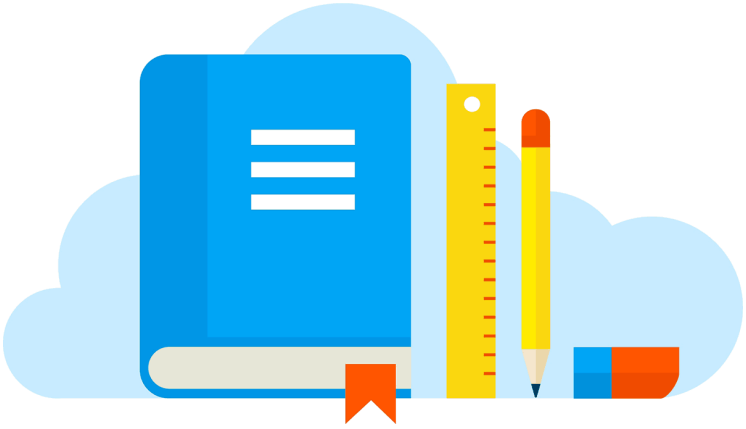


Data Science Notes by Sarowar Ahmed

📊 Chapter: Probability Theory

📊 Topic: Hypergeometric Distribution

🔎 Hello, GitHub family! Today, we're diving into the fascinating world of hypergeometric distribution. This concept might sound complex, but I'll explain it in a way that's easy to understand for everyone. Whether you're a student, professional, or just curious about probability, this post will help you grasp the hypergeometric distribution with real-life examples and a neat visualization!

🎯 What is Hypergeometric Distribution?

▪ Imagine you have a box filled with colored balls: some are red, and some are blue. You want to know the probability of picking a specific

number of red balls if you draw a few balls without putting them back in the box. This scenario, where you sample without replacement, is where the hypergeometric distribution comes into play!

📐 Formula of Hypergeometric Distribution:

▪ The probability of drawing exactly k successes (red balls) in n draws, from a finite population of size N containing exactly K successes, is given by:

$P(X=k) = (K \text{ C } k)\ (N-K \text{ C } n-k) / (N \text{ C } n)$

🌍 Real-Life Example:

▪ Scenario: Suppose a box contains 20 balls, 8 of which are red and 12 are blue. You draw 5 balls randomly without putting them back. What's the probability of drawing exactly 3 red balls?
▪ Using the Formula:

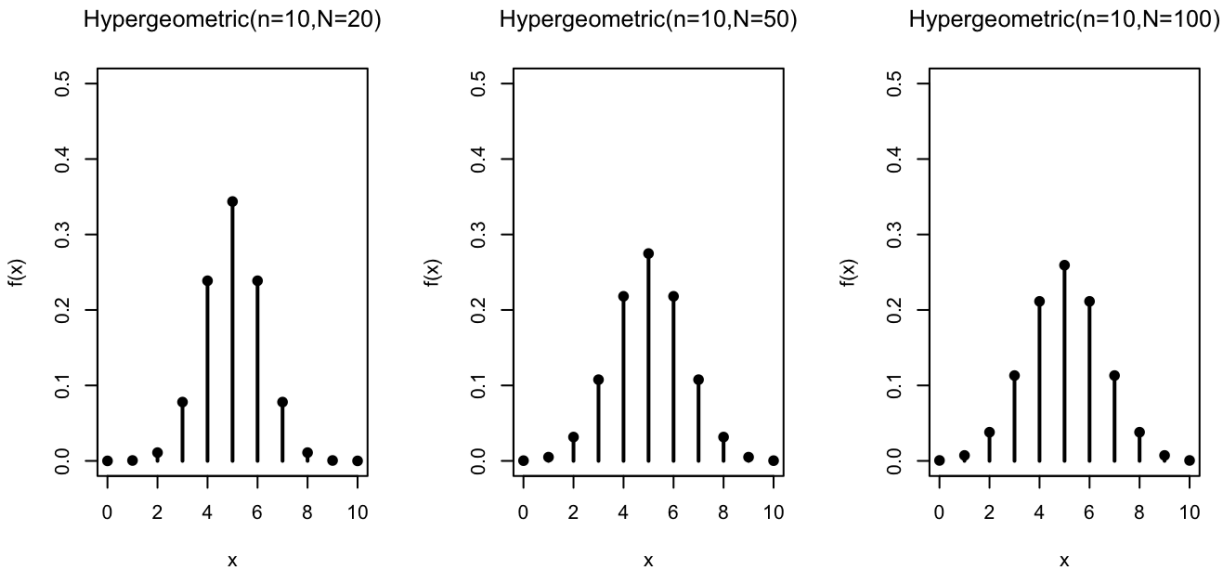
N=20 (total number of balls)
K=8 (total number of red balls)
n=5 (number of balls drawn)
k=3 (number of red balls to find)

$P(X=3) = (8 \text{ C } 3)(12 \text{ C } 2)/(20 \text{ C } 5)$

$P(X=3) = 56\times66/15504\approx0.239$

So, there's approximately a 23.9% chance of drawing exactly 3 red balls.

🖼️ Visualizing the Hypergeometric Distribution:



▪ This visual aid helps to illustrate the sequential process of updating beliefs based on new evidence, making it easier to grasp the concept.

🚀 Why Does This Matter?

▪ Understanding the hypergeometric distribution is crucial in many fields, including quality control, ecological studies, and even card games, where the outcome is based on sampling without replacement. It helps us calculate probabilities in scenarios where each draw significantly affects the outcomes of the subsequent draws.

Got any questions about Hypergeometric Distribution!? Feel free to ask me via Linkedin! Let's keep learning together.

My LinkedIn Date: 19/04/2024