assumes that each trial is independent and the probability of success remains constant.
- However, in many real-life problems, selections are made without replacement.

Introduction

- In the previous sections, we studied distributions such as the binomial distribution.
- The binomial distribution assumes that each trial is independent and the probability of success remains constant.
- However, in many real-life problems, selections are made without replacement.
- In such cases, the trials are not independent, and the hypergeometric distribution is used.

Real-life Scenario

- Suppose we have 10 tickets numbered 1 to 10.

Real-life Scenario

- Suppose we have 10 tickets numbered 1 to 10.
- We draw 3 tickets without replacement.

Real-life Scenario

- Suppose we have 10 tickets numbered 1 to 10.
- We draw 3 tickets without replacement.
- Let "success" be the event of drawing an odd-numbered ticket.

Real-life Scenario

- Suppose we have 10 tickets numbered 1 to 10.
- We draw 3 tickets without replacement.
- Let "success" be the event of drawing an odd-numbered ticket.
- Total odd numbers = 5 (1, 3, 5, 7, 9)

Real-life Scenario

- Suppose we have 10 tickets numbered 1 to 10.
- We draw 3 tickets without replacement.
- Let "success" be the event of drawing an odd-numbered ticket.
- Total odd numbers = 5 (1, 3, 5, 7, 9)
- If we replace the ticket each time, the trials are independent and we use the Binomial Distribution.

Real-life Scenario

- Suppose we have 10 tickets numbered 1 to 10.
- We draw 3 tickets without replacement.
- Let "success" be the event of drawing an odd-numbered ticket.
- Total odd numbers $= 5$ (1, 3, 5, 7, 9)
- If we replace the ticket each time, the trials are independent and we use the Binomial Distribution.
- If we do not replace the tickets, the trials are dependent and we use the Hypergeometric Distribution.

Basic Definition

- The hypergeometric distribution gives the probability of obtaining exactly x successes in n draws,

Basic Definition

- The hypergeometric distribution gives the probability of obtaining exactly x successes in n draws,
- From a finite population of size N containing M successes and $(N - M)$ failures,

Basic Definition

- The hypergeometric distribution gives the probability of obtaining exactly x successes in n draws,
- From a finite population of size N containing M successes and $(N - M)$ failures,
- When sampling is done without replacement.

$$P(X = x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$$

Parameters of Hypergeometric Distribution

- $\boldsymbol{N}$ = Total number of items in the population

Parameters of Hypergeometric Distribution

- N = Total number of items in the population
- M = Number of success items in the population

Parameters of Hypergeometric Distribution

- N = Total number of items in the population
- M = Number of success items in the population
- n = Number of draws (sample size)

Parameters of Hypergeometric Distribution

- $\boldsymbol{N}$ = Total number of items in the population
- $\boldsymbol{M}$ = Number of success items in the population
- $\boldsymbol{n}$ = Number of draws (sample size)
- $\boldsymbol{x}$ = Number of observed successes in the sample

Mean and Variance

- The mean or expected value is given by

$$E(X) = n\frac{M}{N}$$

- The mean or expected value is given by

$$E(X) = n\frac{M}{N}$$

- The variance is

$$Var(X) = n\frac{M}{N}\frac{N-M}{N}\frac{N-n}{N-1}$$

Mean and Variance

- The mean or expected value is given by

$$E(X) = n\frac{M}{N}$$

- The variance is

$$Var(X) = n\frac{M}{N}\frac{N-M}{N}\frac{N-n}{N-1}$$

- The term $\frac{N-n}{N-1}$ is called the finite population correction factor.

Example 1: Ticket Problem

- 10 tickets are numbered from 1 to 10.

Example 1: Ticket Problem

- 10 tickets are numbered from 1 to 10.
- 5 tickets have odd numbers (successes) and 5 have even numbers (failures).

Example 1: Ticket Problem

- 10 tickets are numbered from 1 to 10.
- 5 tickets have odd numbers (successes) and 5 have even numbers (failures).
- 3 tickets are drawn without replacement.

Example 1: Ticket Problem

- 10 tickets are numbered from 1 to 10.
- 5 tickets have odd numbers (successes) and 5 have even numbers (failures).
- 3 tickets are drawn without replacement.
- Find the probability that exactly 2 tickets have odd numbers.

$$P(X = 2) = \frac{\binom{5}{2}\binom{5}{1}}{\binom{10}{3}} = \frac{10 \times 5}{120} = \frac{1}{2.4} \approx 0.4167$$

Example 2: Jury Selection Problem

- A jury of 12 members is chosen from a pool of 20 people.

Example 2: Jury Selection Problem

- A jury of 12 members is chosen from a pool of 20 people.
- 8 are men (successes) and 12 are women (failures).

Example 2: Jury Selection Problem

- A jury of 12 members is chosen from a pool of 20 people.
- 8 are men (successes) and 12 are women (failures).
- Find the probability that the jury contains exactly 5 men.

$$P(X = 5) = \frac{\binom{8}{5}\binom{12}{7}}{\binom{20}{12}}$$

Example 3: Fish Tank Problem

- A tank contains 200 fish. 60 are tagged and 140 are untagged.

Example 3: Fish Tank Problem

- A tank contains 200 fish. 60 are tagged and 140 are untagged.
- A sample of 10 fish is drawn without replacement.

