



**Data Science
Academy**

www.datascienceacademy.com.br

Business Analytics

Distribuição Normal e Inferência Estatística



O Teorema Central do Limite é um importante resultado da estatística e a demonstração de muitos outros teoremas estatísticos dependem dele. Em teoria das probabilidades, esse teorema afirma que quando o tamanho da amostra aumenta, a distribuição amostral da sua média aproxima-se cada vez mais de uma distribuição normal. Este resultado é fundamental na teoria da inferência estatística.

Na inferência estatística a utilidade do teorema central do limite vai desde estimar os parâmetros como a média populacional ou o desvio padrão da média populacional, a partir de uma amostra aleatória dessa população, ou seja, da média amostral e do desvio padrão da média amostral até calcular a probabilidade de um parâmetro ocorrer dado um intervalo, sua média amostral e o desvio padrão da média amostral.

O teorema central do limite afirma que a média de uma amostra de n elementos de uma população tende a uma distribuição normal. Pode-se pensar de forma empírica que ao nos distanciarmos da média, a probabilidade de ocorrência diminui, ou seja, é mais provável ocorrer um evento que se encontra próximo da média do que um evento de um dos extremos. Além disso, uma distribuição pode ganhar a forma de curva normal se possuir diferentes combinações para cada resultado possível do espaço amostral. Isso é válido (em se tratando de amostras discretas), para amostras suficientemente grandes da população. O suficientemente grande, varia de acordo com a população, para populações com distribuição quase simétrica, a amostra pode ser menor do que para populações cuja distribuição seja assimétrica. A curva normal obtida pode então ser convertida em uma curva binomial ou em uma curva de Poisson, e posteriormente pode-se ainda realizar uma correção de continuidade. A precisão da correção de continuidade também pode ser medida.

Assim, é permitido inferir sobre a população através da média amostral e do desvio padrão amostral. Se extraíssemos todos os elementos da população, os dados sobre a amostra seriam exatamente iguais aos da população, mas isso pode ser demasiadamente custoso e/ou lento e/ou impossível (é impossível medir a resistência máxima de qualquer produto para todos os elementos da população).