Example 3: Fish Tank Problem

- A tank contains 200 fish. 60 are tagged and 140 are untagged.
- A sample of 10 fish is drawn without replacement.
- Find the probability that exactly 4 tagged fish are drawn.

$$P(X = 4) = \frac{\binom{60}{4}\binom{140}{6}}{\binom{200}{10}}$$

Probability of Acceptance Without Further Inspection

- A lot of 25 units contains 10 defective units.

Probability of Acceptance Without Further Inspection

- A lot of 25 units contains 10 defective units.
- An engineer inspects 2 randomly selected units from the lot.

Probability of Acceptance Without Further Inspection

- A lot of 25 units contains 10 defective units.
- An engineer inspects 2 randomly selected units from the lot.
- The lot is accepted if both selected units are non-defective.

Probability of Acceptance Without Further Inspection

- A lot of 25 units contains 10 defective units.
- An engineer inspects 2 randomly selected units from the lot.
- The lot is accepted if both selected units are non-defective.
- We are to find the probability that the lot is accepted without further inspection.

Solution

- Total units, $N = 25$, defective units $M = 10$, and sample size $n = 2$.

Solution

- Total units, $N = 25$, defective units $M = 10$, and sample size $n = 2$.
- Desired probability:

$$P(\text{no defective units in sample}) = \frac{\binom{15}{2}}{\binom{25}{2}}$$

Solution

- Total units, $N = 25$, defective units $M = 10$, and sample size $n = 2$.
- Desired probability:

$$P(\text{no defective units in sample}) = \frac{\binom{15}{2}}{\binom{25}{2}}$$

- Calculate the values:

$$\frac{15 \times 14}{25 \times 24} = \frac{210}{600} = 0.35$$

Solution

- Total units, $N = 25$, defective units $M = 10$, and sample size $n = 2$.
- Desired probability:

$$P(\text{no defective units in sample}) = \frac{\binom{15}{2}}{\binom{25}{2}}$$

- Calculate the values:

$$\frac{15 \times 14}{25 \times 24} = \frac{210}{600} = 0.35$$

- **Answer:** The probability that the lot is accepted without further inspection is $\boxed{0.35}$.

Important Properties

- Trials are dependent since sampling is without replacement.

Important Properties

- Trials are dependent since sampling is without replacement.
- The total population N is finite.

Important Properties

- Trials are dependent since sampling is without replacement.
- The total population N is finite.
- The random variable X can take integer values within

$$\max(0, n - (N - M)) \leq X \leq \min(n, M)$$

Important Properties

- Trials are dependent since sampling is without replacement.
- The total population N is finite.
- The random variable X can take integer values within

$$\max(0, n - (N - M)) \leq X \leq \min(n, M)$$

- The sum of all probabilities is 1:

$$\sum_{x} P(X = x) = 1$$

Comparison with Binomial Distribution

- The Binomial Distribution assumes independent trials with constant probability $\boldsymbol{p}$.

Comparison with Binomial Distribution

- The Binomial Distribution assumes independent trials with constant probability $\boldsymbol{p}$.
- The Hypergeometric Distribution assumes dependent trials without replacement.

Comparison with Binomial Distribution

- The Binomial Distribution assumes independent trials with constant probability p.
- The Hypergeometric Distribution assumes dependent trials without replacement.
- When N is large and n is small, the hypergeometric distribution approximates the binomial distribution.

Summary

- Used when sampling is done without replacement.

Summary

- Used when sampling is done without replacement.
- Suitable for finite populations.

Summary

- Used when sampling is done without replacement.
- Suitable for finite populations.
- Depends on $\boldsymbol{N}$, $\boldsymbol{M}$, and $\boldsymbol{n}$.

Summary

- Used when sampling is done without replacement.
- Suitable for finite populations.
- Depends on $\boldsymbol{N}$, $\boldsymbol{M}$, and $\boldsymbol{n}$.
- Probability formula:

$$P(X = x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$$

Summary

- Used when sampling is done without replacement.
- Suitable for finite populations.
- Depends on $\boldsymbol{N}$, $\boldsymbol{M}$, and $\boldsymbol{n}$.
- Probability formula:

$$\boldsymbol{P(X = x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}}$$

- Mean and variance summarize the distribution's behavior.

Reach PostNetwork Academy

Website

www.postnetwork.co

Reach PostNetwork Academy

Website

www.postnetwork.co

YouTube Channel

www.youtube.com/@postnetworkacademy

Reach PostNetwork Academy

Website

www.postnetwork.co

YouTube Channel

www.youtube.com/@postnetworkacademy

Facebook Page

www.facebook.com/postnetworkacademy

Reach PostNetwork Academy

Website

www.postnetwork.co

YouTube Channel

www.youtube.com/@postnetworkacademy

Facebook Page

www.facebook.com/postnetworkacademy

LinkedIn Page

www.linkedin.com/company/postnetworkacademy

Reach PostNetwork Academy

Website

www.postnetwork.co

YouTube Channel

www.youtube.com/@postnetworkacademy

Facebook Page

www.facebook.com/postnetworkacademy

LinkedIn Page

www.linkedin.com/company/postnetworkacademy

GitHub Repositories

www.github.com/postnetworkacademy

Thank You!